Universidad del Valle de Guatemala

FACULTAD DE CIENCIAS Y HUMANIDADES

DEPARTAMENTO DE MATEMÁTICA

BASES DE GRÖBNER Y SU APLICACIÓN A LA DEMOSTRACIÓN AUTOMATIZADA DE TEOREMAS GEOMÉTRICOS: UNA PRUEBA DEL TEOREMA DEL HEXAGRAMA DE PASCAL

TRABAJO DE GRADUACIÓN PRESENTADO POR

FRANCISCO JOSÉ MARTÍNEZ FIGUEROA

PARA OPTAR AL GRADO

DE LICENCIADO EN MATEMÁTICA

GUATEMALA

2015

BASES DE GRÖBNER Y SU APLICACIÓN A LA DEMOSTRACIÓN AUTOMATIZADA DE TEOREMAS GEOMÉTRICOS: UNA PRUEBA DEL TEOREMA DEL HEXAGRAMA DE PASCAL

Universidad del Valle de Guatemala

FACULTAD DE CIENCIAS Y HUMANIDADES

DEPARTAMENTO DE MATEMÁTICA

BASES DE GRÖBNER Y SU APLICACIÓN A LA DEMOSTRACIÓN AUTOMATIZADA DE TEOREMAS GEOMÉTRICOS: UNA PRUEBA DEL TEOREMA DEL HEXAGRAMA DE PASCAL

TRABAJO DE GRADUACIÓN PRESENTADO POR

FRANCISCO JOSÉ MARTÍNEZ FIGUEROA

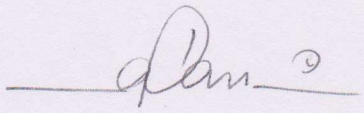
PARA OPTAR AL GRADO

DE LICENCIADO EN MATEMÁTICA

GUATEMALA

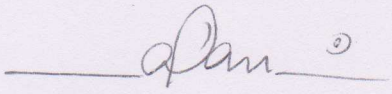
2015

Vo.Bo.:

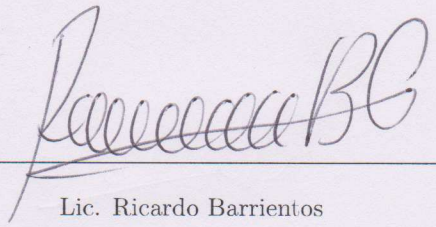


MSc. Alan Reyes

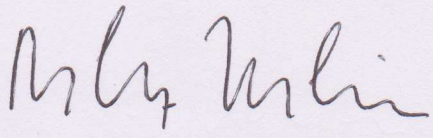
Tribunal Examinador:



MSc. Alan Reyes



Lic. Ricardo Barrientos



Dr. Roberto Molina

Fecha de aprobación: Guatemala, 31 de julio de 2015

Índice general

Índice de figuras

Índice de tablas

Resumen

Las bases de Gröbner y el algoritmo para su cálculo, introducidos por B. Buchberger en la década de los 60's, provocó una revolución en la geometría algebraica. Los nuevos métodos permitieron el uso de las computadoras como herramienta para calcular bases de ideales de polinomios, dando así solución a innumerables problemas asociados. Parte de la popularidad de la nueva teoría, se debe a su utilidad en una amplia gama de problemas: desde la resolución de sistemas de ecuaciones polinomiales, hasta la aplicación menos evidente en la demostración automatizada de teoremas geométricos. En este trabajo, se presenta la teoría básica de las bases de Gröbner y su aplicación a la demostración automatizada de teoremas geométricos. El método mostrado se ejemplifica con una demostración al caso general del teorema del hexagrama de Pascal. Se prueba también un caso degenerado de dicho teorema: el teorema de Pappus, aunque el resto de casos degenerados no son examinados a fondo. Para un próximo trabajo de investigación, se recomienda examinar más de cerca los casos degenerados del teorema del hexagrama de Pascal, y determinar si es posible aplicar el mismo método para su demostración.

1 INTRODUCCIÓN

Las bases de Gröbner fueron introducidas en 1965 por Bruno Buchberger en su tesis doctoral. En dicho trabajo, Buchberger desarrolló las ideas de su asesor, Wolfgan Gröbner, para resolver sistemas de ecuaciones polinomiales en varias variables. El trabajo de Buchberger no sólo formalizó estas ideas, sino que demostró que el objeto fundamental de su tesis, las bases de Gröbner, se pueden calcular para cualquier ideal de polinomios en varias variables, proporcionando además un algoritmo para su cálculo.

Desde entonces, las bases de Gröbner han sido ampliamente estudiadas, relacionándose generalmente con la geometría algebraica y temas afines. Las aplicaciones de las mismas son innumerables, comenzando por muchos de los problemas en matemática y matemática aplicada que llevan al planteamiento de sistemas de ecuaciones polinomiales.

En este trabajo se presenta la teoría básica de Bases de Gröbner y el algoritmo de Buchberger, y su relación con algunos conceptos de geometría algebraica. Con ayuda de la misma, muchos temas de geometría algebraica y álgebra conmutativa que antes se estudiaban únicamente en cursos de posgrado, se incluyen de manera asequible en este texto.

El texto comienza recopilando los resultados más importantes de anillos, campos e ideales que se utilizan a lo largo del trabajo en el Capítulo 2. De esta forma, basta con que el lector este familiarizado con un poco de álgebra lineal para poder comprender el desarrollo de este trabajo. Aunque no se demuestran todos los resultados de este capítulo, se incluyen en las referencias las fuentes que se pueden consultar.

En el Capítulo 3, se construye y estudia el anillo de polinomios de una variable, con coeficientes en un campo. El lector notará que para resolver un sistema de ecuaciones de polinomios en una variable, basta con obtener el máximo común divisor de los polinomios con el algoritmo de Euclides y luego utilizar el teorema fundamental del álgebra para garantizar la existencia de soluciones del mismo. En este capítulo se muestra este desarrollo, comenzando por mostrar el algoritmo de la división que es indispensable para poder utilizar el algoritmo de Euclides.

El interés por desarrollar la base de Gröbner de un ideal se hace evidente en el Capítulo 4, en el cuál se introducen los anillos de polinomios en varias variables. Al tratar de crear un algoritmo de la división para varias variables, se evidencia que primero es necesario introducir un orden para los monomios, sin embargo esto no es suficiente. Con un orden de monomios fijo, es posible tener un algoritmo de reducción, similar al de la división para una variable, pero éste falla en la unicidad del residuo que sí tiene el algoritmo de la división.

Así, en el Capítulo 5, se introducen las bases de Gröbner buscando tener residuos, o formas normales, únicas para poder trabajar en el anillo cociente $\mathbb{K}[x_1, \cdots, x_n]/I$, con I un ideal.

En el Capítulo 6 se introducen los S-polinomios y el criterio de Buchberger, el cual tiene como consecuencia inmediata el algoritmo de Buchberger. Se discuten también un segundo y tercer criterios de Buchberger, que permiten mejorar el algoritmo en sentido de recursos computacionales, pues permiten evitar algunas operaciones redundantes. Se discute también brevemente la complejidad computacional del algoritmo que, aunque resulta ser muy alta, para la mayoría de ejemplos en las aplicaciones es lo suficientemente manejable.

El Capítulo 7 introduce algunos conceptos básicos de geometría algebraica, mostrando así el poder de las bases de Gröbner al demostrar el importante teorema de *Nullstellensatz* o teorema de los ceros de Hilbert, el cual permite generalizar el teorema fundamental del álgebra al caso de varias variables. Como una de las aplicaciones más importantes de las bases de Gröbner, se finaliza el capítulo mostrando su utilización para la resolución de sistemas de ecuaciones polinomiales en varias variables, cuando dichos sistemas tienen un número finito de soluciones.

Finalmente, una de las aplicaciones más interesantes de esta teoría es la de la demostración automatizada de teoremas de geometría euclideana. Este trabajo finaliza con dicha aplicación en el Capítulo 8. Así, se recopila la metodología descrita por Cox, Little, y O'Shea (2007) y se aplica para demostrar los casos *genéricos* del teorema del hexagrama de Pascal, así como su caso degenerado, el teorema de Pappus. A su vez, se incluye anexo el código para implementarlo en el paquete computacional SAGE (Stein *et al.*, 2015), así como los resultados obtenidos del mismo.

2 ANILLOS, CAMPOS E IDEALES

En este capítulo se presentan los conceptos y resultados más importantes de la teoría general de anillos, campos e ideales que serán de utilidad en los próximos capítulos. Para una presentación más detallada y las demostraciones omitidas, refiérase a Herstein (1975), entre otros.

2.1. Anillos

Definición 2.1 (Anillo)**.** *Dado un conjunto no vacío R y dos operaciones binarias $+, \cdot : R \times R \to R$, se dice que $(R, +, \cdot)$ (o simplemente R cuando no hay confusión sobre $+$ y $\cdot$)* ***es un anillo*** *si $\forall\ a, b, c \in R$ se cumplen:*

1. *$a + b \in R$ (Cerradura de la suma)*
2. *$a + b = b + a$ (Conmutatividad de la suma)*
3. *$(a + b) + c = a + (b + c)$ (Asociatividad de la suma)*
4. *Existe un elemento 0 en R tal que $a + 0 = a\ \forall a \in R$ (Neutro aditivo)*
5. *Para cada $a \in R$ existe $-a \in R$ tal que $a + (-a) = (-a) + a = 0$ (Inversos aditivos)*
6. *$a \cdot b \in R$ (Cerradura del producto)*
7. *$a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (Asociatividad del producto)*
8. *$a \cdot (b + c) = a \cdot b + a \cdot c$ y $(b + c) \cdot a = b \cdot a + c \cdot a$ (Leyes distributivas)*

Observación. A las operaciones $+$ y $\cdot$ se les denomina **suma** y **producto**, respectivamente. Es usual no escribir el símbolo de producto entre dos variables, entendiendo su yuxtaposición como el producto: $ab = a \cdot b$.

Definición 2.2. *Sea R, un anillo.*

- *Si $\forall a, b \in R\ ab = ba$ (i.e. su producto es conmutativo), se dice que R* ***es un anillo conmutativo****.*
- *Si existe un elemento $1 \in R$ tal que $1 \cdot a = a \cdot 1 = a\ \forall a \in R$, se dice que R* ***es un anillo con identidad****.*

***Ejemplo* 2.1**

Algunos anillos importantes:

1. $R = \mathbb{Z}$ el conjunto de números los enteros, $Z = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$ con la suma y multiplicación usuales, es un *anillo conmutativo con identidad*.

2. $R = 2\mathbb{Z}$ el conjunto de enteros pares, $2\mathbb{Z} = \{\cdots, -4, -2, 0, 2, 4, \cdots\}$ con la suma y multiplicación usuales, es un *anillo conmutativo*.

3. $R = \mathbb{Z}/n\mathbb{Z}$ el conjunto de enteros módulo n, para $n \in \mathbb{Z}^+, n \geq 2$, $\mathbb{Z}/n\mathbb{Z} = \{0, 1, 2, \cdots, n-1\}$ donde la suma y multiplicación devuelven el residuo de dividir entre n el resultado de la suma o multiplicación usual de $\mathbb{Z}$, es un *anillo conmutativo con identidad*.

4. $R = \mathbb{R}[x]$ el conjunto de polinomios con coeficientes en $\mathbb{R}$ con la suma y multiplicación usuales es un *anillo conmutativo con identidad*. El estudio de este anillo constituye el tema del próximo capítulo.

Proposición 2.3. Si R es un anillo, entonces para todo $a, b \in R$ se cumple:

1. $a \cdot 0 = 0 \cdot a = 0$

2. $a(-b) = (-a)b = -(ab)$

3. $(-a)(-b) = ab$

Si además R es un anillo con identidad, entonces:

4. $(-1)a = -a$

5. $(-1)(-1) = 1$

Demostración. Se omite la prueba de este resultado por ser estándar. Veáse Herstein (1975) para más detalles. ■

Definición 2.4. *Sea R un anillo conmutativo, y sea $a \in \mathbb{R}$, $a \neq 0$. Si existe $b \in R$, $b \neq 0$ tal que $ab = 0$, se dice que a es un **divisor de cero**. Si en un anillo conmutativo R no hay divisores de cero, se dice que R es un **dominio entero** o **dominio de integridad**.*

Proposición 2.5. En un dominio entero, si $ac = bc$ y $c \neq 0 \Rightarrow a = b$.

Demostración.

$$ac = bc \Leftrightarrow ac - bc = 0 \Leftrightarrow (a - b)c = 0$$

como un dominio entero no tiene divisores de cero y $c \neq 0$ entonces $a - b = 0 \Leftrightarrow a = b$

■

2.2. Campos

Definición 2.6 (Campo)**.** *Dado un anillo conmutativo con identidad R, se dice que es un* ***campo*** *si todos sus elementos no cero tienen inverso multiplicativo, i.e. si* $\forall a \in R \setminus \{0\}\,,\ \exists \frac{1}{a} \in R$ *tal que* $a \cdot \frac{1}{a} = 1$.

***Ejemplo* 2.2**

Algunos campos importantes:

1. $\mathbb{Q}$ el conjunto de números racionales, $\mathbb{R}$, de números reales y $\mathbb{C}$, de números complejos, son campos.
2. $\mathbb{Q}(x)$, el conjunto de funciones racionales en variable x es un campo.
3. $\mathbb{R}[[x]]$, el conjunto de series de potencias $\sum a_i x^i$ con $a_0 \neq 0$, es un campo.

Teorema 2.7. Todo dominio entero finito es un campo

Demostración. Para el anillo trivial $D = \{0\}$, la identidad es el mismo elemento 0, y, por vacuidad, todos sus elementos no cero tienen inverso multiplicativo. Por lo que $\{0\}$ es un campo.

Sea ahora $D = \{a_1, a_2, \cdots, a_n\}$ un dominio entero finito ($|D| = n > 1$). Sea $a \in D \setminus \{0\}$ un elemento cualquiera $\Rightarrow$ para cualesquiera $a_i, a_j \in D$ con $i \neq j$ se debe tener que $aa_i \neq aa_j$, pues de lo contrario la proposición 2.5 implicaría que $a_i = a_j$ ($\rightarrow\leftarrow$). Considérese el conjunto $D' = \{aa_1, aa_2, \cdots, aa_n\} \subseteq D$, como los aa_i's son distintos, $|D'| = n \Rightarrow D' = D$. Como $a \in D = D' \Rightarrow \exists k$ tal que $aa_k = a_k a = a$. Luego, si $b \in D$ entonces existe r tal que $b = aa_r$ entonces:

$$a_k b = a_k(aa_r) = (a_k a)a_r = aa_r = b$$

Por lo que $a_k = 1$. Finalmente, como $1 \in D$ entonces existe $b \in D$ tal que $ab = 1 = ba$, como a era un elemento no nulo arbitrario, D tiene identidad e inversos multiplicativos y es un campo. ∎

***Ejemplo* 2.3**

Para un $n \in \mathbb{Z}^+$, con $n \geq 2$, en el ejemplo 2.1 se construyó el anillo conmutativo con identidad $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$, que es finito. Del teorema anterior es inmediato que $\mathbb{Z}_n$ es un campo si y sólo si es un dominio entero.

Sean $a, b \in \mathbb{Z}_n \setminus \{0\}$ entonces $ab = 0 \Leftrightarrow ab \equiv 0 \mod n \Leftrightarrow ab$ es múltiplo de n. Si n es compuesto, claramente existen a, b no nulos en $\mathbb{Z}_n$ tales que $n = ab$, por lo que $\mathbb{Z}_n$ no puede ser dominio entero. Si en cambio n fuera un número primo p, sus únicos múltiplos serían $p, 2p, 3p, \cdots$ que no pueden escribirse como el producto de dos números en $\mathbb{Z}_p$, por lo que $\mathbb{Z}_p$ no tiene divisores de cero y es un dominio entero.

Por lo tanto, Z_n es un campo finito si y sólo si $n \geq 2$ es un número primo. Como ejemplo concreto, note que en la tabla de multiplicaciones de $\mathbb{Z}_5$ (Tabla 2.1) todos los elementos no nulos tienen inverso multiplicativo.

Tabla 2.1: Tabla de multiplicación en $\mathbb{Z}_5$

$\times$	**1**	**2**	**3**	**4**
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

Teorema 2.8. Todo dominio entero se puede sumergir en un campo.

Es decir, dado un dominio entero D, existe un campo $\mathcal{F}$ tal que $\mathcal{F}$ contiene una copia de D (véase isomorfismos en Herstein, 1975, p. 131) y tal que las operaciones $+$ y $\cdot$ en $\mathcal{F}$ son extensiones de las definidas en D.

Observación. Al menor campo que contiene al dominio entero D se le llama **campo de cocientes de** D. A continuación se presenta un esbozo de la demostración de este resultado. Para más detalles véase Herstein (1975, p. 140).

Demostración. Esbozo:

1. Se construye el conjunto $D' = D \times (D \setminus \{0\}) = \{(a,b) | a, b \in D, b \neq 0\}$.

2. Se define la relación $\mathcal{R} \subseteq D'$ como

$$(a,b)\mathcal{R}(c,d) \Leftrightarrow ad = bc.$$

3. Se prueba que $\mathcal{R}$ es una relación de equivalencia y se construye el conjunto de clases de equivalencia $\mathcal{F} = D'/\mathcal{R}$.

4. Se denota por $\frac{a}{b}$ a la clase de equivalencia de (a,b) en $\mathcal{F}$.

5. Se definen la suma y el producto como:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd} \text{ y } \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

6. Se prueba que $\mathcal{F}$ es un campo al notar que $\frac{a}{a}$ para cualquier a es la identidad de $\mathcal{F}$ y que para $\frac{a}{b}$ su inverso es $\frac{b}{a}$ pues

$$\frac{a}{b} \cdot \frac{b}{a} = \frac{ba}{ab} = \frac{ab}{ab} = \frac{a}{a}.$$

7. Finalmente se nota que D es isomorfo al conjunto $\left\{\frac{ab}{b} \in \mathcal{F} | a \in D\right\}$ para algún $b \neq 0$ fijo. Con lo que hay una copia de D en $\mathcal{F}$ y este último extiende a D. ∎

***Ejemplo* 2.4**

$\mathbb{Q}$ es el campo de cocientes del dominio entero $\mathbb{Z}$. $\mathbb{K}(x)$ es el campo de cocientes de $\mathbb{K}[x]$.

2.3. Ideales

Definición 2.9. *Dado un anillo R, se dice que un subconjunto no vacío $I \subseteq R$ es un* ***ideal de*** *R si cumple con:*

1. $\forall\ a, b \in I \Rightarrow\ a + b \in I$.
2. $\forall\ a \in I$, *su inverso aditivo también está en* I, *i.e.* $-a \in I$.
3. *El neutro aditivo de* R *está en* I, *i.e.* $0 \in I$.
4. *Para cada* $u \in I$ *y* $r \in R$, $ur \in I$ *y* $ru \in I$.

Proposición 2.10. Sean R un anillo con identidad y $I \subseteq R$ un subconjunto no vacío. I es un ideal de R si y sólo si se cumplen:

I) Si $a, b \in I \Rightarrow\ a - b \in I$

II) Para cada $u \in I$ y $r \in R$, $ur \in I$ y $ru \in I$.

Demostración. ($\Rightarrow$) Si $I \subseteq R$ es un ideal. Sean $a, b \in I$, la condición 2 asegura que $-b \in I$, luego por la condición 1, $a + (-b) = a - b \in I$.

($\Leftarrow$) Si se cumplen I y II. Sea $a \in I$ entonces $a - a = 0 \in I$ y también $0 - a = -a \in I$. Además, si $a, b \in I \Rightarrow\ -b \in I \Rightarrow\ a - (-b) = a + b \in I$, por lo que I es un ideal de R. ∎

Definición 2.11. *Dados un anillo conmutativo con identidad R y un subconjunto no vacío $S \subseteq R$, el* ***ideal generado por*** *S, denotado por $\langle S \rangle$, es el conjunto de combinaciones lineales finitas de elementos de S:*

$$\langle S \rangle = \{r_1 s_1 + r_2 s_2 + \cdots r_n s_n | r_i \in R, s_i \in S, n \in \mathbb{N}\}$$

Si $S = \{s_1, \cdots, s_n\}$ es un conjunto finito, $\langle S \rangle$ se puede denotar también como $\langle s_1, \cdots, s_n \rangle$.

Proposición 2.12. Dado un anillo conmutativo con identidad R y un subconjunto no vacío $S \subseteq R$, $\langle S \rangle$ es el menor ideal que contiene a S.

Demostración. Se prueba primero que $\langle S \rangle$ es un ideal utilizando la proposición 2.10.

I) Sean $a, b \in \langle S \rangle$ con $a = r_1 s_1 + \cdots + r_n s_n$ y $b = r'_1 s'_1 + \cdots + r'_m s'_m$ donde $s_i, s'_i \in S$, $r_i, r'_i \in R$ y $n, m \in \mathbb{N}$. Entonces $a - b = r_1 s_1 + \cdots + r_n s_n - r'_1 s'_1 - \cdots - r'_m s'_m$ evidentemente es elemento de $\langle S \rangle$.

II) Sean $a \in \langle S \rangle$ y $r \in R$ con $a = r_1 s_1 + \cdots + r_n s_n$, donde $r_i \in R$ y $s_i \in S$ entonces

$$ar = ra = r(r_1 s_1 + \cdots + r_n s_n) = (rr_1)s_1 + \cdots + (rr_n)s_n$$

pertenece a $\langle S \rangle$ pues $rr_i \in R$. De donde $\langle S \rangle$ es un ideal de R.

Como R tiene identidad, entonces para todo $s \in S$, $1 \cdot s \in \langle S \rangle \Rightarrow\ S \subseteq \langle S \rangle$.

Finalmente, supóngase que existe un ideal I de R tal que $S \subseteq I$. Sea $a = r_1 s_1 + \cdots + r_n s_n \in \langle S \rangle$, como cada $s_i \in S \subseteq I$ entonces I debe atrapar los productos $r_i s_i \in I$ para $i = 1, 2, \cdots, n$. Como además I es cerrado bajo la suma, entonces $r_1 s_1 + \cdots + r_n s_n = a \in I$, por lo que $\langle S \rangle \subseteq I$ y el ideal $\langle S \rangle$ es el menor que contiene a S. ∎

Definición 2.13. *Un ideal I del anillo R es un* ***ideal principal*** *si existe un elemento $a \in R$ tal que $I = \langle a \rangle$. Si para un anillo R todos sus ideales son principales, entonces se dice que R es un* ***anillo de ideales principales*** *o* ***anillo principal****.*

Observación. Es claro que los ideales triviales $\{0\} = \langle 0 \rangle$ y $R = \langle 1 \rangle$ son ideales principales.

Definición 2.14 (Divisibilidad)**.** *Dados un anillo conmutativo con identidad R y $a, b \in R$, se dice que a* ***divide*** *b (o que b* ***es múltiplo de*** *a), denotado $a|b$, si existe otro elemento $c \in R$ tal que $b = ac$.*

En la siguiente proposición se enuncian sin demostración algunas propiedades de la relación de *divisibilidad* que serán de utilidad en los próximos capítulos. Refiérase a Herstein (1975, p. 144) para más detalles.

Proposición 2.15. Dados un anillo conmutativo con identidad R y $a, b, c \in R$ se cumplen:

1. $1|a$ y $a|0$, $\forall\ a \in R$.
2. $a|a$, $\forall\ a \in R$.
3. Si $a|b$ entonces $a|bc$.
4. Si $a|b$ y $b|c$ entonces $a|c$.
5. Si $a|b$ y $a|c$ entonces $a|(\lambda b + \mu c)$.

2.4. Cálculo de ideales

En el siguiente lema se introducen algunas operaciones de ideales cuyo resultado produce de nuevo un ideal. Tales operaciones son de utilidad para manipular ideales en los próximos capítulos.

Lema 2.16. Sea R un anillo. Sean $I, J \subseteq R$ ideales de R y $I_1 \subseteq I_2 \subseteq \cdots \subseteq R$ una familia contable de ideales de R, entonces:

1. $H = I \cap J$.
2. $H = \bigcup_i I_i$.
3. $H = I + J = \{c + d | c \in I, d \in J\}$.

4. $H = I \cdot J = \left\{ \sum_{i=1}^{k} c_i d_i \middle| \, c_i \in I, \ d_i \in J \right\}$.

5. $H = (I : J) = \{r \in R | rc \in I \ \forall \ c \in J\}$.

son todos ideales de R.

Demostración. Para todos los casos, se usa la proposición 2.10 para probar que H es un ideal:

1. Sean $a, b \in I \cap J \Rightarrow a, b \in I$ y $a, b \in J \ \Rightarrow a - b \in I$ y $a - b \in J \ \Rightarrow a - b \in I \cap J$. De igual forma, si $a \in I \cap J$ y $r \in R \ \Rightarrow a \in I$ y $a \in J$, entonces $ar, ra \in I$ y $ar, ra \in J$ por lo que $ar, ra \in I \cap J$. Por lo tanto, $I \cap J$ es un ideal.

2. Sean $a, b \in \bigcup_i I_i$ entonces existen I_k, I_r tales que $a \in I_k$, $b \in I_r$ y $I_k \subseteq I_r$ o viceversa. Supóngase sin pérdida de la generalidad que este es el caso, entonces $a, b \in I_r \Rightarrow a - b \in I_r \subseteq \bigcup_i I_i$. Luego, para cualquier $r \in R$, $ar, ra \in I_r \subseteq \bigcup_i I_i$, por lo que es un ideal.

3. Sean $a, b \in I + J \Rightarrow$ existen $c_1, c_2 \in I$ y $d_1, d_2 \in J$ tales que $a = c_1 + d_1$ $b = c_2 + d_2$. Entonces $a - b = (c_1 + d_1) - (c_2 + d_2) = (c_1 - c_2) + (d_1 - d_2)$, donde $c_1 - c_2 \in I$ y $d_1 - d_2 \in J$, por lo que $a - b \in I + J$. Igualmente, para $r \in R$, $ra = r(c_1 + d_1) = rc_1 + rd_1$ y $ar = (c_1 + d_1)r = c_1 r + d_1 r$ donde $rc_1, c_1 r \in I$ y $rd_1, d_1 r \in J$ por lo que $ra, ar \in I + J$ y es un ideal.

4. Sean $a, b \in I \cdot J$ es evidente que $a - b \in I \cdot J$. Si $a = \sum_{i=1}^{k} c_i d_i$ con $c_i \in I, d_i \in J$ entonces para $r \in R$,

$$ra = r \sum_{i=1}^{k} c_i d_i = \sum_{i=1}^{k} r(c_i d_i) = \sum_{i=1}^{k} (rc_i) d_i \quad \text{, con } rc_i \in I, \ d_i \in J \text{ y}$$

$$ar = \left(\sum_{i=1}^{k} c_i d_i \right) r = \sum_{i=1}^{k} (c_i d_i) r = \sum_{i=1}^{k} c_i (d_i r) \quad \text{, con } c_i \in I, \ d_i r \in J$$

   Entonces $ar, ra \in I \cdot J$ por lo que $I \cdot J$ es un ideal.

5. Si $a, b \in (I : J) \Rightarrow \ ad, bd \in I \ \forall \ d \in J \Rightarrow \ (a - b)d = ad - bd \in I \ \forall \ d \in J \Rightarrow \ a - b \in (I : J)$. Además, si $r \in R$ entonces $(ra)d = r(ad) \in I \ \forall \ d \in J \Rightarrow \ ra \in (I : J)$ y $(ar)d = a(rd)$ donde $rd \in J \ \forall \ d \in J$ por lo que $(ar)d = a(rd) \in I \ \forall \ d \in J \Rightarrow \ ar \in (I : J)$. Entonces $(I : J)$ es un ideal. ∎

Lema 2.17. Sean R un anillo conmutativo con identidad y $S, T \subseteq R$ subconjuntos no vacíos. Entonces se tienen las identidades:

1. $\langle S \rangle + \langle T \rangle = \langle S \cup T \rangle$.

2. $\langle S \rangle \cdot \langle T \rangle = \langle st | \ s \in S, t \in T \rangle$.

Demostración.

1. Si $a \in \langle S\rangle + \langle T\rangle$ entonces existen $\sigma \in \langle S\rangle$ y $\tau \in \langle T\rangle$ tales que $a = \sigma + \tau$. Como $\sigma = r_1s_1 + \cdots + r_ns_n$ y $\tau = r'_1t_1 + \cdots + r'_mt_m$ con $r_i, r'_i \in R$, $s_i \in S$ y $t_i \in T$, entonces $a = r_1s_1 + \cdots + r_ns_n + r'_1t_1 + \cdots + r'_mt_m \in \langle S \cup T\rangle$, de donde $\langle S\rangle + \langle T\rangle \subseteq \langle S \cup T\rangle$. Por otro lado, $S \subseteq \langle S\rangle + \langle T\rangle$ y $T \subseteq \langle S\rangle + \langle T\rangle$, por lo que $S \cup T \subseteq \langle S\rangle + \langle T\rangle$. Como $\langle S \cup T\rangle$ es el menor ideal que contiene a $S \cup T$, entonces $\langle S \cup T\rangle \subseteq \langle S\rangle + \langle T\rangle$ y se tiene la igualdad.

2. Sea $H = \langle st | s \in S, t \in T\rangle$. Si $a \in H \Rightarrow\ a = r_1(s_1t_1)+\cdots+r_n(s_nt_n) = (r_1s_1)t_1+\cdots+(r_ns_n)t_n$, donde los $r_is_i \in S$, por lo que $a = s'_1t_1 + \cdots + s'_nt_n \in \langle S\rangle \cdot \langle T\rangle$, de donde $H \subseteq \langle S\rangle \cdot \langle T\rangle$. Si $a \in \langle S\rangle \cdot \langle T\rangle \Rightarrow\ a = s_1t_1 + \cdots + s_nt_n \in H \Rightarrow\ \langle S\rangle \cdot \langle T\rangle \subseteq H$. Por lo que $(\langle S\rangle : \langle T\rangle) = \langle st|\ s \in S, t \in T\rangle$. ∎

Finalmente, el siguiente resultado resume algunas de las propiedades de estas operaciones de ideales, la demostración puede ser consultada en Fröberg (1997, p. 18).

Proposición 2.18. Sean I, J, H ideales en el anillo R. Entonces

1. $(I \cap H) + (J \cap H) \subseteq (I + J) \cap H$.
2. $(I : H) + (J : H) \subseteq ((I + J) : H)$.
3. $(I : (J + H)) = (I : J) \cap (I : H)$.
4. $I + J = J + I$, $I + (J + H) = (I + J) + H$.
5. $I \cdot (J \cdot H) = (I \cdot J) \cdot H$.
6. $I \cdot (J + H) = I \cdot J + I \cdot H$, $(I + J) \cdot H = I \cdot H + J \cdot H$.
7. $((I \cap J) : H) = (I : H) \cap (J : H)$.

2.5. Anillos cocientes

Definición 2.19. *Sean R un anillo e $I \subseteq R$ un ideal. Dados $a, b \in R$, se dice que a* ***es congruente con*** *b* ***módulo*** *I si y sólo si $a - b \in I$, lo que se denota por $a \equiv b \mod I$.*

Lema 2.20. Dados un anillo R y un ideal $I \subseteq R$, la relación *ser congruente a módulo I* ($\equiv$ mód I) es una relación de equivalencia.

Demostración.

1. **Reflexividad:** si $a \in R \Rightarrow\ a - a = 0 \in I$ pues I, siendo ideal, contiene al neutro aditivo.

   Entonces $a \equiv a \mod I\ , \forall\ a \in R$.

2. **Transitividad:** si $a \equiv b \mod I$ y $b \equiv c \mod I$ entonces $a - b \in I$ y $b - c \in I$, por lo que, dado que I es cerrado respecto a la adición, $a - c = (a - b) + (b - c) \in I \Rightarrow a \equiv c \mod I$.

3. **Simetría:** si $a \equiv b \mod I \Rightarrow a - b \in I$. Por la conmutatividad de la suma en I, $b - a = -a + b = -(a - b) \in I \Rightarrow b \equiv a \mod I$

■

Definición 2.21. *Dados un anillo R y un ideal $I \subseteq R$, al anillo cociente de R con la relación de equivalencia congruencia módulo I ($R/_{\equiv \text{ mód } I}$) se le denomina* ***anillo cociente de*** R ***con*** I *y se denota por R/I. A las clases de equivalencia de R/I se les denomina* ***clases laterales de*** I*. La clase lateral de I de un elemento $a \in R$ se denota por $[a]$ o bien $a + I$.*

Teorema 2.22. Dados un anillo R y un ideal $I \subseteq R$, el anillo cociente de R con I (R/I) es un anillo con las operaciones para $a + I, b + I \in R/I$ definidas por:

- $(a + I) + (b + I) = (a + b) + I$
- $(a + I) \cdot (b + I) = ab + I$

Demostración. La demostración de este resultado es estándar, por lo que se omite aquí. La clave de la prueba radica en utilizar el mapa canónico $\varphi : R \to R/I$ definido por $\varphi : a \mapsto a + I$, el cuál es un homomorfismo sobreyectivo, *i.e.* $\varphi(a + b) = \varphi(a) + \varphi(b)$, $\varphi(ab) = \varphi(a)\varphi(b)$, y mostrar que las propiedades de anillo de R se "heredan" en R/I. Refiérase a Herstein (1975, p. 135) para la prueba completa. ■

3 ANILLOS DE POLINOMIOS

En este capítulo se estudian los anillos de polinomios de una sola variable (generalmente x), como un caso particular de los anillos de polinomios de varias variables que se estudiarán en los próximos capítulos. El desarrollo de este capítulo se centra en el estudio de los ideales de estos anillos y particularmente en la solución a las interrogantes: (1) cómo determinar si un polinomio dado pertenece o no a un ideal, y (2) cómo son los elementos de los anillos cocientes que estos ideales inducen.

3.1. Polinomios sobre un anillo

La construcción aquí presentada corresponde a la utilizada por Garcia y Lequain (2003).

Definición 3.1. *Dado un anillo R, un* ***polinomio de una variable sobre*** *R es una sucesión $(r_0, r_1, \cdots, r_n, \cdots)$, con $r_i \in R$ para todo $i = 0, 1, 2, \cdots$ y tal que $r_i \neq 0$ únicamente para un número finito de índices.*

Se definen además las siguientes operaciones de *suma* y *producto* entre los polinomios sobre R.

- **Suma:** $(a_0, a_1, a_2, \cdots) + (b_0, b_1, b_2, \cdots) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \cdots)$.
- **Producto:** $(a_0, a_1, a_2, \cdots) \cdot (b_0, b_1, b_2, \cdots) = (c_0, c_1, c_2, \cdots)$, donde

$$\begin{cases} c_0 = a_0 b_0 \\ c_1 = a_0 b_1 + a_1 b_0 \\ \vdots \\ c_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_{n-1} b_1 + a_n b_0 \\ \vdots \end{cases}$$

Proposición 3.2. La suma y el producto de dos polinomios sobre un anillo R, son también polinomios sobre R.

Demostración. Sean $a = (a_0, a_1, \cdots)$ y $b = (b_0, b_1, \cdots)$ polinomios sobre R. Como únicamente un número finito de sus coeficientes son distintos de cero, existen $m, n \in \mathbb{N}$ tales que $a_i = 0 \ \forall \ i > m$ y $b_i = 0 \ \forall \ i > n$. Entonces:

- Sea $c = a + b$. Si $i_0 > \text{máx}(m, n)$, entonces

$$c_{i_0} = a_{i_0} + b_{i_0} = 0 + 0 = 0$$

Por lo que $c_i = 0 \ \forall \ i > \text{máx}(m, n)$ y la suma de a y b es un polinomio sobre R.

- Sea $c = a \cdot b$. Si $i_0 > m + n$, entonces

$$c_{i_0} = a_0 b_{i_0} + a_1 b_{i_0 - 1} + \cdots a_{i_0 - 1} b_1 + a_{i_0} b_0$$

donde todos los sumandos son de la forma $a_j b_k$ con $j + k = i_0 > m + n$, por lo que $j > m$ ó $k > n$, de donde $a_j = 0$ ó $b_k = 0 \Rightarrow a_j b_k = 0$. Por lo tanto, $c_{i_0} = 0$, $\forall\ i_0 > m + n$ y el producto de a y b es un polinomio sobre R. ■

Proposición 3.3. Los polinomios sobre un anillo conmutativo con identidad R, con la suma y el producto recién definidos, es un anillo conmutativo con identidad.

Demostración.

- La cerradura de la suma y el producto se probaron en la proposición anterior.
- (Conmutatividad de la suma)

$$a + b = (a_0 + b_0, \cdots, a_n + b_n, \cdots) = (b_0 + a_0, \cdots, b_n + a_n, \cdots) = b + a$$

- (Asociatividad de la suma)

$$\begin{aligned} a + (b + c) &= (a_0, \cdots, a_n, \cdots) + (b_0 + c_0, \cdots, b_n, c_n, \cdots) = \\ &= (a_0 + (b_0 + c_0), \cdots a_n + (b_n + c_n), \cdots) = \\ &= ((a_0 + b_0) + c_0, \cdots, (a_n + b_n) + c_n, \cdots) = \\ &= (a_0 + b_0, \cdots, a_n, b_n) + (c_0, \cdots, c_n, \cdots) = (a + b) + c \end{aligned}$$

- (Neutro aditivo) Si $0 = (0, 0, \cdots)$, claramente un polinomio sobre R, se cumple

$$0 + a = a + 0 = (a_0, a_1, \cdots) + (0, 0, \cdots) = (a_0 + 0, a_1 + 0, \cdots) = a$$

- (Inverso aditivo) Para $a = (a_0, a_1, \cdots)$, sea $-a = (-a_0, -a_1, \cdots)$, entonces

$$(-a) + a = a + (-a) = (a_0, a_1, \cdots) + (-a_0, -a_1, \cdots) = (a_0 - a_0, a_1 - a_1, \cdots) = (0, 0, \cdots) = 0$$

- (Asociatividad del producto)

$$\begin{aligned} ((ab)c)_m &= \sum_{i=0}^{m} (ab)_i c_{m-i} = \sum_{i=0}^{m} \left(\sum_{j=0}^{i} a_j b_{i-j} \right) c_{m-i} = \\ &= \sum_{i=0}^{m} \sum_{j=0}^{i} (a_j b_{i-j} c_{m-i}) = \sum_{j=0}^{m} \sum_{i=j}^{m} (a_j b_{i-j} c_{m-i}) = \\ &= \sum_{j=0}^{m} a_j \left(\sum_{i=0}^{m-j} b_{(i-j)+j} c_{m-(i+j)} \right) = \sum_{j=0}^{m} a_j \left(\sum_{i=0}^{m-j} b_i c_{(m-j)-i} \right) = \\ &= \sum_{j=0}^{m} a_j (bc)_{m-j} = (a(bc))_m \end{aligned}$$

Para todo m, por lo que $(ab)c = a(bc)$.

- (Conmutatividad del producto)

  $$(ab)_m = a_0b_m + a_1b_{m-1} + \cdots + a_{m-1}b_1 + a_mb_0 = b_ma_0 + b_{m-1}a_1 + \cdots + b_1a_{m-1} + b_0a_m =$$
  $$= b_0a_m + b_1a_{m-1} + \cdots + b_{m-1}a_1 + b_ma_0 = (ba)_m$$

  Para todo m, por lo que $ab = ba$.

- (Ley distributiva)

  $$(a(b+c))_m = \sum_{i=0}^{m} a_i(b+c)_{m-i} = \sum_{i=0}^{m} a_i(b_{m-i} + c_{m-i}) =$$
  $$= \sum_{i=0} a_ib_{m-i} + a_ic_{m-i} = \sum_{i=0}^{m} a_ib_{m-i} + \sum_{i=0}^{m} a_ic_{m-i} =$$
  $$= (ab)_m + (ac)_m$$

  Para todo m, por lo que $a(b+c) = ab + ac$.

- (Identidad) Si $1 = (1, 0, 0, \cdots)$, 1 es la identidad multiplicactiva, pues

  $$1 \cdot a = a \cdot 1 = (a_0, a_1, \cdots, a_n, \cdots) \cdot (1, 0, 0, \cdots) = (a_0, a_1, \cdots, a_n, \cdots)$$

Por lo tanto, los polinomios sobre un anillo conmutativo con identidad, forman un anillo conmutativo con identidad. ■

Observación. Denótese por p^n a la multiplicación de n copias de p, es decir

$$p^n = \underbrace{p \cdot p \cdots p}_{n \text{ veces}}$$

lo cuál está bien definido gracias a la asociatividad del producto. Por simplicidad, se denotará con la variable x al polinomio $(0, 1, 0, \cdots)$. Nótese ahora:

$$x^n = (0, 1, 0, \cdots)^n = (0, \cdots, 0, \underbrace{1}_{\text{posición } n+1}, 0, \cdots)$$

y que

$$(0, \cdots, 0, \underbrace{a_n}_{\text{posición } n+1}, 0, \cdots) = (a_n, 0, 0, \cdots) \cdot (0, \cdots, 0, \underbrace{1}_{\text{posición } n+1}, 0, \cdots) = (a_n, 0, 0, \cdots)x^n$$

Por lo tanto

$$(a_0, a_1, \cdots, a_n, 0, 0, \cdots) = (a_0, 0, \cdots) + (a_1, 0, 0, \cdots)x + (a_2, 0, 0, \cdots)x^2 + \cdots + (a_n, 0, 0, \cdots)x^n$$

Además, en lugar de escribir $(a_i, 0, \cdots)$ se escribirá únicamente a_i. Se hace esta convención sin peligro de crear confusiones, pues es posible mostrar que los polinomios de la forma $(a_i, 0, \cdots)$ son una copia isomórfica del anillo R, por lo que por simplicidad se dirá que $(a_i, 0, 0, \cdots) = a_i \in R$. Bajo estas convenciones, se escribe el polinomio a como:

$$a(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

donde a los sumandos $a_i x^i$ que conforman a se les demonina **términos**, y, en cada término, se dice que a_i es su **coeficiente** y x^i su **monomio**. Por simplicidad se escribe $1x^i = x^i$. Adoptando finalmente la notación $R[x]$ para el **anillo de polinomios sobre R en variable x**.

Una vez definido el anillo de polinomios $R[x]$ con la notación y operaciones usuales, se procede ahora a nombrar algunas partes importantes de los polinomios.

Definición 3.4. *Sea $p = a_0 + a_1 x + \cdots + a_n x^n$, con $p \neq 0$ un polinomio. Se definen*

- *El **término principal**, $\mathbf{lt}(p)$, como el término de p en el que la variable x tiene el mayor exponente, $\mathbf{lt}(p) = a_n x^n$*
- *El **coeficiente principal**, $\mathbf{lc}(p)$, como el coeficiente del término principal de p, $\mathbf{lc}(p) = a_n$*
- *El **monomio principal**, $\mathbf{lm}(p)$, como el monomio del término principal de p, $\mathbf{lm}(p) = x^n$*
- *El **grado de** p, $\mathbf{deg}(p)$, como el exponente de x en el monomio principal de p, $\mathbf{deg}(p) = n$.*

*Además, si el coeficiente principal de un polinomio es 1, se dice que es un polinomio **mónico**. Si $\mathbf{deg}(p) = 0$, se dice que es una **constante** y se tiene que $p \in R \subseteq R[x]$ (en el sentido de la observación anterior).*

***Ejemplo* 3.1**
Sea $p \in \mathbb{Z}[x]$ tal que $p = 2 + 3x - 8x^3 + 3x^5 + 4x^6$. El término principal de p es el sumando en el que la variable x tiene mayor exponente, en este caso será $\mathbf{lt}(p) = 4x^6$. Luego es evidente que

- El coeficiente principal es $\mathbf{lc}(p) = 4$.
- El monomio principal es $\mathbf{lm}(p) = x^6$.
- Y el grado de p es $\mathbf{deg}(p) = 6$.

Proposición 3.5. Dados $p, q \in R[x] \setminus \{0\}$. Entonces $\mathbf{deg}(p+q) \leq \text{máx}(\mathbf{deg}(p), \mathbf{deg}(q))$.

Demostración. Sean $p = a_0 + a_1 x + \cdots + a_n x^n$, $a_n \neq 0$ y $q = b_0 + b_1 x + \cdots + b_m x^m$, $b_m \neq 0$. Sea $r = p + q$, con $r = c_0 + c_1 x + \cdots + c_k x^k$, $c_k \neq 0$. Entonces $\mathbf{deg}(p) = n$, $\mathbf{deg}(q) = m$ y $\mathbf{deg}(r) = k$. Como se hizo en la demostración de la proposición 3.2, $r_i = 0$ $\forall i > \text{máx}(m, n)$, dado que $r_k \neq 0$, la única opción posible es que $\mathbf{deg}(r) = k \leq \text{máx}(n, m) = \text{máx}(\mathbf{deg}(p), \mathbf{deg}(q))$ ∎

Aunque es posible estudiar los polinomios sobre un anillo cualquiera R, para los propósitos de este trabajo se requerirá trabajar con polinomios sobre un campo $\mathbb{K}$. Para mayor generalidad de los resultados, las siguientes proposiciones se trabajan sobre un dominio entero $\mathbb{D}$, haciendo el cambio definitivo a un campo en la próxima sección.

Proposición 3.6. Sea $\mathbb{D}$ un dominio entero. Si $p, q \in \mathbb{D}[x] \setminus \{0\}$, entonces $\mathbf{deg}(pq) = \mathbf{deg}(p) + \mathbf{deg}(q)$ y $\mathbf{lt}(pq) = \mathbf{lt}(p)\,\mathbf{lt}(q)$.

Demostración. Sean $p = \sum_{i=0}^{m} a_i x^i$ y $q = \sum_{j=0}^{n} b_j x^j$ con $a_m b_n \neq 0 \Rightarrow \mathbf{deg}(p) = m$ y $\mathbf{deg}(q) = n$. Sea $r = pq = \sum_{i=0}^{k} c_i x^i$ con $c_k \neq 0$ Como se mostró en la proposición 3.2, $c_i = 0 \; \forall \; i > m+n$, por lo que $\mathbf{deg}(r) = k \leq m+n$. Ahora bien,

$$c_{m+n} = \underbrace{a_0 b_{m+n} + a_1 b_{m+n-1} + \cdots}_{b_i = 0 \;\forall\; i > n} + a_m b_n + \underbrace{\cdots + a_{m+n-1} b_1 + a_{m+n} b_0}_{a_i = 0 \;\forall\; i > m} = a_m b_n$$

Como $\mathbb{D}$ es un dominio entero, entonces $c_{m+n} = a_m b_n \neq 0$, lo que significa que $\mathbf{deg}(r) \geq m+n$. Por lo tanto $\mathbf{deg}(r) = m+n = \mathbf{deg}(p) + \mathbf{deg}(q)$. Esto significa que $a_m b_n x^{m+n}$ es el término principal de r, por lo que

$$\mathbf{lt}(r) = a_m b_n x^{m+n} = (a_m x^m)(b_n x^n) = \mathbf{lt}(p)\,\mathbf{lt}(q)$$

■

Teorema 3.7. Si $\mathbb{D}$ es un dominio entero, $\mathbb{D}[x]$ también es un dominio entero.

Demostración. Sean $p, q \in \mathbb{D}[x] \setminus \{0\}$. Entonces p y q tienen cada uno al menos un término no cero, *i.e.* $p = \sum_{i=0}^{m} a_i x^i$ y $q = \sum_{i=0}^{n} b_i x^i$ con $a_m, b_n \neq 0$. Por la proposición anterior, $\mathbf{lt}(pq) = \mathbf{lt}(p)\,\mathbf{lt}(q) = a_m b_n x^{m+n} \neq 0$, por lo que $pq \neq 0$. ■

Observación. La importancia de este resultado se obtiene en conjunto con el teorema 2.8, que permite construir el campo de cocientes $\mathbb{D}(x) \supset \mathbb{D}[x]$, conocido como el **campo de funciones racionales en** x **sobre** $\mathbb{D}$. De ahora en adelante, se utilizará $\mathbb{D} = \mathbb{K}$ un campo, pues es conveniente tener inversos multiplicativos, dado que permiten dividir términos: $\frac{ax^m}{bx^n} = \frac{a}{b} x^{m-n}$, operación que es indispensable para que exista un algoritmo de división.

Proposición 3.8. Sea $\mathbb{K}$ un campo. Los únicos elementos de $\mathbb{K}[x]$ con inverso multiplicativo son las constantes.

Demostración. Sea $p \in \mathbb{K}[x]$ un elemento con inverso, *i.e.* $\exists \; q \in \mathbb{K}[x]$ tal que $pq = 1$. Entonces $\mathbf{deg}(pq) = \mathbf{deg}(1) = 0 \Rightarrow 0 = \mathbf{deg}(pq) = \mathbf{deg}(p) + \mathbf{deg}(q) \geq \mathbf{deg}(p)$ por lo que $\mathbf{deg}(p) = 0$ y $p \in \mathbb{K}$ es una constante. Además, todas las constantes son elementos del campo $\mathbb{K}$ (más precisamente, a una copia isomórfica), por lo que tienen inversos multiplicativos (siempre que no sean cero). ■

3.2. El Algoritmo de la división y de Euclides

Se estudian ahora el algoritmo de la división y el algoritmo de Euclides en el anillo de polinomios $\mathbb{K}[x]$. Estos serán de vital importancia para dar respuesta a los problemas que se plantearon al comienzo del capítulo.

Teorema 3.9 (Algoritmo de la división para una variable)**.** Dados dos polinomios $p, q \in \mathbb{K}[x]$ con $q \neq 0$, existen dos únicos polinomios $c, r \in \mathbb{K}[x]$ tales que $p = cq + r$, donde $r = 0$ o bien $\mathbf{deg}(r) < \mathbf{deg}(q)$.

Demostración.

- **Existencia:** El Algoritmo 3.1 que se muestra a continuación, devuelve dos polinomios p y $c \in \mathbb{K}[x]$ que cumplen lo buscado.
- **Unicidad:** Supóngase que existen dos parejas de polinomios c_1, r_1 y c_2, r_2 que cumplen con el teorema $p = c_1 q + r_1 = c_2 q + r_2$, entonces $(c_1 - c_2)q = r_2 - r_1$. De esta igualdad de polinomios, si $r_2 - r_1 \neq 0$, se deduce la igualdad de sus grados:

  $$\mathbf{deg}(c_1 - c_2) + \mathbf{deg}(q) = \mathbf{deg}((c_1 - c_2)q) = \mathbf{deg}(r_2 - r_1) \leq \text{máx}(\mathbf{deg}(r_1), \mathbf{deg}(r_2)) < \mathbf{deg}(q)$$

  luego, $\mathbf{deg}(c_1 - c_2) < 0$, pero como el grado debe ser un número no negativo, se llega a una contradicción. Entonces se tiene que $r_1 = r_2$. Finalmente como $(c_1 - c_2)q = 0$ y $q \neq 0$, $c_1 - c_2 = 0 \Rightarrow c_1 = c_2$.

■

Algoritmo 3.1 (Algoritmo de la división para una variable)

Input: $p, q \in \mathbb{K}[x]$ con $q \neq 0$

Output: $c, r \in \mathbb{K}[x]$ tales que $p = cq + r$, y $\mathbf{deg}(r) < \mathbf{deg}(q)$ o $r = 0$

begin
  $c \leftarrow 0; r \leftarrow p$;
  while $r \neq 0$ **and** $\mathbf{deg}(q) \leq \mathbf{deg}(r)$ **do**
    $c \leftarrow c + \mathbf{lt}(r)/\mathbf{lt}(q)$;
    $r \leftarrow r - (\mathbf{lt}(r)/\mathbf{lt}(q))q$;
  end
end

Demostración.

- **Finitud:** Nótese que:

  $$\mathbf{lt}((\mathbf{lt}(r)/\mathbf{lt}(q))q) = (\mathbf{lt}(r)/\mathbf{lt}(q))\mathbf{lt}(q) = \mathbf{lt}(r)$$

  lo que significa que en cada pasada por el ciclo **while**, el término principal de r se cancela. Si r se vuelve cero, se termina el algoritmo; en caso contrario, el grado de r disminuye. Dado que el grado inicial de r, $\mathbf{deg}(r) = \mathbf{deg}(p)$ es finito, si no se hace cero, disminuye hasta que $\mathbf{deg}(r) < \mathbf{deg}(q)$ y termina el algoritmo.
- **Correctitud:** La condición de terminación del algoritmo asegura que al final $r = 0$ o $\mathbf{deg}(r) < \mathbf{deg}(q)$. Obsérvese que la identidad $p = cq + r$ se mantiene durante todo el algoritmo. Inicialmente cuando $c = 0$ y $r = p$, se tiene: $p = 0 + p = cq + r$, y con cada redefinición de r' y c' se tiene:

  $$c'q + r' = (c + \mathbf{lt}(r)/\mathbf{lt}(q))q + (r - (\mathbf{lt}(r)/\mathbf{lt}(q))q) = cq + r = p$$

Con lo que el algoritmo devuelve lo requerido.

■

Definición 3.10. *Dados $p, q \in \mathbb{K}[x]$ con $q \neq 0$ y $c, r \in \mathbb{K}[x]$ tales que $p = cq + r$ como resultado del algoritmo de la división, a r se le llama* ***residuo de*** *p* ***entre*** *q y se denota por* $\mathbf{rem}(p, q)$.

***Ejemplo* 3.2**

Se ilustra ahora cómo funciona el algoritmo de la división. En los pasos en los que se hace la división $\mathbf{lt}(r)\,/\mathbf{lt}(q)$ se evidencia la necesidad de que $\mathbb{K}$ sea un campo.

Sean $p = x^4 + 3x^2 - 2x + 1$ y $q = 2x + 1$ en $\mathbb{Q}[x]$. Los pasos del algoritmo de la división se muestran en la Tabla 3.1, y dan como resultado:

$$p = \left(\frac{1}{2}x^3 - \frac{1}{4}x^2 + \frac{13}{8}x - \frac{29}{16}\right)(2x+1) + \left(\frac{45}{16}\right)$$

Tabla 3.1: Ejemplo algoritmo de la división (algoritmo 3.1)

1.	$c \leftarrow 0$ $r \leftarrow p = x^4 + 3x^2 - 2x + 1$	 $\mathbf{deg}(r) = 4 \geq \mathbf{deg}(q) = 1$
2.	$c \leftarrow c + \frac{\mathbf{lt}(r)}{\mathbf{lt}(q)} = 0 + \frac{x^4}{2x} = \frac{1}{2}x^3$ $r \leftarrow r - \left(\frac{\mathbf{lt}(r)}{\mathbf{lt}(q)}\right) q = p - \frac{1}{2}x^3 q = -\frac{1}{2}x^3 + 3x^2 - 2x + 1$	 $\mathbf{deg}(r) = 3 \geq \mathbf{deg}(q) = 1$
3.	$c \leftarrow c + \frac{\mathbf{lt}(r)}{\mathbf{lt}(q)} = c + \frac{-1/2x^3}{2x} = \frac{1}{2}x^3 - \frac{1}{4}x^2$ $r \leftarrow r - \left(\frac{\mathbf{lt}(r)}{\mathbf{lt}(q)}\right) q = r + \frac{1}{4}x^2 q = \frac{13}{4}x^2 - 2x + 1$	 $\mathbf{deg}(r) = 2 \geq \mathbf{deg}(q) = 1$
4.	$c \leftarrow c + \frac{\mathbf{lt}(r)}{\mathbf{lt}(q)} = c + \frac{13/4x^2}{2x} = \frac{1}{2}x^3 - \frac{1}{4}x^2 + \frac{13}{8}x$ $r \leftarrow r - \left(\frac{\mathbf{lt}(r)}{\mathbf{lt}(q)}\right) q = r + \frac{13}{8}xq = -\frac{29}{8}x + 1$	 $\mathbf{deg}(r) = 1 \geq \mathbf{deg}(q) = 1$
5.	$c \leftarrow c + \frac{\mathbf{lt}(r)}{\mathbf{lt}(q)} = c + \frac{-29/8x}{2x} = \frac{1}{2}x^3 - \frac{1}{4}x^2 + \frac{13}{8}x - \frac{29}{16}$ $r \leftarrow r - \left(\frac{\mathbf{lt}(r)}{\mathbf{lt}(q)}\right) q = r + \frac{29}{16}q = \frac{45}{16}$	 $\mathbf{deg}(r) = 1 \not\geq \mathbf{deg}(q) = 1 \hookleftarrow$

Definición 3.11 (MCD)**.** *Sean $p, q \in \mathbb{K}[x]$ dos polinomios distintos de cero. Se dice que el polinomio $h \in \mathbb{K}[x]$ es un* ***máximo común divisor*** *de p y q (denotado $h = mcd(p, q)$) si*

1. *h divide p y h divide q.*
2. *Si h_1 divide p y h_1 divide q, entonces h_1 divide h.*

Proposición 3.12. Dados $p, q \in \mathbb{K}[x]$ distintos de cero, si h_1 y h_2 son máximos comunes divisores de p y q entonces existe una constante $a \in \mathbb{K}$ tal que $h_1 = ah_2$.

Demostración. Como h_2 es máximo común divisor de p y q entonces $h_2|p$ y $h_2|q$. Como h_1 también es máximo común divisor, $h_2|h_1$ entonces existe $a \in \mathbb{K}[x]$ tal que $h_1 = ah_2$. Análogamente, existe $b \in \mathbb{K}[x]$ tal que $h_2 = bh_1$. Entonces

$$h_1 = ah_2 = abh_1 \;\Rightarrow\; (ab-1)h_1 = 0 \Rightarrow\; ab = 1$$

Como las constantes son los únicos elementos de $\mathbb{K}[x]$ con inverso, entonces $a, b \in \mathbb{K}$ y se tiene que $h_1 = ah_2$, $a \in \mathbb{K}$. ∎

Algoritmo 3.2 (Algoritmo de Euclides)

Input: $p, q \in \mathbb{K}[x]$ con $q \neq 0$

Output: $h = mcd(p, q)$

```
begin
    r_0 ← p; r_1 ← q ;
    r_2 ← rem(r_0, r_1);
    while r_2 ≠ 0 do
        r_0 ← r_1; r_1 ← r_2;
        r_2 ← rem(r_0, r_1);
    end
    h ← r_1
end
```

Demostración.

- **Finitud:** El algoritmo de la división asegura que $\mathbf{deg}(r_2) = \mathbf{deg}(\mathbf{rem}(r_0, r_1)) < \mathbf{deg}(r_1)$ y en cada paso por el ciclo **while** r_1 se redefine como r_2, por lo que el grado de r_2 forma una sucesión decreciente de enteros no negativos, que eventualmente llega a cero. Luego de esto, el grado de r_2 no puede disminuir más, por lo que la única opción que deja el algoritmo de la división es que $r_2 = 0$ y entonces se termina el algoritmo de Euclides.

- **Correctitud:** Se comienza por probar que: $p, q \in \mathbb{K}[x] \Rightarrow \ mcd(p, q) = mcd(q, \mathbf{rem}(p, q))$ (salvo constantes).

  Nótese que $p = cq + \mathbf{rem}(p, q)$ para algún $c \in \mathbb{K}[x]$, por lo que $mcd(q, \mathbf{rem}(p, q)) \,|p$ además $mcd(q, \mathbf{rem}(p, q)) \,|q$ entonces $mcd(q, \mathbf{rem}(p, q)) \,|mcd(p, q)$. Por otro lado, $\mathbf{rem}(p, q) = p - cq \ \Rightarrow \ mcd(p, q) \,|\mathbf{rem}(p, q)$ como además $mcd(p, q) \,|q \Rightarrow mcd(p, q) \,|mcd(q, \mathbf{rem}(p, q))$, de donde, por la propiedad 3.12, $mcd(p, q) = mcd(q, \mathbf{rem}(p, q))$ concluyendo lo buscado.

  Se aplica ahora esta propiedad al algoritmo de Euclides. Como notación, en cada paso por el ciclo **while**, se escribirán con primas la variables redefinidas y sin primar las no redefinidas:

  $$mcd(r_0', r_1') = mcd(r_0', r_2) = mcd(r_1, \mathbf{rem}(r_0, r_1)) = mcd(r_1, r_0)$$

  Esto implica que en el último paso se tendrá $r_2' = 0$ y

  $$h = r_1' = mcd(r_1', 0) = mcd(r_1', r_2') = mcd(r_1', \mathbf{rem}(r_0', r_1')) = mcd(r_0', r_1') = \cdots = mcd(p, q)$$

∎

3.3. Ideales de $\mathbb{K}[x]$

Teorema 3.13. $\mathbb{K}[x]$ es un anillo de ideales principales.

Demostración. Sea $I \subseteq \mathbb{K}[x]$ un ideal de $\mathbb{K}[x]$. Considérese $N = \{\mathbf{deg}(p) \mid p \in I\} \subseteq \mathbb{N}$. Por el principio del buen orden N tiene un elemento mínimo n, sea $q \in I$ tal que $\mathbf{deg}(q) = n$. Se probará que $I = \langle q \rangle$.

Como $q \in I$, se tiene que $\langle q \rangle \subseteq I$. Sea ahora $p \in I$ un elemento cualquiera. Por el algoritmo de la división existen $c, r \in \mathbb{K}[x]$ tales que $p = cq + r$ con $r = 0$ o $\mathbf{deg}(r) < \mathbf{deg}(q) = n$. Al despejar para r se obtiene que $r = cq - p \in I$ (pues $p, q \in I$) por lo que $\mathbf{deg}(r) \geq n$ entonces $r = 0$ y $p = cq \in \langle q \rangle$. Dado que p era arbitrario, $I \subseteq \langle q \rangle$ y entonces $I = \langle q \rangle$ ■

El teorema 3.13 permite conocer la estructura de todos los ideales de $\mathbb{K}[x]$, más aún, crea una identificación entre los elementos de $\mathbb{K}[x]$ y sus ideales, lo que permite estudiar a los ideales de $\mathbb{K}[x]$ por medio de los polinomios que los generan. Considérense dos ideales $I, J \subseteq \mathbb{K}[x]$, existen $p, q \in \mathbb{K}[x]$ tales que $I = \langle p \rangle$, $J = \langle q \rangle$. Si $I \subseteq J$ entonces $p \in J = \langle q \rangle \Rightarrow$ existe $c \in \mathbb{K}[x]$ tal que $p = cq \Rightarrow q|p$. Por otro lado, si $q|p \Rightarrow \exists\, c \in \mathbb{K}[x]$ tal que $p = cq \Rightarrow p \in \langle q \rangle$ por lo que $\langle p \rangle \subseteq \langle q \rangle \Rightarrow I \subseteq J$. En otras palabras,

$$\langle p \rangle \subseteq \langle q \rangle \Leftrightarrow q|p$$

En el caso en que $\langle p \rangle = \langle q \rangle$ se tendría $p|q$ y $q|p$ por lo que debe existir una constante $a \in \mathbb{K}$ tal que $p = aq$. De aquí, es evidente que la correspondencia entre ideales y polinomios de $\mathbb{K}[x]$ es única, salvo constantes, e identifica a las relaciones de orden de contención de ideales ($\supseteq$) con la de divisibilidad ($|$) entre polinomios. Si un ideal de $\mathbb{K}[x]$ está dado como el generado de un conjunto finito de polinomios: $\langle p_1, p_2, \cdots, p_m \rangle$, el teorema 3.13 asegura que ese mismo ideal es generado por un sólo elemento. El siguiente teorema permite calcular dicho polinomio explícitamente.

Teorema 3.14. Dado $I = \langle p_1, p_2, \cdots, p_m \rangle \subseteq \mathbb{K}[x]$ entonces $I = \langle mcd(p_1, p_2, \cdots, p_m) \rangle$. Donde $mcd(p_1, p_2, p_3) = mcd(p_1, mcd(p_2, p_3)) = mcd(mcd(p_1, p_2), p_3)$.

Demostración. Por inducción sobre m:

1. Cuando $m = 2$, $I = \langle p_1, p_2 \rangle$. Por el teorema 3.13 existe $h \in I$ tal que $I = \langle h \rangle$. *Comoh* $\in I \Rightarrow \exists\, c_1, c_2$ tales que $h = c_1 p_1 + c_2 p_2$, como $mcd(p_1, q_1)$ divide a p_1 y $p_2 \Rightarrow mcd(p_1, p_2)\,|h$. Por otro lado, el $mcd(p_1, p_2)$ debe ser una combinación lineal de ellos (Corolario 3.15), por lo que $mcd(p_1, p_2) \in I = \langle h \rangle \Rightarrow h|mcd(p_1, p_2)$. De la doble divisibilidad, $\langle p_1, p_2 \rangle = \langle h \rangle = \langle mcd(p_1, p_2) \rangle$

2. Para $m \geq 2$,

$$\begin{aligned}\langle p_1, p_2, \cdots, p_m \rangle &= \langle p_1 \rangle + \langle p_2, \cdots, p_m \rangle = \langle p_1 \rangle + \langle mcd(p_2, \cdots, p_m) \rangle = \\ &= \langle p_1, mcd(p_2, \cdots, p_m) \rangle = \langle mcd(p_1, mcd(p_2, \cdots, p_m)) \rangle = \\ &= \langle mcd(p_1, p_2, \cdots, p_m) \rangle \blacksquare\end{aligned}$$

Corolario 3.15. Dados $p, q \in \mathbb{K}[x]$ con $q \neq 0$, su máximo común divisor h es combinación lineal de ellos, *i.e.* $\exists\ \lambda_1, \lambda_2 \in \mathbb{K}[x]$ tales que $h = \lambda_1 p + \lambda_2 q$

Demostración. Por el teorema anterior, $\langle p, q\rangle = \langle mcd(p, q)\rangle$, lo que significa que $mcd(p, q) \in \langle p, q\rangle$. Por lo tanto, deben existir polinomios $\lambda_1, \lambda_2 \in \mathbb{K}[x]$ tales que

$$mcd(p, q) = \lambda_1 p + \lambda_2 q$$ ■

Luego de estudiar los algoritmos de la división y de Euclides, y de caracterizar los ideales de $\mathbb{K}[x]$, es posible resolver los problemas que se plantearon al inicio del capítulo.

Problema 1. Dados un ideal $I = \langle p_1, p_2, \cdots, p_m\rangle \subseteq \mathbb{K}[x]$ y un polinomio no cero $p \in \mathbb{K}[x]$. Determinar si p es un elemento de I o no.

Solución.

1. Calcular $h = mcd(p_1, p_2, \cdots, p_m)$ aplicando el algoritmo de Euclides (algoritmo 3.2) asociativamente.

2. Determinar $r = \mathbf{rem}(p, h)$ utilizando el algoritmo de la división (algoritmo 3.1).

3. p es elemento de I si y sólo si $r = 0$.

Problema 2. Dado un ideal $I = \langle p_1, p_2, \cdots, p_m\rangle$. Caracterizar el anillo cociente $\mathbb{K}[x]/I$, *i.e.* determinar la forma de los elementos de $\mathbb{K}[x]/I$ y determinar la forma de la suma y el producto.

Solución. Sea $q = mcd(p_1, \cdots, p_m) \in I$, entonces $I = \langle q\rangle$. Sea $[p] \in \mathbb{K}[x]/I$, por el algoritmo de la división se tiene $p = cq + \mathbf{rem}(p, q) \Rightarrow\ p - \mathbf{rem}(p, q) = cq \in I$ por lo que $[p] = [\mathbf{rem}(p, q)]$, con $\mathbf{deg}(\mathbf{rem}(p, q)) < \mathbf{deg}(p) = n$ y $\mathbf{rem}(p, q)$ es único. Es decir, para cada clase de $K[x]/I$ existe un único polinomio de grado menor a q que representa a dicha clase. Por otro lado, es evidente que para cualquier polinomio $p \in \mathbb{K}[x]$ de grado menor que q, la clase $[p] \in \mathbb{K}[x]/I$, por lo que es posible ver a $\mathbb{K}[x]/I$ como:

$$\mathbb{K}[x]/I \approx \{[p] \mid p \in \mathbb{K}[x],\ \mathbf{deg}(p) < n\} = \left\{ \sum_{i=0}^{n-1} a_i x^i \,\middle|\, a_i \in \mathbb{K} \right\}$$

Si $f, g \in \mathbb{K}[x]/I$ con $\mathbf{deg}(f), \mathbf{deg}(g) < n \Rightarrow \mathbf{deg}(f + g) \leq \text{máx}(\mathbf{deg}(f), \mathbf{deg}(g)) < n$ por lo que $f + g \in \mathbb{K}[x]/I$ y no hay necesidad de redefinir la suma. De igual manera, para $\alpha \in \mathbb{K}$, $\mathbf{deg}(\alpha f) = \cancelto{0}{\mathbf{deg}(\alpha)} + \mathbf{deg}(f) = \mathbf{deg}(f) < n$ y $\alpha f \in \mathbb{K}$, por lo que $(\mathbb{K}[x]/I, +, \cdot)$ puede ser visto como un espacio vectorial de dimensión n y base $\mathcal{B} = \left\{1, x^1, x^2, \cdots, x^{n-1}\right\}$ (Herstein, 1975).

Finalmente, sólo falta caracterizar el producto del anillo $\mathbb{K}[x]/I$, el cual está dado por

$$f_1 f_2 := \mathbf{rem}(f_1 f_2, q)\,.$$

3.4. Factorización y ceros de polinomios

Dado p un polinomio de grado positivo en $\mathbb{K}[x]$, para cualquier constante $c \in \mathbb{K} \setminus \{0\}$ se puede escribir, $p = c\frac{1}{c}p$, lo que significa que tanto c como cp siempre son divisores de p. A esta familia de divisores se les denomina **divisores triviales**.

Definición 3.16. *Dado un polinomio p en $\mathbb{K}[x]$, se dice que es **irreducible (sobre $\mathbb{K}$)** si sus únicos polinomios divisores son triviales.*

Lema 3.17. Si $p \in \mathbb{K}[x]$ es irreducible y $p|p_1p_2$, entonces $p|p_1$ ó $p|p_2$.

Demostración. Supóngase que p no divide p_1 y sea $h = mcd(p, p_1)$. Como $h|p$, h debe ser un divisor trivial de p. Obsérvese que h no puede ser cp pues por hipótesis p no divide a p_1; entonces $h = c \in \mathbb{K}$. Más aún, como el máximo común divisor es único salvo por constantes, se puede tomar $h = 1$. Del corolario 3.15 se tiene $1 = \lambda_1 p + \lambda_2 p_1$ entonces $p_2 = \lambda_1 pp_2 + \lambda_2 p_1 p_2$. Luego, como p divide a ambos sumandos del lado derecho, p debe dividir a p_2. ∎

Corolario 3.18. Si $p \in \mathbb{K}[x]$ es irreducible y divide a $p_1p_2 \cdots p_n$, entonces p divide a p_i para algún $1 \leq i \leq n$.

Demostración. Por inducción sobre n.

1. Si $n = 1$. Entonces es inmediato que $p|p_1$.

2. Supóngase ahora que se cumple la propiedad para $n = k - 1$. Entonces, para $n = k$, p es irreducible y divide a $p_1 \cdots p_{k-1}p_k$. Por el lema anterior, $p|(p_1 \cdots p_{k-1}$ ó $p|p_k$. Si $p|p_k$ entonces $i = k$, en caso contrario, $p|(p_1 \cdots p_{k-1}$ y la hipótesis de inducción asegura que $p|p_i$ para algún $1 \leq i \leq k - 1$. ∎

Teorema 3.19. Sea $p \in \mathbb{K}[x]$ tal que $\mathbf{deg}(p) > 0$. Entonces se puede escribir como el producto

$$p = cp_1p_2 \cdots p_k$$

donde los p_i son polinomios mónicos irreducibles y $c = \mathbf{lc}(p) \in \mathbb{K}$. Esta factorización es única salvo el orden de los p_i's.

Demostración.

- **Existencia:** Por inducción sobre el grado de p, n.

  1. Si $n = 1$ los divisores de p deben tener grado 0 ó 1, pero entonces sus únicos divisores son los triviales y p es irreducible $\Rightarrow p = \mathbf{lc}(p)\, p_1$ con $p_1 = \left(\frac{1}{\mathbf{lc}(p)}p\right)$ mónico irreducible.
  2. Supóngase que $n = m$ y que se cumple el teorema para todo polinomio de grado menor que m. Si p es irreducible, $p = \mathbf{lc}(p)\left(\frac{1}{\mathbf{lc}(p)}p\right)$. Si p no es irreducible, entonces $p = q_1q_2$

con $0 < \mathbf{deg}(q_1), \mathbf{deg}(q_2) < m$, por la hipótesis de inducción $q_1 = c_1 p_1 \cdots p_k$ y $q_2 = c_2 p'_1 \cdots p'_{k'}$, con p_i, p'_i polinomios mónicos irreducibles $\Rightarrow p = (c_1 c_2) p_1 \cdots p_k p'_1 \cdots p'_{k'}$ es producto de polinomios mónicos irreducibles.

- **Unicidad:** Supóngase que $p = c_1 p_1 \cdots p_k = c_2 p'_1 \cdots p'_{k'}$. Sin pérdida de generalidad, suponga que $\mathbf{deg}(p_1) \leq \mathbf{deg}(p_2) \cdots \leq \mathbf{deg}(p_k)$. Evidentemente $p_1 | p = c_2 p'_1 \cdots p'_{k'}$ como p_1 es irreducible, por el corolario 3.18, debe dividir a alguno de los p'_i's. Suponga $p_1 | p'_1$ (reordenando los índices si es necesario). Como p'_1 también es irreducible mónico, $p_1 = p'_1$. Entonces $c_1 p_2 \cdots p_k = c_2 p'_2 \cdots p'_{k'}$. Aplicando el mismo proceso k veces, se llega a que $k' = k$ y que las factorizaciones son idénticas. ∎

Lema 3.20. Dado un ideal $I = \langle p \rangle \subseteq \mathbb{K}[x]$, I es maximal, si y sólo si, p es irreducible sobre $\mathbb{K}$.

Demostración. ($\Rightarrow$) Si I es maximal, cualquier ideal J tal que $I \subseteq J \subseteq \mathbb{K}[x]$ debe ser $J = I$ ó $J = \mathbb{K}[x]$. Sea q un divisor de $p \Rightarrow \langle p \rangle \subseteq \langle q \rangle \subseteq \mathbb{K}[x]$. Como $\langle p \rangle$ es maximal, $\langle q \rangle = \langle p \rangle$ ó $\langle q \rangle = \mathbb{K}[x] = \langle 1 \rangle$. En dichos casos, debe existir una constante $c \in \mathbb{K}$ tal que $q = cp$ ó $q = c \cdot 1 = c$, pero entonces q es un divisor trivial de p. Luego los únicos divisores de p son los triviales,$\Rightarrow$ p es irreducible.

($\Leftarrow$) Sea p irreducible. Sea J un ideal tal que $I \subseteq J \subseteq \mathbb{K}[x]$, como $\mathbb{K}[x]$ es un anillo de ideales principales, debe existir q tal que $J = \langle q \rangle$. De allí, $\langle p \rangle \subseteq \langle q \rangle \subseteq \langle 1 \rangle$, por lo que $q | p$, y siendo p irreducible, $q = cp$ o $q = c$ con $c \in \mathbb{K}$ constante $\Rightarrow \langle q \rangle = \langle p \rangle = I$ ó $\langle q \rangle = \langle 1 \rangle = \mathbb{K}[x]$, e $I = \langle p \rangle$ es maximal. ∎

Existe una dualidad entre los polinomios vistos como objetos formales de la estructura algebraica $\mathbb{K}[x]$ y vistos como funciones de $\mathbb{K} \to \mathbb{K}$, en donde la variable x puede ser sustituida por cualquier valor de $\mathbb{K}$ para obtener su imagen. A continuación se define este concepto de *función polinomial*. Por simplicidad, a lo largo de todo este trabajo se denotará tanto a un polinomio p como a la función polinomial que induce por el mismo símbolo. Al hacer referencia a la imagen de algún valor $a \in \mathbb{K}$ bajo p se denotará $p(a)$. Cuando se desee hacer explícito que p es un polinomio sobre la variable x, o bien que la función p es unaria y que depende de x, se denotará como $p(x)$.

Definición 3.21. *Dado un polinomio $p = \sum_{i=0}^{m} a_i x^i$, este induce una* ***función de valuación*** *$p : \mathbb{K} \to \mathbb{K}$ tal que $a \mapsto p(a)$, donde para cada valor $a \in \mathbb{K}$*

$$p(a) = \sum_{i=0}^{m} a_i a^i$$

De esta manera, se tiene un mapa entre $\mathbb{K}[x] \to \mathrm{Pol}(\mathbb{K}, \mathbb{K})$ que a cada polinomio le asigna una función polinomial con coeficientes en $\mathbb{K}$, $p : \mathbb{K} \to \mathbb{K}$.

Definición 3.22. *Dado un polinomio $p \in \mathbb{K}[x] \setminus \{0\}$, se dice que $a_0 \in \mathbb{K}$ es* ***una raíz o cero de*** *p, si $p(a_0) = 0$.*

Lema 3.23. Dados dos polinomios $p, q \in \mathbb{K}[x]$, si $p = q$ (como polinomios) entonces $p(a) = q(a),\ \forall a \in \mathbb{K}$ (como funciones).

Demostración. Sean $p = \sum_{i=0}^{m} a_i x^i$ y $q = \sum_{j=0}^{n} b_j x^j$. Como $p = q \Rightarrow p - q = 0$ y $\mathbf{deg}(p) = m = n = \mathbf{deg}(q)$ entonces $0 = p - q = \sum_{i=0}^{n} (a_i - b_i)x^i \Rightarrow a_i = b_i$ para todo $0 \leq i \leq n$. Luego, para cualquier $a \in \mathbb{K}$,

$$p(a) = \sum_{i=0}^{n} a_i\, a^i = \sum_{i=0}^{n} b_i\, a^i = q(a),\ \forall\ a \in \mathbb{K}.$$

■

***Ejemplo* 3.3**

Aunque parece tentador pensar que el recíproco del lema anterior también se cumple, no siempre es así, particularmente cuando el campo $\mathbb{K}$ es finito, pueden haber dos polinomios distintos que representen a la misma función polinomial.

Haciendo $\mathbb{K} = \mathbb{Z}_3$, los polinomios $p, q \in \mathbb{Z}_3[x]$ con $p = 2x^5$ y $q = x^3 + x$ son evidentemente distintos, sin embargo es sencillo verificar que las funciones que inducen coinciden para todos los valores de $\mathbb{Z}_3 = \{0, 1, 2\}$:

x	$p(x)$	$q(x)$
0	$2(0)^5 = 2 \cdot 0 = 0$	$(0)^3 + 0 = 0 + 0 = 0$
1	$2(1)^5 = 2 \cdot 1 = 2$	$(1)^3 + 1 = 1 + 1 = 2$
2	$2(2)^5 = 2 \cdot 2 = 1$	$(2)^3 + 2 = 2 + 2 = 1$

Teorema 3.24 (Del Residuo)**.** Dado un polinomio $p \in \mathbb{K}[x]$ no cero y $a_0 \in \mathbb{K}$ una constante cualquiera, entonces $p(a_0) = \mathbf{rem}(p, x - a_0)$.

Demostración. El algoritmo de la división asegura la existencia de c y $\mathbf{rem}(p, x - a_0)$ tales que $p = c(x - a_0) + \mathbf{rem}(p, x - a_0)$. Por el lema anterior, $p(a_0) = c(a_0)(a_0 - a_0) + \mathbf{rem}(p, x - a_0)\,(a_0) = \mathbf{rem}(p, x - a_0)\,(a_0)$. Sin embargo, el algoritmo de la división asegura que el residuo es cero o de grado menor que $x - a_0$. Como $\mathbf{deg}(x - a_0) = 1$, esto significa que el residuo es cero o de grado cero, una constante en cualquier caso. Entonces $\mathbf{rem}(p, x - a_0) \in \mathbb{K}$ y

$$p(a_0) = \mathbf{rem}(p, x - a_0)\,(a_0) = \mathbf{rem}(p, x - a_0)$$

■

Corolario 3.25 (Teorema del Factor)**.** Dado un polinomio $p \in \mathbb{K}[x]$ no cero con $\mathbf{deg}(p) \geq 1$, $a_0 \in \mathbb{K}$ es una raíz de p si y sólo si $x - a_0$ es divisor de p.

Demostración. a_0 es raíz de $p \Leftrightarrow p(a_0) = \mathbf{rem}(p, x - a_0) = 0 \Leftrightarrow p = c(x - a_0),\ c \in \mathbb{K}[x] \Leftrightarrow x - a_0 | p$. ■

Corolario 3.26. Un polinomio de grado n en $\mathbb{K}[x]$ no puede tener más de n ceros.

Demostración. Si $c_1, c_2, \cdots, c_r$ son ceros de $p \in \mathbb{K}[x]$ con $\mathbf{deg}(p) = n$ entonces $(x - c_1)(x - c_2)\cdots(x - c_r)$ debe dividir a p por el teorema del factor. Para que esto sea posible, el grado de este producto (r) debe ser menor o igual que el grado de p, por lo que $r \leq n$. ∎

Este resultado permite acotar el número de ceros que tiene un polinomio dado, sin embargo no se ha dicho nada sobre cuando un polinomio tiene raíces o no. Considere por ejemplo los polinomios $p = x - 1$, $q = x^2 - 1$ y $r = x^2 + 1$ en $\mathbb{R}[x]$, es evidente que $p(1) = q(1) = q(-1) = 0$, como p puede tener a lo sumo una raíz, esta única raíz es 1, y las únicas dos raíces de q son 1 y -1. Por otro lado, r tiene como máximo dos raíces, pero estas deben resolver la ecuación

$$x^2 + 1 = 0 \;\Leftrightarrow\; x^2 = -1$$

la cual no tiene solución en $\mathbb{R}$. Si se cambia el campo subyacente por $\mathbb{C}$, y se considera $r \in \mathbb{C}[x]$, r tiene exactamente dos raíces $r = \pm i$. Este resultado se conoce como el Teorema Fundamental del Álgebra, y se enuncia a continuación sin prueba. Puede revisarse su prueba en el libro de Fine (1997), que presenta seis diferentes pruebas para este teorema, usando, por ejemplo, análisis complejo, teoría de Galois y topología algebraica.

Teorema 3.27 (Teorema Fundamental del Álgebra)**.** Todo polinomio $p \in \mathbb{C}[x]$ de grado mayor o igual que uno, tiene al menos una raíz en $\mathbb{C}$.

Corolario 3.28. Dado $p \in \mathbb{C}[x]$ con $\mathbf{deg}(p) = n > 0$, p se descompone en n factores lineales

$$p = c(x - c_1)\cdots(x - c_n)$$

Demostración. Por el Teorema Fundamental del Álgebra, existe $c_1 \in \mathbb{C}$ tal que $(x - c_1)|p \Rightarrow\ p = (x - c_1)p_1$ con $\mathbf{deg}(p_1) = \mathbf{deg}(p) - \mathbf{deg}(x - c_1) = n - 1$. Repitiendo este proceso $n - 1$ veces más, se llega a $p = p_n(x - c_1)\cdots(x - c_n)$, donde $\mathbf{deg}(p_n) = 0$, *i.e.* $p = c(x - c_1)\cdots(x - c_n)$. ∎

Definición 3.29. *Si $(x - c)^m$ divide p pero no $(x - c)^{m+1}$, se dice que c es una **raíz de multiplicidad** m **de** p.*

Del corolario al Teorema Fundamental del Álgebra es evidente entonces el siguiente resultado.

Corolario 3.30. Cualquier $p \in \mathbb{C}[x]$ con $\mathbf{deg}(p) = n > 0$, tiene exactamente n raíces, si éstas se cuentan con multiplicidad.

4 ANILLOS DE POLINOMIOS EN VARIAS VARIABLES

Como se observó al final en la sección 3.1, a partir de cualquier dominio entero $\mathbb{D}$, es posible construir el anillo $\mathbb{D}[x]$, el cual, aplicando el teorema 3.7, es también un dominio entero. En el caso especial en el que $\mathbb{D} = \mathbb{K}[x]$, se construye el anillo $\mathbb{D}[y] = \mathbb{K}[x][y] = \mathbb{K}[x, y]$, *agregando* la variable y al anillo de polinomios $\mathbb{K}[x]$, construyendo así un anillo de polinomios en dos variables. Análogamente, para cualquier conjunto de variables $x_1, x_2, \cdots, x_n$ se puede construir el anillo $\mathbb{K}[x_1, x_2, \cdots, x_n]$ de polinomios en n variables.

En este capítulo, se formaliza la definición de este anillo y se estudian sus propiedades como generalizaciones de las de $\mathbb{K}[x]$. Al igual que en el capítulo anterior, el estudio de este anillo y sus ideales se centra en responder dos preguntas: (1) cómo determinar si un polinomio dado pertenece o no a un ideal, y (2) cómo son los elementos de los anillos cocientes que estos ideales inducen. Aunque $\mathbb{K}[x_1, \cdots, x_n]$ es una generalización directa de $\mathbb{K}[x]$, se verá que la solución a estos problemas no es tan sencilla como lo fue en el capítulo anterior, por lo que sólo será posible darles solución al final del Capítulo 6, luego de definir las *bases de Gröbner.*

4.1. Propiedades de $\mathbb{K}[x_1, \cdots, x_n]$

Definición 4.1. *Dado un campo $\mathbb{K}$, el **anillo de polinomios con coeficientes en $\mathbb{K}$ y variables** $x_1, \cdots, x_n$ es el conjunto $\mathbb{K}[x_1, \cdots, x_n] = \mathbb{K}_{n-1}[x_n]$ donde los anillos $\mathbb{K}_i$ se definen inductivamente como:*

$$\mathbb{K}_1 = \mathbb{K}[x_1] \text{ y } \mathbb{K}_k = \mathbb{K}_{k-1}[x_k] \text{ para } k \geq 2$$

*A los elementos de $\mathbb{K}[x_1, \cdots, x_n]$ se les llama **polinomios en las variables** $x_1, \cdots, x_n$ **y coeficientes en** $\mathbb{K}$. A cualquier producto de las variables $x_1, \cdots, x_n$ se le llama **monomio**, a un elemento de $\mathbb{K}$ que multiplica a un monomio se le llama **coeficiente** y al coeficiente junto al monomio se le denomina **término**.*

Observación. Definiendo la multiplicación de la variables como conmutativa, *i.e.* $xy = yx$ se puede mostrar que $\mathbb{K}[x, y] = \mathbb{K}[y, x]$, y más generalmente, si $x'_1, x'_2, \cdots, x'_n$ es cualquier permutación de $x_1, x_2, \cdots, x_n$ entonces $\mathbb{K}[x_1, \cdots, x_n] = \mathbb{K}[x'_1, \cdots, x'_n]$. Los elementos de $\mathbb{K}[x_1, \cdots, x_n]$ se pueden escribir de la forma:

$$p(x_1, \cdots, x_n) = \sum a_{(i_1, i_2, \cdots, i_n)} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n} = \sum_{\boldsymbol{i} \in \mathbb{N}^n} a_{\boldsymbol{i}} \boldsymbol{x}^{\boldsymbol{i}}$$

como una suma finita de términos, donde $\boldsymbol{i} = (i_1, \cdots, i_n) \in \mathbb{N}^n$ y $\boldsymbol{x^i} = x_1^{i_1} \cdots x_n^{i_n}$. La suma de polinomios se define al sumar los coeficientes de monomios iguales y el producto se puede calcular mediante la propiedad distributiva y la regla de exponentes para los monomios:

$$\left(x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}\right)\left(x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n}\right) = \left(x_1^{i_1+j_1} x_2^{i_2+j_2} \cdots x_n^{i_n+j_n}\right)$$

Nótese que $\mathbb{K} \subseteq \mathbb{K}[x_1, \cdots, x_n]$ y los elementos de $\mathbb{K}$ se denominan **constantes**. Más aún se tiene la cadena de inclusiones

$$\mathbb{K} \subseteq \mathbb{K}[x_1] \subseteq \mathbb{K}[x_1, x_2] \subseteq \cdots \subseteq \mathbb{K}[x_1, \cdots x_n].$$

Definición 4.2. *Dado un monomio $m \in \mathbb{K}[x_1, \cdots, x_n]$, $m = x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$, se definen el* ***multigrado de*** *m como la tupla* $\mathbf{mdeg}(m) = (i_1, i_2, \cdots, i_n) \subseteq \mathbb{N}^n$ *y el* ***grado de*** *m como* $\mathbf{deg}(m) = i_0 + i_1 + \cdots + i_n$. *El* ***grado de la variable*** *x_j* ***en el monomio*** *m es el exponente i_j.*

El grado de una variable x_j que no aparece en un monomio se define como cero. Particularmente, el multigrado de $1 = x_1^0 \cdots x_n^0$ es $\mathbf{mdeg}(1) = (0, 0, \cdots, 0)$ y su grado $\mathbf{deg}(1) = 0$.

Definición 4.3. *Para $p \in \mathbb{K}[x_1, \cdots, x_n] \setminus \{0\}$ se definen el* ***grado del polinomio*** *p,*$\mathbf{deg}(p)$*, como el mayor grado entre sus monomios, y el* ***grado de la variable*** *x_j* ***en el polinomio*** *p como el mayor grado de la variable x_j entre los monomios de p, y se denota por* $\mathbf{deg}_{x_i}(p)$.

Es importante notar que ahora es imposible definir un *grado* como el que se tenía para $\mathbb{K}[x]$. Aunque ahora se definen tres conceptos similares, no conducen directamente a la definición de término y coeficiente principales, pues puede darse el caso de que varios monomios distintos tengan el mismo grado, multigrado, o grado respecto a alguna variable. El concepto de *grado respecto a una variable* será de utilidad a continuación para probar algunas propiedades de $\mathbb{K}[x_1, \cdots, x_n]$ generalizando las de $\mathbb{K}[x]$ al considerarlo como $\mathbb{K}[x_1, \cdots, x_{n-1}][x_n]$, pero luego de eso se dejará de utilizar, pues en este caso los *coeficientes* son elementos de $\mathbb{K}[x_1, \cdots, x_{n-1}]$ que no es un campo, desechando de una vez cualquier intento por definir un algoritmo de la división similar al de $\mathbb{K}[x]$.

Teorema 4.4. Dado un campo $\mathbb{K}$, el anillo de polinomios en varias variables $\mathbb{K}[x_1, \cdots, x_n]$, es un anillo conmutativo con unidad. Más aún, $\mathbb{K}[x_1, \cdots, x_n]$ es un dominio entero.

Demostración. Por inducción sobre n:

1. Para $n = 1$, $\mathbb{K}[x_1]$ es un anillo conmutativo con unidad por la proposición 3.3. Luego, como $\mathbb{K}$ es un dominio entero, el teorema 3.7, asegura que $\mathbb{K}[x_1]$ también lo es.

2. Supóngase que la propiedad es cierta para $n = k \geq 2$. Luego, $\mathbb{D} = \mathbb{K}[x_1, \cdots, x_k]$ es un dominio entero. Aplicando de nuevo la propiedad 3.3 y el teorema 3.7, $\mathbb{D}[x_{k+1}]$ es un anillo conmutativo con unidad, más aún, un dominio entero. Como $\mathbb{D}[x_{k+1}] = \mathbb{K}[x_1, \cdots, x_k, x_{k+1}] = \mathbb{K}[x_1, \cdots, x_k, x_{k+1}]$, se tiene lo buscado. ∎

Proposición 4.5. Dados $p, q \in \mathbb{K}[x_1, \cdots, x_n] \setminus \{0\}$, se tiene que

$$\mathbf{deg}_{x_i}(p+q) \leq \text{máx}\left(\mathbf{deg}_{x_i}(p), \mathbf{deg}_{x_i}(q)\right) \text{ y}$$

$$\mathbf{deg}_{x_i}(pq) = \mathbf{deg}_{x_i}(p) + \mathbf{deg}_{x_i}(q)$$

para $i = 1, \cdots, n$para $i = 1, \cdots, n$.

Demostración. Sin pérdida de la generalidad considere $i = n$, pues siempre se pueden permutar las variables para garantizar que este sea el caso, entonces $\mathbb{K}[x_1, \cdots, x_n] = \mathbb{K}[x_1, \cdots, x_{n-1}][x_n] = \mathbb{D}[x_n]$. Luego, el grado $\mathbf{deg}_{x_n}()$ es simplemente el grado tomando $\mathbb{D} = \mathbb{K}[x_1, \cdots, x_{n-1}]$ como el dominio entero de coeficientes. Entonces, por las propiedades 3.5 y 3.6,

$$\mathbf{deg}_R(p+q) \leq \text{máx}(\mathbf{deg}_R(p), \mathbf{deg}_R(q)) \text{ y } \mathbf{deg}_R(pq) = \mathbf{deg}_R(p) + \mathbf{deg}_R(q)$$

equivalentemente

$$\mathbf{deg}_{x_n}(p+q) \leq \text{máx}(\mathbf{deg}_{x_n}(p), \mathbf{deg}_{x_n}(q)) \text{ y } \mathbf{deg}_{x_n}(pq) = \mathbf{deg}_{x_n}(p) + \mathbf{deg}_{x_n}(q). \quad \blacksquare$$

Corolario 4.6. Los únicos elementos de $\mathbb{K}[x_1, \cdots, x_n]$ con inverso multiplicativo son las constantes.

Demostración. Sea $a \in \mathbb{K} \subseteq \mathbb{K}[x_1, \cdots, x_n] \setminus \{0\}$, como $\mathbb{K}$ es un campo, existe $\frac{1}{a} \in \mathbb{K}$ tal que $a\frac{1}{a} = 1$. Para la otra implicación, sea $a \in \mathbb{K}[x_1, \cdots, x_n] \setminus \{0\}$ un elemento invertible, *i.e.* existe $b \in \mathbb{K}[x_1, \cdots, x_n] \setminus \{0\}$ tal que $ab = 1$ entonces $\mathbf{deg}_{x_i}(ab) = \mathbf{deg}_{x_i}(a) + \mathbf{deg}_{x_i}(b) = \mathbf{deg}_{x_i}(1) = 0$ por lo que $\mathbf{deg}_{x_i}(a) = 0$ para $i = 1, \cdots, n$. Como el grado de a respecto a todas las variables es cero, $a \in \mathbb{K}$ y es una constante. $\blacksquare$

El teorema 4.4 en conjunto con el teorema 2.8 permite construir el campo de cocientes

$$\mathbb{K}(x_1, \cdots, x_n) \supset \mathbb{K}[x_1, \cdots, x_n]$$

conocido como el **campo de funciones racionales en** $x_1, \cdots, x_n$ **sobre** $\mathbb{K}$, consistente de fracciones de polinomios en las variables $x_1, x_2, \cdots, x_n$. De esta forma es posible *dividir* polinomios en varias variables, cuando se extiende el espacio en el cual se está trabajando. El siguiente ejemplo muestra algunos de los inconvenientes para definir un algoritmo de la división en el caso de polinomios en $\mathbb{K}[x_1, \cdots x_n]$, como el que se obtuvo el capítulo anterior para polinomios en $\mathbb{K}[x]$.

***Ejemplo* 4.1**

El multigrado con el que se dotan a los polinomios de varias variables carece del *orden natural* que tiene el grado de polinomios sobre una sola variable. Dado que $\mathbb{K}[x, y] = \mathbb{K}[x][y] = \mathbb{K}[y][x]$ se podría usar el grado de x o de y para esto. Sean $p = x^2y^3 + xy^3 - 2x^2y^2 - y$, $q = x + y - 1$, al utilizar el mismo mecanismo del algoritmo de la división se obtiene:

- Al usar $\mathbf{deg}_y(\cdot)$

$$
\begin{array}{r|l}
 & (x^2+x)y^2-(x^3+2x^2-x)y+(x^4+x^3-3x^2+x-1) \\
\hline
y+(x-1) & (x^2+x)y^3-2x^2y^2-y \\
 & -(x^2+x)y^3-(x-1)(x^2+x)y^2 \\
\hline
 & (-x^3-2x^2+x)y^2-y \\
 & (x^3+2x^2-x)y^2+(x^3+2x^2-x)(x-1)y \\
\hline
 & (x^4+x^3-3x^2+x-1)y \\
 & -(x^4+x^3-3x^2+x-1)y-(x^4+x^3-3x^2+x-1)(x-1) \\
\hline
 & -x^5+4x^3-4x^2+2x-1
\end{array}
$$

- Al usar $\mathbf{deg}_x(\cdot)$

$$
\begin{array}{r|l}
 & (y^3-2y^2)x-(y^4-4y^3+2y^2) \\
\hline
x+(y-1) & (y^3-2y^2)x^2+y^3x-y \\
 & -(y^3-2y^2)x^2-(y^3-2y^2)(y-1)x \\
\hline
 & (-y^4+4y^3-2y^2)x-y \\
 & (y^4-4y^3+2y^2)x+(y^4-4y^3+2y^2)(y-1) \\
\hline
 & y^5-5y^4+6y^3-2y^2-y
\end{array}
$$

No solo en ambos casos se obtiene un resultado distinto, sino que al reducir el grado de y o de x el grado de la otra variable aumenta. Más aún, este ejemplo funciona sólo porque en cada paso fue posible encontrar un polinomio por el cual multiplicar a q para *eliminar* el término de mayor grado (respecto a y o x) de p, lo cual no siempre es posible, por ejemplo si $q = xy - 1$ hubiera sido imposible utilizar este camino respecto a $\mathbf{deg}_y(\cdot)$:

$$
\begin{array}{r|l}
 & (x+1)\,y^2 \\
\hline
xy-1 & (x^2+x)y^3-2x^2y^2-y \\
 & -(x^2+x)y^3+(x+1)\,y^2 \\
\hline
 & (-2x^2+x+1)\,y^2-y
\end{array}
$$

donde el proceso debe terminar aunque el grado del *residuo* no sea menor que el del divisor, pues $x \nmid (-2x^2+x+1)$.

4.2. Órdenes de monomios

El orden de los números naturales permite ordenar directamente los monomios de $\mathbb{K}[x]$, lo que se utilizó en el Capítulo 3 para definir el *término principal* de un polinomio, el cuál fue necesario para el algoritmo de la división (algoritmo 3.1). Para poder encontrar un algoritmo que de alguna forma imite el algoritmo de la división en $\mathbb{K}[x_1, \cdots, x_n]$ y que al restringirlo al caso

de una sola variable dé exactamente el algoritmo 3.1, es necesario introducir un *orden* para $\mathbb{N}^n$ que permita ordenar los monomios de $\mathbb{K}[x_1, \cdots, x_n]$ según su multigrado y que sea compatible con la multiplicación de monomios. En la discusión que sigue, se denotará por $\mathcal{M}^n$ al conjunto de monomios de $\mathbb{K}[x_1, \cdots, x_n]$, *i.e.*

$$\mathcal{M}^n = \{\, \boldsymbol{x}^{\boldsymbol{\alpha}} | \, \boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_n) \in \mathbb{N}^n \text{ y } \boldsymbol{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1} \cdots x_n^{\alpha_n} \}\,.$$

Definición 4.7 (Orden de monomios)**.** *Una relación $\prec$ sobre $\mathcal{M}^n \subseteq \mathbb{K}[x_1, \cdots, x_n]$, se dice que es un* ***orden de monomios en*** $\mathbb{K}[x_1, \cdots, x_n]$ *si cumple:*

1. *Para cualesquiera $\boldsymbol{x}^{\boldsymbol{\alpha}}, \boldsymbol{x}^{\boldsymbol{\beta}} \in \mathcal{M}^n$, $\boldsymbol{x}^{\boldsymbol{\alpha}} \prec \boldsymbol{x}^{\boldsymbol{\beta}}$, $\boldsymbol{x}^{\boldsymbol{\beta}} \prec \boldsymbol{x}^{\boldsymbol{\alpha}}$ ó $\boldsymbol{x}^{\boldsymbol{\alpha}} = \boldsymbol{x}^{\boldsymbol{\beta}}$ (i.e. $\prec$ es total).*
2. *Si $\boldsymbol{x}^{\boldsymbol{\alpha}} \prec \boldsymbol{x}^{\boldsymbol{\beta}}$ y $\boldsymbol{x}^{\boldsymbol{\beta}} \prec \boldsymbol{x}^{\boldsymbol{\gamma}}$ entonces $\boldsymbol{x}^{\boldsymbol{\alpha}} \prec \boldsymbol{x}^{\boldsymbol{\gamma}}$ (i.e. $\prec$ es transitiva).*
3. *$1 \prec \boldsymbol{x}^{\boldsymbol{\alpha}}$ para todo $\boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathcal{M}^n$, $\boldsymbol{x}^{\boldsymbol{\alpha}} \neq 1$ (1 es el mínimo respecto $\prec$).*
4. *Si $\boldsymbol{x}^{\boldsymbol{\alpha}} \prec \boldsymbol{x}^{\boldsymbol{\beta}}$ y $\boldsymbol{x}^{\boldsymbol{\gamma}} \in \mathcal{M}^n$ entonces $\boldsymbol{x}^{\boldsymbol{\alpha}}\boldsymbol{x}^{\boldsymbol{\gamma}} \prec \boldsymbol{x}^{\boldsymbol{\beta}}\boldsymbol{x}^{\boldsymbol{\gamma}}$ (i.e. $\prec$ es monótona respecto del producto).*

Proposición 4.8. Dados $\boldsymbol{x}^{\boldsymbol{\alpha}}, \boldsymbol{x}^{\boldsymbol{\beta}} \in \mathcal{M}^n$ tales que $\boldsymbol{x}^{\boldsymbol{\alpha}} | \boldsymbol{x}^{\boldsymbol{\beta}}$ entonces $\boldsymbol{x}^{\boldsymbol{\alpha}} \prec \boldsymbol{x}^{\boldsymbol{\beta}}$ o $\boldsymbol{x}^{\boldsymbol{\alpha}} = \boldsymbol{x}^{\boldsymbol{\beta}}$, para cualquier orden de monomios $\prec$.

Demostración. Sea $\prec$ un orden de monomios fijo. Sean $\boldsymbol{x}^{\boldsymbol{\alpha}}, \boldsymbol{x}^{\boldsymbol{\beta}} \in \mathcal{M}^n$ tales que $\boldsymbol{x}^{\boldsymbol{\alpha}} | \boldsymbol{x}^{\boldsymbol{\beta}}$, entonces existe $\boldsymbol{x}^{\boldsymbol{\gamma}} \in \mathcal{M}^n$ tal que $\boldsymbol{x}^{\boldsymbol{\beta}} = \boldsymbol{x}^{\boldsymbol{\alpha}}\boldsymbol{x}^{\boldsymbol{\gamma}}$. Si $\boldsymbol{x}^{\boldsymbol{\gamma}} = 1$ entonces $\boldsymbol{x}^{\boldsymbol{\alpha}} = \boldsymbol{x}^{\boldsymbol{\beta}}$ y se tiene el resultado. Supóngase entonces que $\boldsymbol{x}^{\boldsymbol{\gamma}} \neq 1$. Por la condición 3 de la definición 4.7, $1 \prec \boldsymbol{x}^{\boldsymbol{\gamma}}$ luego por la condición 4 se tiene $\boldsymbol{x}^{\boldsymbol{\alpha}} \prec \boldsymbol{x}^{\boldsymbol{\gamma}}\boldsymbol{x}^{\boldsymbol{\alpha}} = \boldsymbol{x}^{\boldsymbol{\beta}}$. ∎

A continuación se definen algunos de los órdenes de polinomios más utilizados para trabajar con polinomios en varias variables. La demostración de que cumplen con las condiciones de la definición 4.7 se puede encontrar en Cox *et al.* (2007).

Definición 4.9. *(Órdenes de Monomios usuales) Dados $\boldsymbol{x}^{\boldsymbol{\alpha}}, \boldsymbol{x}^{\boldsymbol{\beta}} \in \mathcal{M}^n$ con $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^n$, se definen los órdenes de monomios* ***lexicográfico*** *($\prec_{Lex}$),* ***lexicográfico con grado*** *($\prec_{DegLex}$) y* ***lexicográfico inverso con grado*** *($\prec_{DegRevLex}$) como:*

$$\boldsymbol{x}^{\boldsymbol{\alpha}} \prec_{Lex} \boldsymbol{x}^{\boldsymbol{\beta}} \Leftrightarrow \text{ en el vector } \boldsymbol{\beta} - \boldsymbol{\alpha} \text{ la primera entrada no cero es positiva.}$$

$$\boldsymbol{x}^{\boldsymbol{\alpha}} \prec_{DegLex} \boldsymbol{x}^{\boldsymbol{\beta}} \Leftrightarrow \begin{cases} \mathbf{deg}(\boldsymbol{x}^{\boldsymbol{\alpha}}) < \mathbf{deg}(\boldsymbol{x}^{\boldsymbol{\beta}}) & o \\ \mathbf{deg}(\boldsymbol{x}^{\boldsymbol{\alpha}}) = \mathbf{deg}(\boldsymbol{x}^{\boldsymbol{\beta}}) & y\ \boldsymbol{x}^{\boldsymbol{\alpha}} \prec_{Lex} \boldsymbol{x}^{\boldsymbol{\beta}} \end{cases}$$

$$\boldsymbol{x}^{\boldsymbol{\alpha}} \prec_{DegRevLex} \boldsymbol{x}^{\boldsymbol{\beta}} \Leftrightarrow \begin{cases} \mathbf{deg}(\boldsymbol{x}^{\boldsymbol{\alpha}}) < \mathbf{deg}(\boldsymbol{x}^{\boldsymbol{\beta}}) & o \\ \mathbf{deg}(\boldsymbol{x}^{\boldsymbol{\alpha}}) = \mathbf{deg}(\boldsymbol{x}^{\boldsymbol{\beta}}) & y \text{ en el vector } \boldsymbol{\beta} - \boldsymbol{\alpha} \\ \text{la última entrada no cero es positiva.} \end{cases}$$

Observaciones.

1. Note que los tres órdenes que se acaban de definir incluyen una comparación entre los multigrados de los polinomios, en la cual lo relevante es el signo de la primera o última entrada que no es cero. Esto significa que el orden de los monomios depende del orden en el que se consideren las variables. Por ejemplo, usando orden lexicográfico en $\mathbb{K}[x,y]$ y en $\mathbb{K}[y,x]$ para x y y se tiene:

   - Para $x,y \in \mathbb{K}[x,y] \ \Rightarrow x = x^1y^0,\ y = x^0y^1 \Rightarrow\ y \prec_{\text{Lex}} x$ pues $(1,0)-(0,1)=(1,-1)$ tiene primera coordenada no cero positiva.
   - Para $x,y \in \mathbb{K}[y,x] \ \Rightarrow x = y^0x^1,\ y = y^1x^0 \Rightarrow\ x \prec_{\text{Lex}} y$ pues $(1,0)-(0,1)=(1,-1)$ tiene primera coordenada no cero positiva

   El orden lexicográfico posee el inconveniente que aun cuando $\mathbb{K}[x,y] = \mathbb{K}[y,x]$ (como anillos), sus órdenes lexicográficos son distintos. Para evitar este tipo de percances, es necesario identificar cuál orden lexicográfico se está utilizando, entendiéndose cuando no se haga la distinción explícita que $y \prec_{\text{Lex}} x$, el orden usual de $\mathbb{K}[x,y]$. Así, en la definición anterior, se está dando la definición para cuando $x_n \prec x_2 \prec \cdots \prec x_1$, caso que se sobreentenderá a menos que se especifique lo contrario. Esta misma observación es válida para los tres órdenes recién definidos.

2. En el orden lexicográfico es importante notar que una variable domina a todas las variables después de ella, sin importar el grado que tengan. Si $\boldsymbol{x}^{\boldsymbol{\alpha}} = x_i = \boldsymbol{x}^{(0,\cdots,0,1,0,\cdots,0)}$ y $\boldsymbol{x}^{\boldsymbol{\beta}} = x_{i+1}^{\beta_{i+1}} x_{i+2}^{\beta_{i+2}} \cdots x_n^{\beta_n} = \boldsymbol{x}^{(0,\cdots,0,\beta_{i+1},\beta_{i+2},\cdots,\beta_n)}$ entonces

$$\boldsymbol{\alpha} - \boldsymbol{\beta} = (0,\cdots,0,1,-\beta_{i+1},-\beta_{i+2},\cdots,-\beta_n) \text{ por lo que}$$

$$\boldsymbol{x}^{\boldsymbol{\beta}} \prec_{\text{Lex}} \boldsymbol{x}^{\boldsymbol{\alpha}} \Leftrightarrow\ x_{i+1}^{\beta_{i+1}} x_{i+2}^{\beta_{i+2}} \cdots x_n^{\beta_n} \prec_{\text{Lex}} x_i$$

3. En los órdenes DegLex y DegRevLex en cambio, el grado total de los monomios se utiliza para ordenarlos, y únicamente cuando hay empates se resuelven con Lex (para DegLex) y revisando el signo de la última entrada no cero de la resta de los multigrados para DegRevLex.

4. De aquí en adelante se utilizará solo un orden de monomios a la vez, por lo que a menos de que haya posibilidad de confusiones se suprimirán los subíndices de los símbolos $\prec$.

***Ejemplo* 4.2**

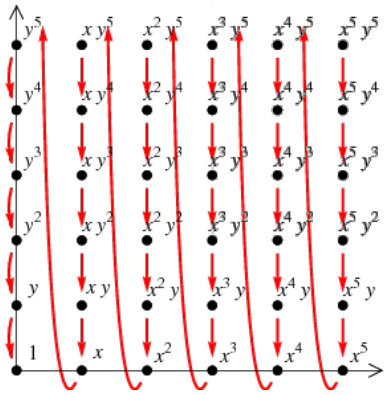
A continuación se muestra cómo quedan ordenados los monomios de $\mathbb{K}[x,y]$ en los tres órdenes definidos. Las figuras 4.1 y 4.2 muestran los mismos órdenes de forma gráfica para $\mathbb{K}[x,y]$ y $\mathbb{K}[x,y,z]$, las flechas apuntan en dirección al monomio más pequeño.

- En Lex:

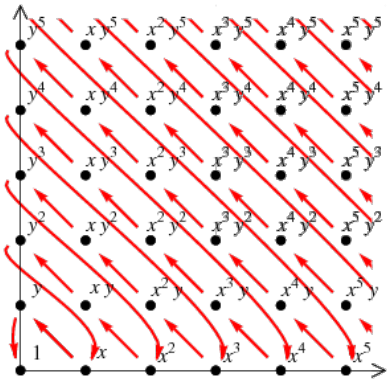
$$1 \prec y \prec y^2 \prec y^3 \prec \cdots \prec x \prec xy \prec xy^2 \prec \cdots \prec x^2 \prec x^2y \prec x^2y^2 \prec \cdots \prec \cdots$$

Figura 4.1: Órdenes de monomios en $\mathbb{K}[x, y]$ con $y \prec x$

(a) Orden Lexicográfico



(b) Orden Lexicográfico con grado (usual e inverso)



- DegLex y DegRevLex resultan ser el mismo orden:

$$1 \prec y \prec x \prec y^2 \prec xy \prec x^2 \prec y^3 \prec xy^2 \prec x^2y \prec x^3 \prec \cdots$$

Además de estos tres órdenes monomiales y de los $n!$ que resultan al permutar el orden de las variables en $\mathbb{K}[x_1, \cdots, x_n]$, se pueden generar muchos más. Así, por ejemplo, se pueden crear **órdenes de bloque** al separar las variables en dos bloques y comparar cada bloque con un orden distinto, o, cómo se hizo con DegLex, se le puede agregar una comparación de grados como primer paso a otro orden para crear un **orden con grado** Adams y Loustaunau (1994).

El siguiente teorema permite caracterizar todos los órdenes monomiales como conjuntos de vectores ortogonales. Aunque sea de utilidad poder caracterizar los órdenes monomiales pues, como se verá más adelante, las propiedades de las bases de Gröbner dependen del orden específico que se esté utilizando; en lo que resta de este estudio y para el problema que se busca resolver, basta con las ordenaciones introducidas anteriormente. Así, este teorema se incluye únicamente como una curiosidad sin demostración. Para más información sobre el tema, refiérase a Adams y Loustaunau (1994) o a Fröberg (1997).

Teorema 4.10. Todo orden de monomios en $\mathcal{M}^n$ está definido por una secuencia de vectores ortogonales $u_1, u_2, \cdots, u_r \in \mathbb{R}^n$ con $r \leq n$ como :

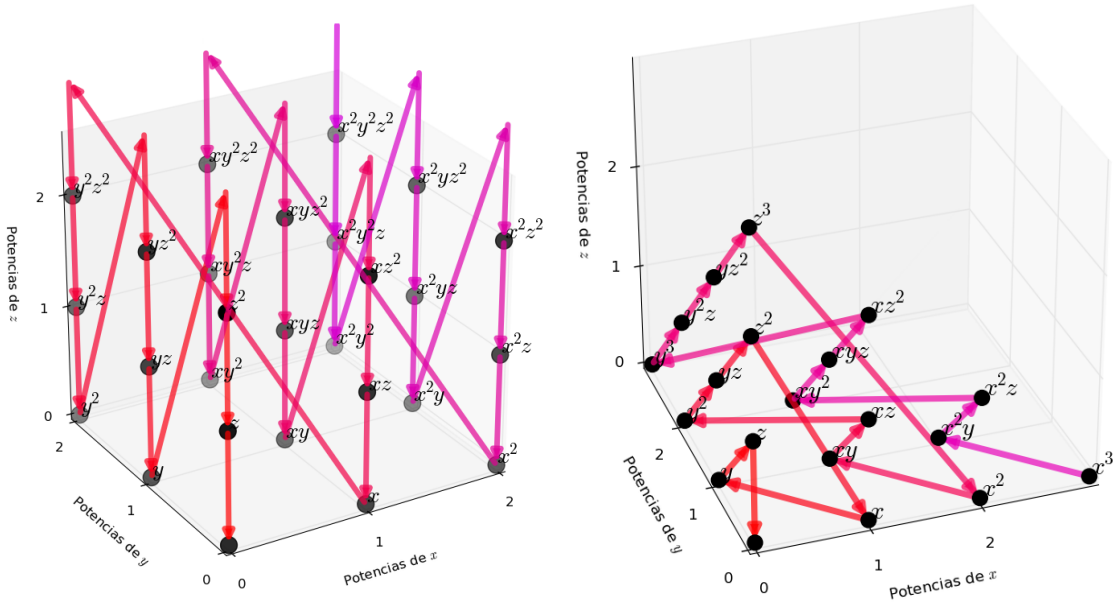
$$\boldsymbol{x}^{\boldsymbol{\alpha}} \prec \boldsymbol{x}^{\boldsymbol{\beta}} \Leftrightarrow \begin{cases} u_1 \cdot \boldsymbol{\alpha} < u_1 \cdot \boldsymbol{\beta} \ \text{ ó} \\ u_1 \cdot \boldsymbol{\alpha} = u_1 \cdot \boldsymbol{\beta} \ \text{ y } u_2 \cdot \boldsymbol{\alpha} < u_2 \cdot \boldsymbol{\beta} \ \text{ ó} \\ u_1 \cdot \boldsymbol{\alpha} = u_1 \cdot \boldsymbol{\beta} \ \text{ y } u_2 \cdot \boldsymbol{\alpha} = u_2 \cdot \boldsymbol{\beta} \ \text{ y } u_3 \cdot \boldsymbol{\alpha} < u_3 \cdot \boldsymbol{\beta} \ \text{ ó} \\ \quad \vdots \\ u_1 \cdot \boldsymbol{\alpha} = u_1 \cdot \boldsymbol{\beta} \ \text{ y } u_2 \cdot \boldsymbol{\alpha} = u_2 \cdot \boldsymbol{\beta} \ \text{ y } \cdots \text{ y } u_{r-1} \cdot \boldsymbol{\alpha} = u_{r-1} \cdot \boldsymbol{\beta} \ \text{ y } u_r \cdot \boldsymbol{\alpha} < u_r \cdot \boldsymbol{\beta} \end{cases}$$

Donde $u_i \cdot \boldsymbol{\alpha}$ denota al producto punto usual de $\mathbb{R}^n$, *i.e.* $u_i \cdot \boldsymbol{\alpha} = u_{i1}\alpha_1 + u_{i2}\alpha_2 + \cdots + u_{in}\alpha_n$.

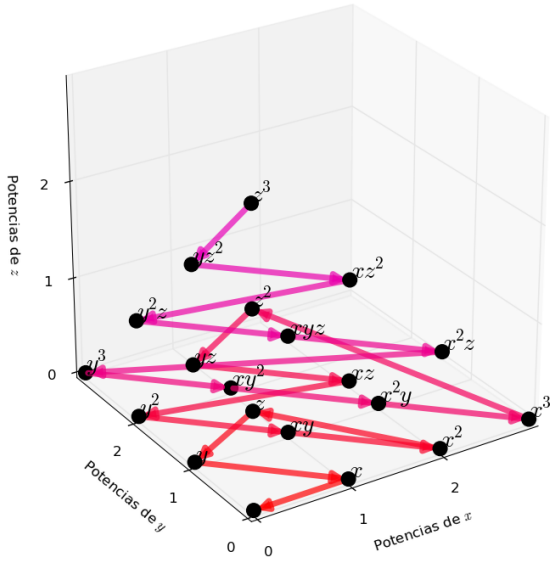
Figura 4.2: Órdenes de monomios en $\mathbb{K}[x, y, z]$ con $z \prec y \prec x$

(a) Orden Lexicográfico

(b) Orden Lexicográfico con grado



(c) Orden Lexicográfico inverso con grado



***Ejemplo* 4.3**

- El orden lexicográfico está definido por los n vectores $u_1 = (1, 0, 0, \cdots, 0), u_2 = (0, 1, 0, \cdots, 0), \cdots,$ $u_n = (0, 0, 0, \cdots, 1)$.
- DegLex está definido por los n vectores $u_1 = (1, 1, 1, \cdots, 1), u_2 = (n-1, -1, -1, \cdots, -1),$ $u_3 = (0, n-2, -1, \cdots, -1), \cdots, u_n = (0, 0, 0, \cdots, 1, -1)$. Donde es evidente que $u_1 = (1, 1, 1, \cdots, 1)$ caracteriza la comparación inicial de grados.
- DegRevLex está definido por los n vectores $u_1 = (1, 1, \cdots, 1, 1), u_2 = (1, 1, \cdots, 1, 1-n),$ $u_3 = (1, 1, \cdots, 1, 2-n, 0), \cdots, u_n = (-1, 0, 0, \cdots, 0)$.

Definición 4.11. *Sea $\prec$ un orden de monomios fijo en $\mathcal{M}^n$. Para $p \in \mathbb{K}[x_1, \cdots, x_n]$ se definen (respecto $\prec$)*

- *El **monomio principal de** p, $\mathbf{lm}(p)$, como el mayor monomio (según $\prec$) de p.*
- *El **término principal de** p, $\mathbf{lt}(p)$, como el término que le corresponde al monomio principal de p.*
- *El **coeficiente principal de** p, $\mathbf{lc}(p)$, como el coeficiente del término principal de p.*
- *El **multigrado de** p, $\mathbf{mdeg}(p)$, como el multigrado del monomio principal de p.*

Lema 4.12. Dados $p, q \in \mathbb{K}[x_1, \cdots, x_n] \setminus \{0\}$, se cumplen [1]

I) $\mathbf{mdeg}(p+q) \leq \text{máx}(\mathbf{mdeg}(p), \mathbf{mdeg}(q))$.

II) $\mathbf{mdeg}(pq) = \mathbf{mdeg}(p) + \mathbf{mdeg}(q)$.

III) $\mathbf{lt}(pq) = \mathbf{lt}(p)\,\mathbf{lt}(q)$.

Demostración. Sean $p, q \in \mathbb{K}[x_1, \cdots, x_n] \setminus \{0\}$. La propiedad I) es evidente, pues los monomios de $p+q$ provienen de p o de q. Para II) y III), escríbanse p y q como:

$$p = a_0 + a_1 \boldsymbol{x}^{\boldsymbol{\alpha}_1} + a_2 \boldsymbol{x}^{\boldsymbol{\alpha}_2} + \cdots + a_k \boldsymbol{x}^{\boldsymbol{\alpha}_k},\ a_k \neq 0,$$

$$q = b_0 + b_1 \boldsymbol{x}^{\boldsymbol{\beta}_1} + + b_2 \boldsymbol{x}^{\boldsymbol{\beta}_2} + \cdots + b_r \boldsymbol{x}^{\boldsymbol{\beta}_r},\ b_r \neq 0$$

con $1 \prec \boldsymbol{x}^{\boldsymbol{\alpha}_1} \prec \boldsymbol{x}^{\boldsymbol{\alpha}_2} \prec \cdots \prec \boldsymbol{x}^{\boldsymbol{\alpha}_k}$ y $1 \prec \boldsymbol{x}^{\boldsymbol{\beta}_1} \prec \boldsymbol{x}^{\boldsymbol{\beta}_2} \prec \cdots \prec \boldsymbol{x}^{\boldsymbol{\beta}_r}$, por lo que $\mathbf{mdeg}(p) = \boldsymbol{\alpha}_k$ y $\mathbf{mdeg}(q) = \boldsymbol{\beta}_r$. En pq, por las propiedades de distributividad, aparecerán todos los monomios de la forma $\boldsymbol{x}^{\boldsymbol{\alpha}_i}\boldsymbol{x}^{\boldsymbol{\beta}_j} = \boldsymbol{x}^{\boldsymbol{\alpha}_i + \boldsymbol{\beta}_j}$, cuyos coeficientes al simplificar no se vuelvan cero. Particularmente, $\boldsymbol{x}^{\boldsymbol{\alpha}_k + \boldsymbol{\beta}_r}$ es uno de los candidatos a estar en la expansión de pq. Si $\boldsymbol{x}^{\boldsymbol{\alpha}_i} \prec \boldsymbol{x}^{\boldsymbol{\alpha}_k}$ y $\boldsymbol{x}^{\boldsymbol{\beta}_j} \prec \boldsymbol{x}^{\boldsymbol{\beta}_r}$ entonces $\boldsymbol{x}^{\boldsymbol{\alpha}_i}\boldsymbol{x}^{\boldsymbol{\beta}_j} \prec \boldsymbol{x}^{\boldsymbol{\alpha}_k}\boldsymbol{x}^{\boldsymbol{\beta}_j}$ y $\boldsymbol{x}^{\boldsymbol{\alpha}_k}\boldsymbol{x}^{\boldsymbol{\beta}_j} \prec \boldsymbol{x}^{\boldsymbol{\alpha}_k}\boldsymbol{x}^{\boldsymbol{\beta}_r}$ por lo que $\boldsymbol{x}^{\boldsymbol{\alpha}_i}\boldsymbol{x}^{\boldsymbol{\beta}_j} \prec \boldsymbol{x}^{\boldsymbol{\alpha}_k}\boldsymbol{x}^{\boldsymbol{\beta}_r}$, lo que significa que el monomio $\boldsymbol{x}^{\boldsymbol{\alpha}_k + \boldsymbol{\beta}_r}$ sólo se puede generar al multiplicar los términos principales de p y q, más

[1] Se entiende $\mathbf{mdeg}(a) \leq \mathbf{mdeg}(b) \Leftrightarrow \boldsymbol{x}^{\mathbf{mdeg}(a)} \preceq \boldsymbol{x}^{\mathbf{mdeg}(b)}$.

aún, significa que cualquier otro monomio del producto pq es menor que este. Como $\mathbf{lc}\left(\boldsymbol{x}^{\boldsymbol{\alpha}_k+\boldsymbol{\beta}_r}\right) = a_k b_r \neq 0$, este monomio es el mayor que aparece en pq y se tiene:

$$\mathbf{mdeg}(pq) = \boldsymbol{\alpha}_k + \boldsymbol{\beta}_r = \mathbf{mdeg}(p) + \mathbf{mdeg}(q)\,,$$

$$\mathbf{lt}(pq) = a_k b_r \boldsymbol{x}^{\boldsymbol{\alpha}_k+\boldsymbol{\beta}_r} = \left(a_k \boldsymbol{x}^{\boldsymbol{\alpha}_k}\right)\left(b_r \boldsymbol{x}^{\boldsymbol{\beta}_r}\right) = \mathbf{lt}(p)\,\mathbf{lt}(q)\,. \qquad \blacksquare$$

De ahora en adelante, a menos que se diga lo contrario, se supondrá que se trabaja con un orden de monomios cualquiera fijo $\prec$, y que **mdeg**(), **lt**(), **lc**() y **lm**() se calculan respecto este orden.

4.3. Lema de Dickson

Definición 4.13. *Sea $I \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un ideal. Se dice que I es un **ideal de monomios** cuando existe un conjunto $S \subseteq \mathcal{M}^n$ de monomios que lo genera $I = \langle S \rangle$.*

Observación. $\mathbb{K}[x_1, \cdots, x_n] = \langle x_1, \cdots, x_n \rangle$ es un ideal de monomios. Es importante notar que, contrario a lo que sugiere el nombre, un ideal de monomios no contiene únicamente monomios, pues debe contener todas las combinaciones lineales de sus elementos. Los únicos casos que cumplen con ser ideales formados únicamente de monomios son los ideales de la forma $\langle x_i \rangle$.

Lema 4.14. Sea $I = \langle S \rangle \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un ideal de monomios con $S \subseteq \mathcal{M}^n$. Entonces, un monomio $\boldsymbol{x}^{\boldsymbol{\beta}} \in I$ si y sólo si $\boldsymbol{x}^{\boldsymbol{\alpha}} | \boldsymbol{x}^{\boldsymbol{\beta}}$ para algún $\boldsymbol{x}^{\boldsymbol{\alpha}} \in S$.

Demostración. ($\longleftarrow$) Si $\boldsymbol{x}^{\boldsymbol{\beta}}$ es divisible por $\boldsymbol{x}^{\boldsymbol{\alpha}} \in S$, entonces pertenece a $\langle S \rangle$ por definición.

($\longrightarrow$) Si $\boldsymbol{x}^{\boldsymbol{\beta}} \in I = \langle S \rangle$ entonces $\boldsymbol{x}^{\boldsymbol{\beta}} = \sum h_i \boldsymbol{x}^{\boldsymbol{\alpha}_i}$ con $h_i \in \mathbb{K}[x_1, \cdots, x_n]$ y $\boldsymbol{x}^{\boldsymbol{\alpha}_i} \in S$. Si se expandieran los h_i's en sus términos, se vería que cada monomio de la expresión resultante es mútiplo de algún $\boldsymbol{x}^{\boldsymbol{\alpha}_i} \in S$. Como $\boldsymbol{x}^{\boldsymbol{\beta}}$ debe ser igual a esta expresión luego de simplificar términos semejantes entonces también debe aparecer en este punto antes de simplificar, por lo que es múltiplo de algún $\boldsymbol{x}^{\boldsymbol{\alpha}_i}$. $\blacksquare$

Lema 4.15. Sea $I \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un ideal de monomios. Sea $p = a_0 + a_1 \boldsymbol{x}^{\boldsymbol{\alpha}_1} + \cdots + a_m^{\boldsymbol{\alpha}_m}$ con $a_i \in \mathbb{K} \setminus \{0\}$. Si $p \in I$, entonces todos los monomios de p pertenecen también a I, *i.e.* $\boldsymbol{x}^{\boldsymbol{\alpha}_i} \in I$ para todo i.

Demostración. Sea $S \subseteq \mathcal{M}^n$ tal que $I = \langle S \rangle$. Si $p \in I$ entonces $p = g_1 m_1 + g_2 m_2 + \cdots + g_k m_k$ con $m_i \in S$ para algunos polinomios $g_i = \sum a_{ij} \boldsymbol{x}^{\boldsymbol{\alpha}_{ij}} \in \mathbb{K}[x_1, \cdots, x_n]$. Entonces

$$a_0 + a_1 \boldsymbol{x}^{\boldsymbol{\alpha}_1} + \cdots + a_m \boldsymbol{x}^{\boldsymbol{\alpha}_m} = \left(\sum a_{1j} \boldsymbol{x}^{\boldsymbol{\alpha}_{1j}}\right) m_1 + \cdots + \left(\sum a_{kj} \boldsymbol{x}^{\boldsymbol{\alpha}_{kj}}\right) m_k$$

por lo que $a_i \boldsymbol{x}^{\boldsymbol{\alpha}_i}$ en el lado izquierdo de la igualdad, debe ser la suma de todos los monomios $a_{ij} \boldsymbol{x}^{\boldsymbol{\alpha}_{ij}} m_i$ del lado derecho que tengan el mismo multigrado $\boldsymbol{\alpha}_i$. Por lo tanto, $a_i \boldsymbol{x}^{\boldsymbol{\alpha}_i}$ es una combinación lineal de m_i's, entonces $\boldsymbol{x}^{\boldsymbol{\alpha}_i} \in I$, para todo $i = 1, \cdots k$. $\blacksquare$

Definición 4.16 (MCD y MCM de monomios)**.** *Sean* $\boldsymbol{x^{\alpha}}, \boldsymbol{x^{\beta}} \in \mathcal{M}^n$ *con* $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_n)$*, y* $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_n)$*, entonces se definen:*

1. *El* ***máximo común divisor de*** $\boldsymbol{x^{\alpha}}$ ***y*** $\boldsymbol{x^{\beta}}$ *como* $mcd\big(\boldsymbol{x^{\alpha}}, \boldsymbol{x^{\beta}}\big) = \boldsymbol{x^{\gamma}}$
   donde $\boldsymbol{\gamma} = (\text{mín}(\alpha_1, \beta_1), \cdots, \text{mín}(\alpha_n, \beta_n))$

2. *El* ***mínimo común múltiplo de*** $\boldsymbol{x^{\alpha}}$ ***y*** $\boldsymbol{x^{\beta}}$ *como* $mcm\big(\boldsymbol{x^{\alpha}}, \boldsymbol{x^{\beta}}\big) = \boldsymbol{x^{\gamma}}$
   donde $\boldsymbol{\gamma} = (\text{máx}(\alpha_1, \beta_1), \cdots, \text{máx}(\alpha_n, \beta_n))$

Observación. Esta definición coincide con la usual para máximo común divisor y mínimo común múltiplo en cualquier anillo en los que existan (véase Herstein, 1975). Sin embargo, la exposición de los mismos se limita al caso de monomios, pues es lo necesario para el presente trabajo. En este caso, la misma definición provee el algoritmo para encontrar el MCM y el MCD de dos monomios, haciendo n comparaciones de sus exponentes. La siguiente propiedad constituye una caracterización importante del máximo común divisor y mínimo común múltiplo. Se enuncia únicamente como referencia, la demostración de que es equivalente a la definición 4.16, puede encontrarse en Herstein (1975).

Proposición 4.17. Sean $\boldsymbol{x^{\alpha}}, \boldsymbol{x^{\beta}} \in \mathcal{M}^n$.

1. Si $\boldsymbol{x^{\gamma}} \in \mathcal{M}^n$ cumple con:

   I) $\boldsymbol{x^{\alpha}} | \boldsymbol{x^{\gamma}}$ y $\boldsymbol{x^{\beta}} | \boldsymbol{x^{\gamma}}$

   II) Para todo $\boldsymbol{x^{\delta}}$ tal que $\boldsymbol{x^{\alpha}} | \boldsymbol{x^{\delta}}$ y $\boldsymbol{x^{\beta}} | \boldsymbol{x^{\delta}}$, vale $\boldsymbol{x^{\gamma}} | \boldsymbol{x^{\delta}}$.

   Entonces $\boldsymbol{x^{\gamma}} = mcm\big(\boldsymbol{x^{\alpha}}, \boldsymbol{x^{\beta}}\big)$.

2. Si $\boldsymbol{x^{\gamma}} \in \mathcal{M}^n$ cumple con:

   I) $\boldsymbol{x^{\gamma}} | \boldsymbol{x^{\alpha}}$ y $\boldsymbol{x^{\gamma}} | \boldsymbol{x^{\beta}}$

   II) Para todo $\boldsymbol{x^{\delta}}$ tal que $\boldsymbol{x^{\delta}} | \boldsymbol{x^{\alpha}}$ y $\boldsymbol{x^{\delta}} | \boldsymbol{x^{\beta}}$, vale $\boldsymbol{x^{\delta}} | \boldsymbol{x^{\gamma}}$.

   Entonces $\boldsymbol{x^{\gamma}} = mcd\big(\boldsymbol{x^{\alpha}}, \boldsymbol{x^{\beta}}\big)$.

Lema 4.18. Sea $I_0 \subseteq I_1 \subseteq I_2 \subseteq \cdots \subseteq \mathbb{K}[x_1, \cdots, x_n]$ una cadena de ideales de monomios. Entonces $I = \bigcup_{i=0}^{\infty} I_i$ es también un ideal de monomios.

Demostración. Sean $S_0, S_1, S_2, \cdots \subseteq \mathcal{M}^n$ conjuntos de monomios tales que $I_i = \langle S_i \rangle$ para $i = 0, 1, \cdots$. Sea $S = \bigcup_{i=0}^{\infty} S_i$. A probar: $I = \langle S \rangle$.

Sea $p \in I$, entonces existe $k \in \mathbb{N}$ tal que $p \in I_k = \langle S_k \rangle$. Como $S_k \subseteq S$, entonces $\langle S_k \rangle \subseteq \langle S \rangle \Rightarrow p \in \langle S \rangle$. Entonces $I \subseteq \langle S \rangle$.

Por otro lado, sea $p \in \langle S \rangle$. Entonces $p = g_1 s_1 + \cdots + g_r s_r$ con $g_i \in \mathbb{K}[x_1, \cdots, x_n]$ y $s_i \in S$ para $i = 1, \cdots, r$. Como los s_i's son elementos de S, para cada subíndice i existe $k_i \in \mathbb{N}$ tal que

$s_i \in S_{k_i}$. Luego, $g_i s_i \in \langle S_{k_i} \rangle = I_{k_i} \subseteq I$ para todo $i = 1, \cdots, r$. Entonces $p \in I$ y $\langle S \rangle \subseteq I$. Por lo tanto, $I = \langle S \rangle$ e I es un ideal de monomios. ∎

Lema 4.19. Sean $I \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un ideal de monomios y $\boldsymbol{x^\alpha} \in \mathcal{M}^n$ un monomio cualquiera. Entonces, el ideal $(I : \langle \boldsymbol{x^\alpha} \rangle)$ es de monomios.

Demostración. Sea $S \subseteq \mathcal{M}^n$ tal que $I = \langle S \rangle$, y sea

$$T = \left\{ \frac{s}{mcd(s, \boldsymbol{x^\alpha})} \middle| \ s \in S \right\} \subseteq \mathcal{M}^n.$$

A probar: $(I : \langle \boldsymbol{x^\alpha} \rangle) = \langle T \rangle$

Sea $p \in \langle T \rangle \Rightarrow p = g_1 \frac{s_1}{mcd(s_1, \boldsymbol{x^\alpha})} + \cdots + g_k \frac{s_k}{mcd(s_k, \boldsymbol{x^\alpha})}$, con $g_i \in \mathbb{K}[x_1, \cdots, x_n]$ para $i = 1, \cdots, k$. Entonces

$$\begin{aligned} p\boldsymbol{x^\alpha} &= g_1 \frac{s_1}{mcd(s_1, \boldsymbol{x^\alpha})} \boldsymbol{x^\alpha} + \cdots + g_k \frac{s_k}{mcd(s_k, \boldsymbol{x^\alpha})} \boldsymbol{x^\alpha} \\ &= \left(g_1 \frac{\boldsymbol{x^\alpha}}{mcd(s_1, \boldsymbol{x^\alpha})} \right) s_1 + \cdots + \left(g_k \frac{\boldsymbol{x^\alpha}}{mcd(s_k, \boldsymbol{x^\alpha})} \right) s_k = g_1' s_1 + \cdots + g_k' s_k \end{aligned}$$

donde, como $mcd(s_i, \boldsymbol{x^\alpha}) \,|\boldsymbol{x^\alpha}$ se tiene $g_i' = g_i \frac{\boldsymbol{x^\alpha}}{mcd(s_i, \boldsymbol{x^\alpha})} \in \mathbb{K}[x_1, \cdots, x_n]$ para todo $i = 1, \cdots, k$. Entonces $p\boldsymbol{x^\alpha} \in \langle S \rangle = I \Rightarrow \ p \in (I : \langle \boldsymbol{x^\alpha} \rangle)$.

Por otro lado, si $p = b_0 + b_1 \boldsymbol{x^{\beta_1}} + \cdots + b_m \boldsymbol{x^{\beta_m}} \in (I : \langle \boldsymbol{x^\alpha} \rangle)$, entonces $p\boldsymbol{x^\alpha} \in I$. Por ser I un ideal de monomios, cada monomio $\boldsymbol{x^{\beta_i}} \boldsymbol{x^\alpha}$ de $p\boldsymbol{x^\alpha}$ pertenece a I. Entonces, existe $s_j \in S$ tal que $s_j | \left(\boldsymbol{x^{\beta_i}} \boldsymbol{x^\alpha} \right) \Rightarrow$ existe un monomio $m \in \mathcal{M}^n$ tal que $s_j m = \boldsymbol{x^{\beta_i}} \boldsymbol{x^\alpha}$. Luego,

$$\boldsymbol{x^{\beta_i}} \left(\frac{\boldsymbol{x^\alpha}}{mcd(s_j, \boldsymbol{x^\alpha})} \right) = m \frac{s_j}{mcd(s_j, \boldsymbol{x^\alpha})}.$$

Nótese que $\frac{\boldsymbol{x^\alpha}}{mcd(s_j, \boldsymbol{x^\alpha})}$ y $\frac{s_j}{mcd(s_j, \boldsymbol{x^\alpha})}$ no tienen ningún factor mayor que 1 en común, pues en caso contrario, $mcd(s_j, \boldsymbol{x^\alpha})$ no sería su máximo común divisor. De donde, como

$$\frac{\boldsymbol{x^\alpha}}{mcd(s_j, \boldsymbol{x^\alpha})} \Bigg| m \frac{s_j}{mcd(s_j, \boldsymbol{x^\alpha})},$$

debe darse que

$$\frac{\boldsymbol{x^\alpha}}{mcd(s_j, \boldsymbol{x^\alpha})} \Bigg| m.$$

Entonces

$$\boldsymbol{x^{\beta_i}} = \left(\frac{m}{\boldsymbol{x^\alpha} / mcd(s_j, \boldsymbol{x^\alpha})} \right) \frac{s_j}{mcd(s_j, \boldsymbol{x^\alpha})} \in \langle T \rangle$$

para todo $i = 1, \cdots, k$, por lo que $p \in \langle T \rangle$. Por lo tanto, $(I : \langle \boldsymbol{x^\alpha} \rangle) = \langle T \rangle$ y es un ideal de monomios como se buscaba. ∎

Lema 4.20. Sea $I \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un ideal de monomios. Entonces el ideal

$$I \cap \mathbb{K}[x_1, \cdots, x_{n-1}] \subseteq \mathbb{K}[x_1, \cdots, x_{n-1}],$$

es un ideal de monomios de $\mathbb{K}[x_1, \cdots, x_{n-1}]$.

Demostración. Sea $S \subseteq \mathcal{M}^n$ tal que $I = \langle S \rangle$. A probar: $I \cap \mathbb{K}[x_1, \cdots, x_{n-1}] = \langle S \cap \mathbb{K}[x_1, \cdots, x_{n-1}] \rangle$. Sea $p \in \langle S \cap \mathbb{K}[x_1, \cdots, x_{n-1}] \rangle$, entonces $p = g_1 s_1 + \cdots + g_k s_k$, con $g_i \in \mathbb{K}[x_1, \cdots, x_{n-1}] \subseteq \mathbb{K}[x_1, \cdots, x_n]$ y $s_i \in S \cap \mathbb{K}[x_1, \cdots, x_{n-1}] \subseteq S$ para $i = 1, \cdots, k$; por lo que es inmediato que $p \in \langle S \rangle \subseteq \mathbb{K}[x_1, \cdots, x_n]$.

Por otro lado, sea $p = a_0 + a_1 \boldsymbol{x}^{\boldsymbol{\alpha}_1} + \cdots + a_r \boldsymbol{x}^{\boldsymbol{\alpha}_r} \in I \cap \mathbb{K}[x_1, \cdots, x_{n-1}] \Rightarrow p \in \mathbb{K}[x_1, \cdots, x_{n-1}]$ por lo que ningún monomio $\boldsymbol{x}^{\boldsymbol{\alpha}_i}$ de p contiene potencias de x_n. Además, $p \in I$ que es un ideal de monomios, por lo que cada monomio $\boldsymbol{x}^{\boldsymbol{\alpha}_i}$ de p pertenece a I. Entonces existen $s_i \in S$ y $m_i \in \mathcal{M}^n$ tales que $m_i s_i = \boldsymbol{x}^{\boldsymbol{\alpha}_i}$. Como en $\boldsymbol{x}^{\boldsymbol{\alpha}_i}$ no aparece ninguna potencia de x_n, tampoco debe hacerlo en $m_i s_i$, lo que significa que $m_i \in \mathcal{M}^{n-1}$ y $s_i \in S \cap \mathbb{K}[x_1, \cdots, x_{n-1}]$, para $i = 1, \cdots, r$. Entonces

$$p = a_0 g_0 s_0 + a_1 g_1 s_1 + \cdots + a_r g_r s_r \in \langle S \cap \mathbb{K}[x_1, \cdots, x_{n-1}] \rangle \subseteq \mathbb{K}[x_1, \cdots, x_{n-1}].$$

Con lo que se concluye la igualdad $I \cap \mathbb{K}[x_1, \cdots, x_{n-1}] = \langle S \cap \mathbb{K}[x_1, \cdots, x_{n-1}] \rangle$ y $I \cap \mathbb{K}[x_1, \cdots, x_{n-1}]$ es un ideal de monomios de $\mathbb{K}[x_1, \cdots, x_{n-1}]$ como se quería. ∎

Teorema 4.21 (Lema de Dickson)**.** Todo ideal de monomios en $\mathbb{K}[x_1, \cdots, x_n]$ es generado por un conjunto finito.

Demostración. Procediendo por inducción sobre n.

1. Si $n = 1$, $\mathbb{K}[x_1]$ es un anillo de ideales principales (teorema 3.13) entonces $I = \langle p \rangle$ para algún $p \in \mathbb{K}[x_1]$. Como I es un ideal de monomios, debe existir $S \subseteq \mathcal{M}^1$ tal que $I = \langle p \rangle = \langle S \rangle$, si $s_1 \in S \Rightarrow\ s \in \langle p \rangle$ entonces $p | s$ lo que implica que p debe ser un monomio. Por lo tanto, I es generado por un único monomio.

2. Si $n > 1$, supóngase que se cumple el teorema para $n - 1$. Para cada $j \in \mathbb{N}$ sea

   $$J_j = (I : \langle x_n^j \rangle) \cap \mathbb{K}[x_1, \cdots, x_{n-1}].$$

   Por los lemas 4.19 y 4.20, J_j es un ideal de monomios de $\mathbb{K}[x_1, \cdots, x_{n-1}]$, luego por la hipótesis de inducción es finitamente generado $J_j = \langle S_j \rangle$. Sea $p \in J_j\ \Rightarrow p \in \mathbb{K}[x_1, \cdots, x_{n-1}]$ y $p \in (I : \langle x_n^j \rangle)$ entonces $pc \in I\ \forall c \in \langle x_n^j \rangle$, particularmente $p \cdot (x_n^j x_n c) \in I$ para cualquier $c \in \mathbb{K}[x_1, \cdots, x_n] \Rightarrow p \cdot (x_n^j x_n c) = p x_n^{j+1} c = pc' \in I$ para todo $c' = x_n^{j+1} c \in \langle x_n^{j+1} \rangle$ entonces $p \in J_{j+1} \Rightarrow J_j \subseteq J_{j+1}$. Entonces se tiene la cadena de inclusiones $J_0 \subseteq J_1 \subseteq J_2 \subseteq \cdots$, luego por el lema 4.18, $J = \bigcup_j J_j$ es un ideal de monomios de $\mathbb{K}[x_1, \cdots, x_{n-1}]$ y debe ser finitamente generado $\Rightarrow J = \langle S \rangle$ con S un conjunto finito.

   Como $S \subseteq J$ y es finito, debe existir un r tal que $S \subseteq J_r$ entonces $J = \langle S \rangle \subseteq J_r \subseteq J \Rightarrow J = J_i$ para todo $i \geq r$ lo que también significa que $S = S_i$ para todo $i \geq r$. Considérese ahora $m \in I$ un monomio, entonces $m = m' x_n^k$ con $m' \in \mathcal{M}^{n-1}$ para algún k. Como $m' x_n^k \in I\ \Rightarrow\ \ m' \in (I : \langle x_n^k \rangle) \cap \mathbb{K}[x_1, \cdots, x_{n-1}] = \langle S_k \rangle$ entonces $m \in \langle x_n^k S_k \rangle$. En consecuencia, $S' = S_0 \cup x_n S_1 \cup x_n^2 S_2 \cup \cdots$ es un conjunto que genera a I. Como a partir de r todos los S_i's

son iguales y $\left\langle x_n^{i+1} S_r \right\rangle \subseteq \left\langle x_n^i S_r \right\rangle$. Por tanto, el conjunto $S'' = S_0 \cup x_n S_1 \cup x_n^2 S_2 \cup \cdots \cup x_n^r S_r$ es un conjunto finito de monomios que genera a I. ∎

Lema 4.22. Cualquier orden de monomios $\prec$ en $\mathcal{M}^n \in \mathbb{K}[x_1, \cdots, x_n]$ es un buen ordenamiento, *i.e.* cada conjunto no vacío $S \subseteq \mathcal{M}^n$ tiene un elemento mínimo (que es menor que todos los demás).

Demostración. Sea $S \subseteq \mathcal{M}^n$. Considere $I = \langle S \rangle$ ideal de monomios, que por el Lema de Dickson debe tener un generador finito $S' = \{m_1, m_2, \cdots, m_k\} \Rightarrow\ I = \langle S \rangle = \langle m_1, m_2, \cdots, m_k \rangle$. Sea m_0 el menor de los m_i's en S' con el orden $\prec$. Luego, cualquier monomio $m \in S \subseteq I = \langle m_1, \cdots, m_k \rangle$ debe ser divisible por algún $m_i \Rightarrow\ m_i | m\ \Rightarrow m \succeq m_i \succeq m_0$. Por lo tanto m_0 es el elemento mínimo de $\langle S \rangle$. Ahora bien, $m_0 \in S' \Rightarrow\ m_0 \in \langle S \rangle \Rightarrow$ existe $m \in S$ tal que $m | m_0 \Rightarrow m \preceq m_0$ donde la única opción es que $m = m_0 \Rightarrow m_0 \in S$. Por lo que $m_0 \in S$ es el elemento mínimo de S. ∎

Corolario 4.23. Una cadena estrictamente decreciente de monomios en cualquier orden de monomios $\prec$ es finita.

Demostración. Sea $m_1 \succ m_2 \succ \cdots$ una cadena decreciente de monomios en $\mathcal{M}^n$ y sea $S = \{m_i | i = 1, 2, \cdots\}$, como S tiene un elemento mínimo, la cadena debe ser finita. ∎

4.4. El Algoritmo de reducción

En el capítulo anterior, se estudió cómo el algoritmo de la división permite resolver el problema de pertenencia a ideales, concluyendo que un polinomio p pertenece a un ideal $I = \langle q \rangle \subseteq \mathbb{K}[x]$ si y sólo si, $\mathbf{rem}(p, q) = 0$. Más aún, se caracterizaron los anillos cocientes al utilizar como representantes de las clases laterales los residuos de dividir dentro de q. Esta metodología fue suficiente pues, como se demostró en el teorema 3.13, $\mathbb{K}[x]$ es un anillo de ideales principales y basta con considerar al generador de cada ideal para caracterizarlo.

Considérese en cambio $\mathbb{K}[x, y]$ y el ideal $I = \left\langle x, y^2 \right\rangle \subseteq \mathbb{K}[x, y]$. Como I es un ideal de monomios, si $y \in I$ entonces x o y^2 debería dividir a y, por lo que $y \notin I \Rightarrow I \neq \mathbb{K}[x, y] = \langle 1 \rangle$. Supóngase que existe un polinomio $p \in \mathbb{K}[x, y]$ tal que $I = \langle p \rangle \Rightarrow p | x$ y $p | y^2$, sin embargo es claro que el único divisor común de x y y^2 es 1, y como se acaba de ver $I \neq \langle 1 \rangle$, lo que es imposible, por lo que el ideal $\left\langle x, y^2 \right\rangle$ no se puede generar por un sólo elemento de $\mathbb{K}[x, y]$. Entonces $\mathbb{K}[x, y]$ y, en general, $\mathbb{K}[x_1, \cdots, x_n]$, no son de ideales principales, por lo que el algoritmo de reducción que se introduzca debe permitir escribir los polinomios como combinaciones lineales de múltiples *divisores* en lugar de sólo uno, como en el caso de $\mathbb{K}[x]$. El algoritmo de reducción que se introduce a continuación permite reescribir un polinomio p en la forma

$$p = c_1 q_1 + c_2 q_2 + \cdots + c_k q_k + r,$$

donde el *residuo* r es *irreducible* por los q_i's, según la definición que procede.

Definición 4.24. *Dados un orden de monomios $\prec$ en $\mathcal{M}^n \subseteq \mathbb{K}[x_1, \cdots, x_n]$ fijo y una secuencia de polinomios $q_1, q_2, \cdots, q_k \in \mathbb{K}[x_1, \cdots, x_n]$, se dice que un polinomio $r \in \mathbb{K}[x_1, \cdots, x_n]$ está* ***reducido respecto los*** *q_i**'s** si $r = 0$ o ninguno de los monomios $\mathbf{lm}(q_1), \mathbf{lm}(q_2), \cdots, \mathbf{lm}(q_k)$ divide a algún monomio de r.*

Teorema 4.25 (Algoritmo de reducción)**.** Sea $\prec$ un orden de monomios fijo en $\mathcal{M}^n$. Dados $p \in \mathbb{K}[x_1, \cdots, x_n]$ y la secuencia ordenada $q_1, q_2, \cdots, q_k \in \mathbb{K}[x_1, \cdots, x_n]$, entonces existen $r, c_1, c_2, \cdots, c_k \in \mathbb{K}[x_1, \cdots, x_n]$ tales que

$$p = c_1 q_1 + c_2 q_2 + \cdots + c_k q_k + r,$$

con r reducido respecto $q_1, \cdots, q_k$. Además si $c_i, q_i \neq 0$ entonces $\mathbf{lm}(p) \succeq \mathbf{lm}(c_i q_i)$.

Demostración. El algorimo 4.1 encuentra los polinomios c_i's y r buscados. ∎

Algoritmo 4.1 (Algoritmo de reducción)

Input: $p, q_1, \cdots, q_k \in \mathbb{K}[x_1, \cdots, x_n] \setminus \{0\}$.

Output: $c_1, \cdots, c_k, r \in \mathbb{K}[x_1, \cdots, x_n]$ tales que $p = c_1 q_1 + \cdots + c_k q_k + r$, r es reducido respecto $q_1, \cdots, q_k$ y máx$(\mathbf{lm}(c_1 q_1), \mathbf{lm}(c_2 q_2), \cdots, \mathbf{lm}(c_k q_k), \mathbf{lm}(r)) = \mathbf{lm}(p)$.

begin
  $c_1 \leftarrow 0, c_2 \leftarrow 0, \cdots, c_k \leftarrow 0, r \leftarrow 0, h \leftarrow p$;
  while $h \neq 0$ **do**
    if *existe i tal que* $\mathbf{lm}(q_i)$ *divide a* $\mathbf{lm}(h)$ **then**
      tómese el menor i tal que $\mathbf{lm}(q_i)$ divide a $\mathbf{lm}(h)$;
      $c_i \leftarrow c_i + \dfrac{\mathbf{lt}(h)}{\mathbf{lt}(q_i)}$;
      $h \leftarrow h - \dfrac{\mathbf{lt}(h)}{\mathbf{lt}(q_i)} q_i$;
    else
      $r \leftarrow r + \mathbf{lt}(h)$;
      $h \leftarrow h - \mathbf{lt}(h)$;
    end
  end
end

Demostración.

- **Finitud:** Nótese que $\mathbf{lt}\left(\dfrac{\mathbf{lt}(h)}{\mathbf{lt}(q_i)} q_i\right) = \dfrac{\mathbf{lt}(h)}{\mathbf{lt}(q_i)} \mathbf{lt}(q_i) = \mathbf{lt}(h)$ por lo que en cada iteración del ciclo **while**, sin importar si se pasa por el **if** o por el **else**, se elimina el término principal de h, lo que, si no hace que h sea cero, deja a h con un monomio principal menor al que tenía originalmente. Denotando por $h^{(t)}$ al polinomio h luego de t iteraciones del ciclo **while**, esto significa que $\mathbf{lm}\big(h^{(0)}\big) \succ \mathbf{lm}\big(h^{(1)}\big) \succ \mathbf{lm}\big(h^{(2)}\big) \succ \cdots$ que es una cadena decreciente de monomios. Por el corolario 4.23, esta cadena debe ser finita y de tamaño T. Así, luego de un

máximo de T pasos, el monomio principal de h no decrece, sino que h debe volverse cero y el algoritmo termina.

- **Correctitud:** Denótese con un exponente de la forma $\{\cdot\}^{(t)}$ a cada polinomio después de la t-ésima iteración del ciclo **while**, entendiendo por $t = 0$ a la inicialización. Nótese entonces que:

$$c_1^{(0)} q_1 + \cdots + c_k^{(0)} q_k + r^{(0)} + h^{(0)} = h^{(0)} = p \tag{4.1}$$

Supóngase que en el paso t esta expresión sigue siendo válida, luego en el paso $t+1$ se pueden dar dos casos:

I) Si existe i tal que $\mathbf{lm}(q_i) \,|\mathbf{lm}(h)$, se elige el menor i con dicha propiedad y entonces:

$$\begin{aligned} &c_1^{(t+1)} q_1 + \cdots + c_i^{(t+1)} q_i + \cdots + c_k^{(t+1)} q_k + r^{(t+1)} + h^{(t+1)} = \\ &\quad = c_1^{(t)} q_1 + \cdots + \left(c_i^{(t)} + \frac{\mathbf{lt}\big(h^{(t)}\big)}{\mathbf{lt}(q_i)} \right) q_i + \cdots + c_k^{(t)} q_k + r^{(t)} + \left(h^{(t)} - \frac{\mathbf{lt}\big(h^{(t)}\big)}{\mathbf{lt}(q_i)} q_i \right) = \\ &\quad = c_1^{(t)} q_1 + \cdots + c_i^{(t)} q_i + \cdots + c_k^{(t)} q_k + r^{(t)} + h^{(t)} = p \end{aligned}$$

II) Si no existe i tal que $\mathbf{lm}(q_i) \,|\mathbf{lm}(h)$, el algoritmo pasa por el **else** y entonces:

$$\begin{aligned} &c_1^{(t+1)} q_1 + \cdots + c_k^{(t+1)} q_k + r^{(t+1)} + h^{(t+1)} = \\ &= c_1^{(t)} q_1 + \cdots + c_k^{(t)} q_k + \left(r^{(t)} + \mathbf{lt}\Big(h^{(t)}\Big) \right) + \left(h^{(t)} - \mathbf{lt}\Big(h^{(t)}\Big) \right) = \\ &= c_1^{(t)} q_1 + \cdots + c_i^{(t)} q_i + \cdots + c_k^{(t)} q_k + r^{(t)} + h^{(t)} = p \end{aligned}$$

En ambos casos, la expresión 4.1 sigue siendo válida, por lo que es válida para cualquier tiempo t. Luego, al finalizar el algoritmo, $h = 0$ y la expresión se reduce a

$$c_1 q_1 + \cdots + c_k q_k + r = p.$$

Obsérvese ahora la construcción de r. Inicialmente $r = 0$, en cada iteración del ciclo **while**, si para ningún i existe q_i tal que $\mathbf{lm}(q_i) \mid \mathbf{lm}(h)$ se pasa al **else** en donde se le agrega el término $\mathbf{lt}(h)$ a r. Esto quiere decir que r está formado por la combinación lineal de monomios que no son divisibles dentro de ningún $\mathbf{lt}(q_i)$, o $r = 0$ si nunca se entra al **else**. En ambos casos, r es reducido respecto a los q_i's.

Finalmente, los c_i's se inicializan como cero, y se les agrega un término de la forma $\dfrac{\mathbf{lt}(h)}{\mathbf{lt}(q_i)}$ cada vez que i es el menor índice tal que $\mathbf{lt}(q_i) \,|\mathbf{lt}(h)$, por lo que

$$\mathbf{lm}(c_i q_i) = \mathbf{lm}(c_i)\,\mathbf{lm}(q_i) \prec \mathbf{lm}\left(\frac{\mathbf{lt}(h)}{\mathbf{lt}(q_i)} \right) \mathbf{lm}(q_i) = \frac{\mathbf{lm}(h)}{\mathbf{lm}(q_i)} \mathbf{lm}(q_i) = \mathbf{lm}(h) \preceq \mathbf{lm}(p)$$

Como $p = c_1 q_1 + \cdots + c_k q_k + r$, por lo menos uno de los sumandos debe tener el mismo monomio principal que p y entonces $\text{máx}(\mathbf{lm}(c_1 q_1), \mathbf{lm}(c_2 q_2), \cdots, \mathbf{lm}(c_k q_k)) = \mathbf{lm}(p)$. ∎

***Ejemplo* 4.4**

Como ejemplo considérense $p = x^2y + xy^2 + y^2$, $q_1 = xy - 1$ y $q_2 = y^2 - 1$ con el orden DegLex. La Tabla 4.1 muestra los pasos del algoritmo 4.1 aplicado a este caso, devolviendo como resultado $c_1 = x + y$, $c_2 = 1$ y $r = x + y + 1$.

Tabla 4.1: Algoritmo de reducción 4.1 aplicado a $p = x^2y + xy^2 + y^2$, $q_1 = xy - 1$ y $q_2 = y^2 - 1$

Paso	$\boldsymbol{h}$	$\boldsymbol{c_1}$	$\boldsymbol{c_2}$	$\boldsymbol{r}$	**Comentario**
0	$p = x^2y + xy^2 + y^2$	0	0	0	
1	$h \leftarrow h - xq_1 = xy^2 + y^2 + x$	x	0	0	$\mathbf{lt}(q_1) = xy \vert \mathbf{lt}(h) = x^2y$
2	$h \leftarrow h - yq_1 = y^2 + x + y$	$x + y$	0	0	$\mathbf{lt}(q_1) = xy \vert \mathbf{lt}(h) = xy^2$
3	$h \leftarrow h - 1 \cdot q_2 = x + y + 1$	$x + y$	1	0	$\mathbf{lt}(q_2) = y^2 \vert \mathbf{lt}(h) = y^2$
4	$h \leftarrow h - \mathbf{lt}(h) = y + 1$	$x + y$	1	x	$\mathbf{lt}(q_i) \nmid \mathbf{lt}(h) = x$ para $i = 1, 2$
5	$h \leftarrow h - \mathbf{lt}(h) = 1$	$x + y$	1	$x + y$	$\mathbf{lt}(q_i) \nmid \mathbf{lt}(h) = x$ para $i = 1, 2$
6	$h \leftarrow h - \mathbf{lt}(h) = 0$	$x + y$	1	$x + y + 1$	
Resultado	0	$x + y$	1	$x + y + 1$	

Considérese ahora el mismo problema pero intercambiando q_1 con q_2, es decir, ahora $q_1' = y^2 - 1$ y $q_2' = xy - 1$ con el mismo orden DegLex. Los pasos del algoritmo de reducción se muestran en la Tabla 4.2, obteniéndose como resultado $c_1' = x + 1$, $c_2' = x$ y $r' = 2x + 1$. Aquí se hace evidente uno de los inconvenientes que presenta el algoritmo de reducción al dividir dentro de secuencias de polinomios cualesquiera: el residuo no es único (ni la representación tampoco).

Tabla 4.2: Algoritmo de reducción 4.1 aplicado a $p = x^2y + xy^2 + y^2$, $q_1 = y^2 - 1$ y $q_2 = xy - 1$

Paso	$\boldsymbol{h}$	$\boldsymbol{c_1'}$	$\boldsymbol{c_2'}$	$\boldsymbol{r'}$	**Comentario**
0	$p = x^2y + xy^2 + y^2$	0	0	0	
1	$h \leftarrow h - xq_2' = xy^2 + y^2 + x$	0	x	0	$\mathbf{lt}(q_2') = xy \vert \mathbf{lt}(h) = x^2y$
2	$h \leftarrow h - xq_1' = y^2 + 2x$	x	x	0	$\mathbf{lt}(q_1') = y^2 \vert \mathbf{lt}(h) = xy^2$
3	$h \leftarrow h - q_1' = 2x + 1$	$x + 1$	x	0	$\mathbf{lt}(q_1') = y^2 \vert \mathbf{lt}(h) = y^2$
4	$h \leftarrow h - \mathbf{lt}(h) = 1$	$x + 1$	x	$2x$	$\mathbf{lt}(q_i') \nmid \mathbf{lt}(h) = 2x$ para $i = 1, 2$
5	$h \leftarrow h - \mathbf{lt}(h) = 0$	$x + 1$	x	$2x + 1$	$\mathbf{lt}(q_i') \nmid \mathbf{lt}(h) = 2x$ para $i = 1, 2$
Resultado	0	$x + 1$	x	$2x + 1$	

***Ejemplo* 4.5**

El algoritmo 4.1 está inspirado en el algoritmo de la división 3.1, por lo que se espera que al aplicarlo en $\mathbb{K}[x]$ y con un único divisor q, coincida con el algoritmo de la división. Note primero

que dado $p \in \mathbb{K}[x]$, el algoritmo de reducción devuelve polinomios $c, r \in \mathbb{K}[x]$ tales que $p = cq + r$ donde r es reducido respecto a q. Que r sea reducido respecto a q significa que $r = 0$ o que ningún monomio de r es divisible dentro de $\mathbf{lm}(q)$, pero obsérvese que en $\mathbb{K}[x]$ para que un monomio sea divisible dentro de otro, es suficiente y necesario que su grado sea mayor que el del divisor, *i.e.* $\mathbf{lm}(q) \,|m$ con m monomio de r si y sólo si $\mathbf{deg}(q) \leq \mathbf{deg}(m) \leq \mathbf{deg}(r)$, por lo que debe suceder que $\mathbf{deg}(r) < \mathbf{deg}(q)$, obteniéndose el mismo resultado que en el algoritmo de la división 3.1 (pues c y r son únicos en el caso de una variable).

Igual como se hizo para $\mathbb{K}[x]$, lo más interesante del algoritmo de la división es el residuo r, por lo que se define una notación particular.

Definición 4.26. *Dado un orden fijo de monomios en $\mathcal{M}^n$ y una secuencia de polinomios $q_1, \cdots, q_k$ en $\mathbb{K}[x_1, \cdots, x_n] \setminus \{0\}$, para $p \in \mathbb{K}[x_1, \cdots, x_n]$ se define el* ***residuo de p módulo la secuencia*** *$q_1, \cdots, q_k$, $\mathbf{rem}(p; q_1, \cdots, q_k)$ como el residuo r de aplicar el algoritmo de reducción a p respecto a $q_1, \cdots, q_k$ (en ese orden).*

Observación. aunque la definición anterior requiere que los polinomios q_i's se tomen en algún orden determinado, por simplicidad de notación se utilizará también la notación $r = \mathbf{rem}(p; G)$ con G un conjunto para denotar que r es un polinomio reducido respecto a G utilizando el algoritmo 4.1 para algún orden de los elementos de G determinado (el orden en el que se enumeran sus elementos, por ejemplo). Cuando se desee explicitar en qué orden se utilizan los polinomios en el algoritmo 4.1 se escribirá la secuencia de forma explícita en $\mathbf{rem}(p; q_1, \cdots, q_k)$.

Proposición 4.27. Dados $p, q_1, \cdots, q_k \in \mathbb{K}[x_1, \cdots, x_k]$, $p - \mathbf{rem}(p; q_1, \cdots, q_k) \in \langle q_1, \cdots, q_k \rangle$. En particular, si $\mathbf{rem}(p; q_1, \cdots, q_k) = 0 \Rightarrow \ p \in \langle q_1, \cdots, q_k \rangle$.

Demostración. Del algoritmo de reducción, $p = c_1 q_1 + \cdots + c_k q_k + \mathbf{rem}(p; q_1, \cdots, q_k)$ entonces $p - \mathbf{rem}(p; q_1, \cdots, q_k) = c_1 q_1 + \cdots + c_k q_k \in \langle q_1, \cdots, q_k \rangle$. Evidentemente si el residuo es cero, $p - 0 = p \in \langle q_1, \cdots, q_k \rangle$. ∎

El ejemplo 4.4 muestra como el residuo módulo un conjunto de polinomios no es único, lo que particularmente imposibilita que el recíproco de la propiedad anterior sea verdadero, al menos para secuencias de polinomios cualesquiera. En el siguiente capítulo se estudia otro conjunto generador del ideal $\langle q_1, \cdots, q_k \rangle$ de tal forma que el residuo de dividir dentro de dicho conjunto sí sea único. La siguiente definición permite referirse a estos residuos como una clase, simplificando la notación y sin exigir un orden determinado para los elementos *divisores*. A diferencia del residuo $\mathbf{rem}(\cdot)$ no se contará con un algoritmo para determinar todos los posibles residuos, sin embargo será de interés teórico para los próximos capítulos.

Definición 4.28. *Dado un orden fijo de monomios en $\mathcal{M}^n$ y un conjunto de polinomios $G = \{q_1, \cdots, q_k\} \subseteq \mathbb{K}[x_1, \cdots, x_n] \setminus \{0\}$, para $p, r \in \mathbb{K}[x_1, \cdots, x_n]$ se dice que p* ***se reduce a r re-***

specto a (o módulo) *G, denotado por $p \xrightarrow{G} r$ si y sólo si r está reducido respecto a G (según la definición 4.24) y existen polinomios $c_i \in \mathbb{K}[x_1, \cdots, x_n]$ tales que $p = c_1 q_1 + \cdots c_k q_k + r$ y $\mathbf{lm}(p) = \text{máx}_i(\mathbf{lm}(c_i q_i), r)$.*

Observación. claramente $p \xrightarrow{G} \mathbf{rem}(p; G)$ con cualquier orden de los elementos de G que se haga la reducción, según garantiza el algoritmo 4.1.

Proposición 4.29. Sean $G = \{q_1, \cdots, q_k\} \subseteq \mathbb{K}[x_1, \cdots, x_n] \setminus \{0\}$ y $p \in \mathbb{K}[x_1, \cdots, x_n]$ entonces se cumplen las propiedades:

1. Si $p \in G$ entonces $p \xrightarrow{G} 0$
2. Si $G \subseteq G'$ finitos y $p \xrightarrow{G} 0$ entonces $p \xrightarrow{G'} 0$
3. Si $p \xrightarrow{G} r$ entonces $p - r \in \langle G \rangle$

Demostración.

1. Si $p \in G$ sin pérdida de generalidad sea $q_1 = p$ entonces $p = 1 \cdot p + 0 \cdot q_1 + \cdots + 0 \cdot q_k + 0$, con 0 reducido respecto a G y $\mathbf{lm}(p) = \text{máx}(\mathbf{lm}(c_i q_i), 0) = \mathbf{lm}(p)$ por lo que $p \xrightarrow{G} 0$.
2. Como $G \subseteq G'$ sea $G' = \{q_1, \cdots, q_k, g_1, \cdots, g_s\}$ con $g_i \in G' \setminus G$. Como $p \xrightarrow{G} 0$, $p = c_1 q_1 + \cdots c_k q_k$ con $\mathbf{lm}(p) = \text{máx}_i(\mathbf{lm}(c_i q_i))$ que es lo mismo que $p = c_1 q_1 + \cdots c_k q_k + 0 \cdot g_1 + \cdots + 0 \cdot g_s$ con $\mathbf{lm}(p) = \text{máx}_i(\mathbf{lm}(c_i q_i), 0)$ por lo que $p \xrightarrow{G'} 0$.
3. Si $p \xrightarrow{G} r \Rightarrow p = c_1 q_1 + \cdots c_k q_k + r$ entonces $p - r = c_1 q_1 + \cdots c_k q_k \in \langle G \rangle$.

■

5 BASES DE GRÖBNER

En este capítulo se define finalmente el concepto fundamental de este trabajo: las *bases de Gröbner*. Como se verá en breve, el nombre *base* hace alusión a que es un conjunto que permite generar el mismo ideal de polinomios de interés, pero que posee propiedades que lo hacen más sencillo de utilizar que cualquier conjunto de generadores. El concepto fue acuñado por Bruno Buchberger en 1965 en su tesis doctoral titulada(traducción al inglés: Buchberger, 2006):

≪ *Un algoritmo para encontrar los elementos de la base del anillo de residuos de un ideal de polinomios cero-dimensional* ≫

en la cual no sólo define las bases de Gröbner sino que da a conocer el algoritmo que ahora lleva su nombre para construirlas, probando la correctitud y finitud del mismo. Buchberger utiliza en su tesis las bases de Gröbner para dar la forma normal de los elementos del anillo de residuos de un ideal cero-dimensional, *i.e.* un ideal de polinomios tal que existe un conjunto finito de soluciones $\mathcal{S}$ tal que para todo elemento p de dicho ideal, $p(a_1, \cdots, a_n) = 0$ para algún$(a_1, \cdots, a_n) \in \mathcal{S}$, aunque después se probó que el método funciona igualmente para anillos de residuos de cualquier dimensión.

Buchberger nombró las bases de Gröbner en honor a su asesor de tesis Wolfgang Gröbner, quién había empezado a desarrollar estas ideas desde unos 20 años antes de la publicación de Buchberger, y quién le encargó como tema de su disertación doctoral el formalizar y delimitar mejor el método. Aunque la definición dada por Buchberger en su trabajo original difiere ligeramente de la presentada aquí (la definición original de Buchberger es exclusiva para el tipo de estructuras con las que trabajó en su tesis (vea Buchberger, 2006)), se prefiere la elegida pues parece ser la más adecuada para los propósitos de este trabajo, en cualquier caso, existen múltiples caracterizaciones de las bases de Gröbner, algunas de las cuales se mencionan en este capítulo.

Como se muestra a continuación, el interés en las bases de Gröbner radica en que sintetizan muchas de las propiedades de los ideales en conjuntos finitos que los generan, haciendo posible así el estudio de estas propiedades. Dado que el ideal trivial $\langle 0 \rangle = \{0\}$ es en sí mismo su propia base de Gröbner, resulta redundante probar los próximos resultados sobre el mismo, por lo que se excluye de la discusión que prosigue. En lo que resta del capítulo, se asumirá que se trabaja con un orden $\prec$ fijo cualquiera sobre $\mathcal{M}^n$.

Definición 5.1. *Dado un subconjunto $S \subseteq \mathbb{K}[x_1, \cdots, x_n]$ con $S \neq \{0\}$, se define el* ***ideal de términos principales****, $\mathbf{lt}(S)$ como el ideal*[1]

$$\mathbf{lt}(S) = \langle \mathbf{lt}(s) \mid s \in S \rangle .$$

Definición 5.2. *Dado un ideal $I \subseteq \mathbb{K}[x_1, \cdots, x_n]$ con $I \neq \{0\}$, un subconjunto $G = \{g_1, \cdots, g_s\} \subseteq I$ es una* ***base de Gröbner del ideal*** *I si, y sólo si, $\mathbf{lt}(I) = \langle \mathbf{lt}(g_1), \mathbf{lt}(g_2), \cdots, \mathbf{lt}(g_s) \rangle$.*

Lema 5.3. Si $G = \{g_1, \cdots, g_s\}$ es una base de Gröbner del ideal I entonces $I = \langle g_1, \cdots, g_s \rangle$.

Demostración. Dado que $G \subseteq I$, evidentemente $\langle G \rangle \subseteq I$. Sea $p \in I$ entonces $\mathbf{lt}(p) \in \mathbf{lt}(I) = \langle \mathbf{lt}(g_1), \cdots, \mathbf{lt}(g_s) \rangle \Rightarrow$ existe i tal que $\mathbf{lt}(p) = c\mathbf{lt}(g_i)$ para algún $\dfrac{c}{\mathbf{lc}(c)} \in \mathcal{M}^n$. Claramente $p - cg_i \in I$ y además[2] $\mathbf{lt}(p - cg_i) \prec \mathbf{lt}(p)$. Aplicando el mismo razonamiento recursivamente, se termina el proceso cuando $p - cg_i - c_2 g_{i2} - \cdots = 0$, por lo que p es un combinación lineal de los g_i's, $\Rightarrow$ $p \in \langle g_1, \cdots, g_s \rangle$ y entonces $I = \langle G \rangle$. ∎

5.1. Teorema de la base de Hilbert

Del lema 5.3 y del lema de Dickson, se obtiene como corolario que todos los ideales de un anillo de polinomios son finitamente generados. Este resultado se conoce como el *teorema de la base de Hilbert*, y fue demostrado por el matemático alemán en 1891.

Corolario 5.4 (Teorema de la base de Hilbert)**.** Todo ideal de un anillo de polinomios $\mathbb{K}[x_1, \cdots, x_n]$ es finitamente generado.

Demostración. Sea $I \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un ideal, $I \neq \{0\}$. Como $\mathbf{lt}(I)$ es un ideal de monomios, es finitamente generado según el lema de Dickson (teorema 4.21). Sea $\mathbf{lt}(I) = \langle m_1, \cdots, m_s \rangle$. Para cualquiera de los m_i's generadores, como $m_i \in \mathbf{lt}(I) = \langle \mathbf{lt}(p) \,|p \in I \rangle$ debe existir p_i tal que $\mathbf{lt}(p_i) \,| m_i \Rightarrow$ existe $c_i \in \mathcal{M}^n$ tal que $m_i = c_i \mathbf{lt}(p_i) = \mathbf{lt}(c_i p_i) = \mathbf{lt}(q_i)$ con $q_i = c_i p_i \in I$. Por lo tanto $I = \langle \mathbf{lt}(q_1), \cdots, \mathbf{lt}(q_s) \rangle$ y $\{q_1, \cdots, q_s\}$ es una base de Gröbner (finita) de I. ∎

Obsérvese que no sólo se acaba de demostrar que todo ideal de $\mathbb{K}[x_1, \cdots, x_n]$ es finitamente generado, sino que es generado por una base de Gröbner. Para futura referencia se enuncia este resultado en el siguiente corolario.

Corolario 5.5. Todo ideal $I \subseteq \mathbb{K}[x_1, \cdots, x_n]$, $I \neq \{0\}$, tiene base de Gröbner.

Definición 5.6. *Dado un anillo R, se dice que R es un* ***anillo noetheriano*** *si, y sólo si, todos sus ideales son finitamente generados.*

---

[1] Observe que $\langle \mathbf{lt}(s) \,| s \in S \rangle = \langle \mathbf{lm}(s) \,| s \in S \rangle$ por lo que en ocasiones se usarán indistintamente, aunque es común encontrar ambas presentaciones, en este caso se prefiere la convención tomada por Fröberg (1997), como principal referencia de este trabajo.

[2] Se entiende que $\mathbf{lt}(p) \prec \mathbf{lt}(q)$ si, y sólo si, $\mathbf{lm}(p) \prec \mathbf{lm}(q)$.

Lema 5.7. Sea R un anillo. Las siguientes condiciones son equivalentes:

1. R es noetheriano.
2. Toda cadena estrictamente ascendente de ideales $I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \cdots$ es estacionaria (A esta propiedad se le conoce como *condición de cadenas ascendentes*).
3. Todo conjunto no vacío de ideales $M = \{I_\alpha\}$ tiene un elemento maximal, *i.e.* existe un ideal $I_{\alpha_0} \in M$ tal que ningún otro elemento de M contiene estrictamente a I_{α_0}.

Demostración.

- (1⇒ 2): Sea $I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \cdots$ una cadena estrictamente ascendente de ideales entonces $I = \bigcup_i I_i$ es un ideal de R, por lo que debe ser finitamente generado. Entonces, la cadena no puede ser infinita, pues debe parar una vez I_n contenga todos los elementos del conjunto generador finito de I.
- (2⇒ 3): Supóngase que existe un conjunto M de ideales que no contiene un ideal maximal, entonces esto significa que se puede construir una cadena estrictamente ascendente infinita, contradiciendo 2. ($\rightarrow\leftarrow$).
- (3⇒ 1): Supóngase que existe un ideal $I \subseteq R$ tal que no es finitamente generado, entonces se puede construir una cadena estrictamente ascendente de ideales $I_1 \subsetneq I_2 \subsetneq \cdots$ al hacer $I_1 = \langle s_1 \rangle$ con $s_1 \in I$, $I_2 = \langle s_1, s_2 \rangle$ con $s_2 \in I \setminus I_1$, $I_3 = \langle s_1, s_2, s_3 \rangle$ con $s_3 \in I \setminus I_2$, ... Sin embargo, el conjunto $M = \{I_1, I_2, \cdots\}$ tiene, por hipótesis, un elemento maximal que no está estrictamente contenido en ninguno de los I_i's ($\rightarrow\leftarrow$). Por lo que todos los ideales de R deben ser finitamente generados. ∎

Observación. el teorema de la base de Hilbert implica que $\mathbb{K}[x_1, \cdots, x_n]$ es un anillo noetheriano, por lo que cumple con el teorema anterior. Particularmente, será de utilidad notar que cumple con la condición de cadenas ascendentes.

5.2. Algunas propiedades

Proposición 5.8. Sean $I \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un ideal y $G = \{g_1, g_2, \cdots, g_s\} \subseteq I$ una base de Gröbner. Entonces $\mathbf{rem}(p; G) = \mathbf{rem}(q; G)$ si, y sólo si, $p - q \in I$. Particularmente, $\mathbf{rem}(p; G) = 0$ si, y sólo si, $p \in I$.

Demostración. ($\Rightarrow$) Sea $r = \mathbf{rem}(p; G) = \mathbf{rem}(q; G)$ entonces $p = c_1 g_1 + \cdots + c_s g_x + r$ y $q = c'_1 g_1 + \cdots + c'_s g_s + r \ \Rightarrow p - q = (c_1 - c'_1) g_1 + \cdots + (c_s - c'_s) g_s \in \langle g_1, \cdots, g_s \rangle = I$.

($\Leftarrow$) Supóngase que $p - q \in I$. Claramente $p - \mathbf{rem}(p; G), q - \mathbf{rem}(q; G) \in I$ luego $\mathbf{rem}(p; G) - \mathbf{rem}(q; G) = (q - \mathbf{rem}(q; G)) - (p - \mathbf{rem}(p; G)) + (p - q) \in I$. Si $\mathbf{rem}(p; G) \neq \mathbf{rem}(q; G)$ entonces $m = \mathbf{lt}(\mathbf{rem}(p; G) - \mathbf{rem}(q; G)) \in \mathbf{lt}(I) = \mathbf{lt}(G)$ lo que significa que algún $\mathbf{lt}(g_i)$ debe dividir a

m, sin embargo $\mathbf{rem}(p;G)$ y $\mathbf{rem}(q;G)$ son reducidos respecto a G, por lo que ninguno de sus términos puede ser divisible dentro de ningún $\mathbf{lt}(g_i)$. De ahí que su diferencia contiene términos que ya aparecían en $\mathbf{rem}(p;G)$ o $\mathbf{rem}(q;G)$ $(\rightarrow\leftarrow)$. Por lo tanto, $\mathbf{rem}(p;G) = \mathbf{rem}(q;G)$. ∎

El siguiente teorema resume algunas de las caracterizaciones de las bases de Gröbner.

Teorema 5.9. Sea $I \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un ideal, $I \neq \{0\}$. Para un conjunto de polinomios distintos de cero $G = \{g_1, \cdots, g_s\} \subseteq I$ son equivalentes.

1. G es una base de Gröbner de I.
2. $p \in I$ si, y sólo si, $\mathbf{rem}(p;G) = 0$.
3. $p \in I$ si, y sólo si, $p = \sum_{i=1}^{s} c_i g_i$ con $\mathbf{lm}(p) = \text{máx}_i(\mathbf{lm}(c_i)\,\mathbf{lm}(g_i))$.

Demostración.

- (1⇒ 2):

  (⇒) Por reducción al absurdo, supóngase que $p \in I$ y $\mathbf{rem}(p;G) \neq 0$. Como $p - \mathbf{rem}(p;G) \in I$ entonces $\mathbf{rem}(p;G) \in I \;\Rightarrow\; \mathbf{lt}(\mathbf{rem}(p;G)) \in \mathbf{lt}(I) = \mathbf{lt}(G)$ entonces alguno de los $\mathbf{lt}(g_i)$'s divide a $\mathbf{lt}(\mathbf{rem}(p;G))$, pero este está reducido respecto a G $(\rightarrow\leftarrow)$, un absurdo. Por lo tanto, $\mathbf{rem}(p;G) = 0$.

  (⇐) Si $\mathbf{rem}(p;G) = 0$ entonces $p = p - 0 = p - \mathbf{rem}(p;G) \in I$.

- (2⇒3): Si $p \in I$, por la propiedad 2 $\mathbf{rem}(p;G) = 0$. Entonces el algoritmo de reducción (algoritmo 4.1) asegura la propiedad. Si $p = \sum_i c_i g_i$ evidentemente $p \in I$.

- (3⇒2): Como $G \subseteq I$ claramente $\mathbf{lt}(G) \subseteq \mathbf{lt}(I)$. Sea $m \in \mathbf{lt}(I)$ con $am = \mathbf{lt}(p)$, $a \in \mathbb{K}$ (*i.e.* $m = \mathbf{lm}(p)$) para algún $p \in I$. Entonces $p = \sum_i^s c_i g_i$, con $m = \mathbf{lm}(p) = \text{máx}(\mathbf{lm}(c_i g_i)) \Rightarrow$ existe i tal que $\mathbf{lt}(g_i) \,|m \;\Rightarrow\; m \in \mathbf{lt}(G)$ por lo que $\mathbf{lt}(I) = \mathbf{lt}(G)$ y G es una base de Gröbner de I. ∎

El siguiente resultado muestra que las bases de Gröbner en efecto cumplen con la propiedad que se buscaba: que generan un residuo único (para un orden de monomios fijo $\prec$ en $\mathbb{K}[x_1, \cdots, x_n]$).

Teorema 5.10. Sean $I \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un ideal no cero y $G = \{g_1, \cdots, g_s\} \subseteq I$ una base de Gröbner de I. Para $p \in \mathbb{K}[x_1, \cdots, x_n]$ sean $r, c_1, \cdots, c_s \in \mathbb{K}[x_1, \cdots, x_n]$ tales que $p = \sum_{i=1}^{s} c_i g_i + r$ con r reducido respecto los g_i's (calculados a partir de cualquier método, no necesariamente mediante el algoritmo de reducción), entonces $r = \mathbf{rem}(p;G)$.

Demostración. Por la propiedad 4.27 $p - \mathbf{rem}(p;G) \in I$, y $p - r = \sum_{i=1}^{s} c_i g_i \in I$ entonces $\mathbf{rem}(p;G) - r = (p-r) - (p - \mathbf{rem}(p;G)) \in I$. Si esta diferencia no fuera cero $\Rightarrow \mathbf{lt}(\mathbf{rem}(p;G) - r) \in \mathbf{lt}(I) = \mathbf{lt}(G)$. Es claro que ningún $g_i \in G$ divide a ningún término de $\mathbf{rem}(p;G)$ ni de r (pues

están reducidos respecto los g_i's), por lo que tampoco dividen a $\mathbf{lt}(\mathbf{rem}(p;G)-r)$ y por ende este no puede ser elemento del ideal de monomios $\mathbf{lt}(G)$ $(\rightarrow\leftarrow)$. Por lo tanto, $\mathbf{rem}(p;G)=r$. ∎

Particularmente, el residuo r se puede calcular con el algoritmo de reducción utilizando una base de Gröbner, sin importar el orden de los elementos en esta base.

Corolario 5.11. Sean $I \subseteq \mathbb{K}[x_1,\cdots,x_n]$ un ideal no cero y $G=\{g_1,\cdots,g_s\} \subseteq I$ una base de Gröbner de I. Si G' es cualquier permutación de la secuencia $g_1,\cdots,g_s$ y cualquier $p \in \mathbb{K}[x_1,\cdots,x_n]$, $\mathbf{rem}(p;G)=\mathbf{rem}(p;G')$. Más aún, G' también es una base de Gröbner de I.

Demostración. Evidentemente $\mathbf{lt}(G)=\mathbf{lt}(G')=\mathbf{lt}(I)$, pues el ideal generado no depende del orden de sus generadores, por lo que G' es también una base de Gröbner de I. Haciendo $r=\mathbf{rem}(p;G')$ en el teorema anterior, se obtiene $\mathbf{rem}(p;G)=\mathbf{rem}(p;G')$. ∎

Con la definición de bases de Gröbner utilizada, se encontró que los residuos obtenidos por el algoritmo de reducción son únicos si se divide dentro de una base de Gröbner (teorema 5.10). En realidad también se cumple su recíproco: si para todo $p \in \mathbb{K}[x_1,\cdots,x_n]$ el residuo de p respecto a un conjunto G es único sin importar el orden de los elementos de G, entonces G debe ser una base de Gröbner[3], para la demostración de esta caracterización refiérase a Adams y Loustaunau (1994, p. 34).

Hasta ahora se ha visto que las bases de Gröbner en efecto resuelven el problema de pertenencia de ideales, y que como devuelven residuos únicos pueden ser utilizadas para encontrar una forma normal (única) de los elementos de los anillos cocientes. El siguiente teorema explora más a fondo la estructura de los anillos cocientes utilizando estas formas normales, para más información veánse Fröberg (1997) y Adams y Loustaunau (1994).

Teorema 5.12. Sea $I \subseteq \mathbb{K}[x_1,\cdots,x_n]$ un ideal con $I \neq \{0\}$. El anillo cociente $\mathbb{K}[x_1,\cdots,x_n]/I$ es un espacio vectorial sobre el campo $\mathbb{K}$ y el conjunto $\mathcal{B}=\{m+I \mid m \in \mathcal{M}^n \setminus \mathbf{lt}(I)\}$ es una base del mismo.

Demostración. Sea $R=\mathbb{K}[x_1,\cdots,x_n]/I$. La prueba de que R sobre el campo $\mathbb{K}$ es un espacio vectorial es estándar por lo que no se presénta aquí (veáse Herstein, 1975). Se muestra únicamente la segunda afirmación. Sea $\mathcal{B}=\{m+I \mid m \in \mathcal{M}^n \setminus \mathbf{lt}(I)\}$ y sea $G=\{g_1,\cdots,g_s\}$ una base de Gröbner para I. Es evidente también que para cualquier elemento $p+I \in R$, $p+I=\mathbf{rem}(p;G)+I$ pues $p-\mathbf{rem}(p;G) \in I$. Escribiendo $\mathbf{rem}(p;G)$ como una combinación lineal de monomios $\mathbf{rem}(p;G)=\sum_i \alpha_i m_i$ con $\alpha_i \in \mathbb{K}$ y $m_i \in \mathcal{M}^n$, como está reducido respecto a G, ninguno de los monomios m_i es divisible por los $\mathbf{lt}(g_i)$'s, lo que significa que $m_i \notin \mathbf{lt}(G)$ $\forall i$ entonces $m_i+I \in \mathcal{B}$ $\forall i$, por lo que cualquier elemento $p+I \in R$ es una combinación lineal de elementos de $\mathcal{B}$ y este es

[3]Se entiende que G es una base de Gröbner sin especificar de qué ideal, como el hecho que G es una base de Gröbner del ideal $\langle G\rangle$.

entonces un conjunto generador de R. Sólo resta probar la independencia lineal de $\mathcal{B}$. Supóngase que existen coeficientes $\alpha_i \in \mathbb{K} \setminus \{0\}$ tales que

$$0 + I = \sum_{m_i \in \mathcal{B}} \alpha_i (m_i + I) = \left(\sum_{m_i \in \mathcal{B}} \alpha_i m_i \right) + I$$

Como $P = \sum_{m_i \in \mathcal{B}} \alpha_i m_i$ es una combinación de monomios distintos y los $\alpha_i \neq 0$ entonces es un polinomio necesariamente distinto de cero, que además es reducido respecto a G. Pero si $0 + I = P + I \Rightarrow P = P - 0 \in I$ y $\mathbf{rem}(P; G) = P \neq 0$ ($\rightarrow\leftarrow$). Por lo tanto $\mathcal{B}$ es linealmente independiente y es una base del espacio vectorial R. ∎

Ahora es posible darles solución a los dos problemas planteados al comienzo del Capítulo 4. Para su solución se asume que ya se conoce una base de Gröbner del ideal. El cómo encontrar una base de Gröbner para un ideal cualquiera se trabajará hasta el próximo capítulo con el algoritmo de Buchberger, por lo que el algoritmo para resolver dichos problemas se proporciona hasta el final del Capítulo 6.

5.3. Bases de Gröbner reducidas

Dado un ideal I con base de Gröbner G, esta base no es única. Por ejemplo, se puede agregar cualquier otro elemento de I a G y el nuevo conjunto G' seguirá siendo una base de Gröbner para I (pues $\mathbf{lt}(G') = \mathbf{lt}(G) = \mathbf{lt}(I)$). Para solventar este problema se definen las *bases de Gröbner reducidas* asegurando así la unicidad de estas para cada ideal (con un orden de monomios fijo). A continuación se da la definición formal de estas, así como la demostración de que son únicas y el algoritmo para convertir cualquier base de Gröbner de un ideal I en la base reducida.

Definición 5.13. *Sea G una base de Gröbner del ideal $I \subseteq \mathbb{K}[x_1, \cdots, x_n]$. Se dice que G es una* ***base de Gröbner reducida*** *si, y sólo si, satisface:*

I) *Si $H \subsetneq G$ entonces $\mathbf{lt}(H) \subsetneq \mathbf{lt}(G) = \mathbf{lt}(I)$.*

II) *g_i es mónico ($\mathbf{lc}(g_i) = 1$) para $i = 1, \cdots, s$.*

III) *Para cualquier i, $1 \leq i \leq s$, el término principal $\mathbf{lt}(g_i)$ no divide a ningún término de otro g_j, para $i \neq j$.*

Proposición 5.14. Sea $I \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un ideal, $I \neq \{0\}$, y fíjese un orden monomial $\prec$, entonces existe una única base de Gröbner reducida de I respecto a $\prec$.

Demostración. Supóngase que existen dos bases de Gröbner reducidas de I, $F = \{f_1, \cdots, f_k\}$ y $G = \{g_1, \cdots, g_s\}$ ordenados tales que $\mathbf{lm}(f_1) \prec \mathbf{lm}(f_2) \prec \cdots \prec \mathbf{lm}(f_k)$ y $\mathbf{lm}(g_1) \prec \mathbf{lm}(g_2) \prec \cdots \prec \mathbf{lm}(g_s)$ y supóngase sin pérdida de la generalidad que $k \leq s$. Como ambas son bases de Gröbner entonces $\mathbf{lt}(I) = \mathbf{lt}(G) = \mathbf{lt}(F)$. Procediendo por inducción sobre r se probará que $f_r = g_r$ para $1 \leq r \leq k$.

Para $r = 1$, si $\mathbf{lt}(g_1) \neq \mathbf{lt}(f_1)$, sin pérdida de la generalidad sean $\mathbf{lt}(f_1) \prec \mathbf{lt}(g_1)$ entonces $\mathbf{lt}(f_1)$ no puede ser elemento de $\mathbf{lt}(G)$ pues todos los $\mathbf{lt}(g_i)$'s son mayores y no lo dividen. Pero $\mathbf{lt}(F) = \mathbf{lt}(G)$ $(\rightarrow\leftarrow)$, por lo que $\mathbf{lt}(f_1) = \mathbf{lt}(g_1)$. Si $f_1 - g_1 \neq 0$, como $f_1 - g_1 \in I \ \Rightarrow \mathbf{lt}(f_1 - g_1) \in \mathbf{lt}(I)$ y de ahí $\mathbf{lt}(f_1 - g_1)$ es divisible dentro de algún $\mathbf{lt}(f_i)$ y algún $\mathbf{lt}(g_j)$, pero esto es imposible, pues $\mathbf{lt}(f_1 - g_1) \prec \mathbf{lt}(f_1)$. Por lo tanto $f_1 = g_1$.

Para $r > 1$, si $\mathbf{lt}(f_r) \neq \mathbf{lt}(g_r)$, sin pérdida de la generalidad sea $\mathbf{lt}(f_r) \prec \mathbf{lt}(g_r)$ entonces $\mathbf{lt}(f_r)$ no puede ser dividido por ningún $\mathbf{lt}(g_i)$ para $i \geq r$, por ser monomios mayores; pero tampoco por ningún $\mathbf{lt}(g_i)$ para $i < r$, pues por hipótesis de inducción ,$\mathbf{lt}(g_i) = \mathbf{lt}(f_i)$ para $i < r$ y el hecho que F es una base reducida. Esto implica que $\mathbf{lt}(f_r) \notin \mathbf{lt}(G) = \mathbf{lt}(F)$ $(\rightarrow\leftarrow)$. Si $f_r - g_r \neq 0$, como $f_r - g_r \in I \ \Rightarrow \mathbf{lt}(f_r - g_r) \in \mathbf{lt}(I)$ por lo que debe ser divisible dentro de algún $\mathbf{lt}(f_i)$ y $\mathbf{lt}(g_j)$, pero esto no puede darse dado que $\mathbf{lm}(f_r - g_r) \prec \mathbf{lm}(f_r) = \mathbf{lm}(g_r)$ y debe ser un monomio de f_r o g_r, que no pueden ser divisibles porque F y G eran bases reducidas; por lo que $f_r = g_r$.

Finalmente, $f_i = g_i$ para $1 \leq r \leq k$ por lo que $F \subseteq G$, pero como $\mathbf{lt}(G)$ debe ser una base minimal de $\mathbf{lt}(I)$ la única opción es que $G = F$. ∎

Corolario 5.15. Dado un orden de monomios $\prec$ en $\mathcal{M}^n$. Dos ideales $I_1, I_2 \subseteq \mathbb{K}[x_1, \cdots, x_n]$, no $\langle 0 \rangle$, son iguales si, y sólo si, las bases reducidas de Gröbner de I_1 y I_2 respecto a $\prec$ son idénticas.

Demostración. Sean G_1 y G_2 bases reducidas de Gröbner de I_1 e I_2 respecto a $\prec$, respectivamente. Si $I_1 = I_2$, entonces el teorema anterior asegura que $G_1 = G_2$.

Por otro lado, si $G_1 = G_2$, entonces $I_1 = \langle G_1 \rangle = \langle G_2 \rangle = I_2$. ∎

El siguiente algoritmo permite transformar una base de Gröbner cualquiera a una base de Gröbner reducida.

Algoritmo 5.1 (Algoritmo para bases de Gröbner reducidas)

Input: $G = g_1, \cdots, g_s$ una base de Gröbner
Output: $H = h_1, \cdots, h_k$ una base de Gröbner reducida con $\langle H \rangle = \langle G \rangle$
begin
  $H \leftarrow G$;
  `/* Se encuentra el generador minimal de` $\mathbf{lt}(G)$ `*/`
  while *existen* $h_i, h_j \in H$ *con* $i \neq j$ *tales que* $\mathbf{lt}(h_i) | \mathbf{lt}(h_j)$ **do**
    $H \leftarrow H \setminus \{h_j\}$
  end
  `/* Se reducen los elementos de` H `respecto el resto */`
  for $i \leftarrow 1$ **to** $|H|$ **do**
    $h_i \leftarrow \mathbf{rem}(h_i; H \setminus \{h_i\})$;
    $h_i \leftarrow \dfrac{1}{\mathbf{lc}(h_i)} h_i$;
  end
end

Demostración.

- **Finitud:** Dado que $H \leftarrow G$ en la inicialización, H es un conjunto finito. Cada vez que se ingresa al ciclo **while** se elimina un elemento del conjunto H, por lo que únicamente se puede ingresar una cantidad finita de veces. Luego, el ciclo **for** se ejecuta exactamente $|H|$ veces, que es un número finito, por lo que el algoritmo evidentemente termina.

- **Correctitud:** Nótese que inicialmente $H = G$, por lo que $\mathbf{lt}(H) = \mathbf{lt}(G) = \langle G \rangle$, luego por cada repetición del ciclo **while** se elimina un elemento h_j de H si existe otro elemento $h_i \in H$ tal que $\mathbf{lt}(h_i) \,|\mathbf{lt}(h_j)$, denótese por $H' = H \setminus \{h_j\}$; como h_i permanece en H', $\mathbf{lt}(h_j) \in \mathbf{lt}(H') \Rightarrow \langle \mathbf{lt}(h_i) \,| h_i \in H \rangle = \langle \mathbf{lt}(h_i) \,| h_i \in H, h_i \neq h_j \rangle \Rightarrow \mathbf{lt}(H') = \mathbf{lt}(H) = \mathbf{lt}(G)$, propiedad que se conserva hasta que se termina el ciclo **while**. Cuando se sale del ciclo **while** ya no existe ningún elemento de H tal que su término principal sea divisible por el término principal de otro elemento de H, por lo que es imposible eliminar más elementos de H sin cambiar $\mathbf{lt}(H)$, cumpliéndose así la condición I de la definición de base de Gröbner reducida.

  Considérese ahora el ciclo **for** del algoritmo. El ciclo se repite para $i = 1, 2, \cdots, |H|$, *i.e.* para todos los elementos de H. Para un i cualquiera, el algoritmo de reducción (algoritmo 4.1) asegura la existencia de polinomios c_j's,$r \in \mathbb{K}[x_1, \cdots, x_n]$ tales que $h_i = \sum_{j \neq i} c_j h_j + r$ con r reducida respecto $H \setminus \{h_i\}$ y tales que $\mathbf{lm}(h_i) = \text{máx}_{j \neq i}(\mathbf{lm}(c_j)\,\mathbf{lm}(h_j), r)$, sin embargo $\mathbf{lm}(h_i)$ no puede ser ninguno de los $\mathbf{lm}(c_j)\,\mathbf{lm}(h_j)$ pues $\mathbf{lm}(h_j) \nmid \mathbf{lm}(h_i)$ para ningún j (esto se eliminó en el ciclo **while**) por lo que $\mathbf{lm}(h_i) = \mathbf{lm}(r) = \mathbf{lm}(\mathbf{rem}(h_i; H \setminus \{h_i\}))$. En la siguiente instrucción se divide h_i dentro de su coeficiente principal para hacerlo mónico, por lo que $\mathbf{lt}(h_i) = \mathbf{lm}(h_i)$. De esta forma el conjunto $\{\mathbf{lm}(h_1), \mathbf{lm}(h_2), \cdots, \mathbf{lm}(h_t)\}$ se conserva durante todo el ciclo **for**. Finalmente, para cada $i = 1, \cdots, |H|$, se reduce h_i respecto a $H \setminus \{h_i\}$ lo que significa que los términos de h_i no son divisibles por ningún $\mathbf{lm}(h_j)$, $j \neq i$, lo cuál se conserva hasta el final pues el ciclo **for** conserva los monomios principales de los elementos de H. Luego h_i se divide entre su coeficiente principal para hacerlo mónico. De esta forma, H cumple las condiciones II y III de la definición de base reducida de Gröbner y H es la base reducida de Gröbner de $\langle G \rangle$. ∎

El algoritmo 5.1 permite reducir una base de Gröbner para obtener una base de Gröbner reducida equivalente, en el sentido que genera el mismo ideal. En la práctica, en lugar de obtener una base de Gröbner cualquiera y luego reducirla, se suele implementar este algoritmo dentro del algoritmo de Buchberger, lo que a su vez hace que este último sea más rápido. Al implementar el algoritmo 5.1 dentro del algoritmo de Buchberger no se le aplica a bases de Gröbner, sino a conjuntos "candidatos" para ser bases de Gröbner. Esto no tiene ningún inconveniente pues se conserva el ideal $\mathbf{lt}(G)$ y únicamente se reducen sus elementos.

6 EL ALGORITMO DE BUCHBERGER

En este capítulo se presenta el algoritmo de Buchberger para calcular bases de Gröbner de ideales. También se incluye una breve discusión sobre su complejidad y algunas mejoras del algoritmo en términos de eficiencia computacional. Aunque la base fundamental del algoritmo es el criterio de Buchberger que él introdujo en su tesis doctoral de 1965 (Buchberger, 2006) junto con el mismo algoritmo, la versión presentada aquí coincide con la presentada por Cox *et al.* (2007) y Becker (1993).

6.1. S-polinomios y el criterio de Buchberger

Definición 6.1 (S-polinomios)**.** *Sean $p, q \in \mathbb{K}[x_1, \cdots, x_n]$ dos polinomios distintos de cero. Se define el* ***S-polinomio de*** *p* ***y*** *q como el polinomio:*

$$S(p,q) = \frac{mcm(\mathbf{lm}(p), \mathbf{lm}(q))}{\mathbf{lt}(p)} \cdot p - \frac{mcm(\mathbf{lm}(p), \mathbf{lm}(q))}{\mathbf{lt}(q)} \cdot q$$

***Ejemplo* 6.1**

Considere $p = x^2y + 2x^2 + y$ y $q = 3xy^2 + x - y + 1$ polinomios de $\mathbb{R}[x, y]$ con el orden DegLex. Su S-polinomio es

$$\begin{aligned} S(p,q) &= \frac{mcm(x^2y, xy^2)}{x^2y}(x^2y + 2x^2 + y) - \frac{mcm(x^2y, xy^2)}{3xy^2}(3xy^2 + x - y + 1) \\ &= \frac{x^2y^2}{x^2y}(x^2y + 2x^2 + y) - \frac{x^2y^2}{3xy^2}(3xy^2 + x - y + 1) \\ &= (x^2y^2 + 2x^2y + y^2) - (x^2y^2 + \tfrac{1}{3}x^2 - \tfrac{1}{3}xy + \tfrac{1}{3}x) \\ &= 2x^2y - \tfrac{1}{3}x^2 + \tfrac{1}{3}xy + y^2 - \tfrac{1}{3}x \end{aligned}$$

Como $S(p,q)$ es una combinación lineal de p y de q, $S(p,q) \in \langle p, q\rangle$. Para que $\{p, q\}$ sea una base de Gröbner, es necesario que $\mathbf{lm}(S(p,q)) \in \mathbf{lt}(\langle p, q\rangle)$, por lo que, siendo un ideal de monomios, $\mathbf{lm}(S(p,q))$ debe ser divisible dentro de $\mathbf{lt}(p)$ ó $\mathbf{lt}(q)$. En este ejemplo $\mathbf{lt}(p) = x^2y | \mathbf{lm}(S(p,q)) = 2x^2y$, por lo que $\{p, q\}$ es un candidato a ser base de Gröbner. Si en cambio ninguno de los términos principales de p ni de q dividiera al del $S(p,q)$, sería inmediato decir que $\{p, q\}$ no es una base de Gröbner.

Así como en el ejemplo anterior, el S-polinomio de dos polinomios es siempre una combinación lineal de los mismos en el que se eliminan sus términos principales. Sea $G = \{g_1, \cdots, g_s\} \subseteq I$ es una posible base de Gröbner para el ideal I; si es posible encontrar una combinación lineal de g_i

y g_j tal que el término principal de dicha combinación sea menor que todos los $\mathbf{lt}(g_i)$, entonces G no podría ser una base de Gröbner, pues dicho término principal no estaría en $\mathbf{lt}(G)$. Con esta motivación, Buchberger ideó un criterio que permite determinar si un conjunto dado es o no una base de Gröbner en un número finito de pasos y que se basa en el cálculo de S-polinomios. Se presenta a continuación una propiedad fundamental de los S-polinomios que permitirá demostrar la correctitud del criterio de Buchberger, siguiendo la demostración de Cox *et al.* (2007).

Lema 6.2. Sean $p_1, p_2, \cdots, p_s \in \mathbb{K}[x_1, \cdots, x_n]$ tales que $\mathbf{lm}(p_i) = m\ , i = 1, 2, \cdots, s$. Supóngase que existen constantes $c_i \in \mathbb{K}$ tales que $\mathbf{lm}\left(\sum_{i=1}^{s} c_i p_i\right) \prec m$. Entonces la suma $\sum_{i=1}^{s} c_i p_i$ puede ser escrita como una combinación lineal con coeficientes en $\mathbb{K}$ de los S-polinomios $S(p_i, p_j), i \neq j$.

Demostración. Sean $d_i = \mathbf{lc}(p_i)$ de tal forma que el coeficiente principal de $c_i p_i$ es $c_i d_i$. Como los $c_i p_i$'s tienen monomio principal m y su suma tiene monomio principal estrictamente menor que m, debe darse que $\sum_{i=1}^{s} c_i d_i = 0$. Defínanse $q_i = \frac{p_i}{d_i}$ para $i = 1, \cdots, s$, siendo los q_i's mónicos, entonces obsérvese que la suma de los $c_i p_i$'s puede reescribirse como una suma telescópica:

$$\sum_{i=1}^{s} c_i p_i = \sum_{i=1}^{s} c_i d_i q_i = c_1 d_1 (q_1 - q_2) + (c_1 d_1 + c_2 d_2)(q_2 - q_3) + (c_1 d_1 + c_2 d_2 + c_3 d_3)(q_3 - q_4) +$$
$$+ \cdots + (c_1 d_1 + \cdots + c_{s-1} d_{s-1})(q_s - q_{s-1}) + (c_1 d_1 + \cdots + c_s d_s) q_s$$

Nótese que $mcm(\mathbf{lm}(p_i), \mathbf{lm}(p_j)) = m$ por lo que

$$S(p_i, p_j) = \frac{m}{\mathbf{lt}(p_i)} \cdot p_i - \frac{m}{\mathbf{lt}(p_j)} \cdot p_j = \frac{m}{d_i m} \cdot p_i - \frac{m}{d_j m} \cdot p_j = q_i - q_j$$

Sean $k_i = \sum_{j=1}^{i} c_j d_j$, luego como $\sum_{i=1}^{s} c_i d_i = 0$ se obtiene lo buscado:

$$\sum_{i=1}^{s} c_i p_i = k_1 S(p_1, p_2) + k_2 S(p_2, p_3) + \cdots + k_{s-1} S(p_{s-1}, p_s)$$ ∎

Teorema 6.3 (Criterio de Buchberger)**.** Sea $G = \{g_1, \cdots, g_s\} \subseteq \mathbb{K}[x_1, \cdots, x_n]$. G es una base de Gröbner de $\langle G \rangle$ si, y sólo si, para todas las parejas $1 \leq i, j \leq s$, $i \neq j$ el S-polinomio $S(g_i, g_j)$ se reduce a cero, *i.e.* $S(g_i, g_j) \xrightarrow{G} 0$.

Demostración. ($\Rightarrow$) Si G es una base de Gröbner, $S(g_i, g_j) \in \langle G \rangle$ pues es una combinación lineal de g_i y g_j, entonces $\mathbf{rem}(S(g_i, g_j); G) = 0\ \forall\ i, j$ por la propiedad 5.8.

($\Leftarrow$) Sea $p \in \langle G \rangle$ un polinomio distinto de cero. Suponiendo que todos los S-polinomios de elementos de G tienen residuo 0 módulo G se probará que $\mathbf{lt}(p) \in \mathbf{lt}(G)$. Como $p \in \langle G \rangle$, deben existir polinomios $c_i \in \mathbb{K}[x_1, \cdots, x_n]$ tales que $p = \sum_{i=1}^{s} c_i g_i$ por lo que

$$\mathbf{lm}(p) \preceq \text{máx}(\mathbf{lm}(c_1)\,\mathbf{lm}(g_1), \cdots, \mathbf{lm}(c_s)\,\mathbf{lm}(g_s)).$$

Considere ahora el conjunto

$$\mathfrak{M} = \left\{ m \,\middle|\, m = \text{máx}(\mathbf{lm}(c_1)\,\mathbf{lm}(g_1), \cdots, \mathbf{lm}(c_s)\,\mathbf{lm}(g_s)) \text{ y } p = \sum_{i=1}^{s} c_i g_i \right\}$$

Como $\mathfrak{M} \subseteq \mathcal{M}^n$ es un conjunto no vacío y el orden monomial $\prec$ es un buen ordenamiento, debe tener un elemento mínimo, m_0 que se alcanza con los polinomios $c_1, \cdots, c_s$. Supóngase por el absurdo que $\mathbf{lm}(p) \prec m_0$ entonces

$$p = \sum_{i=1}^{s} c_i g_i = \sum_{\mathbf{lm}(c_i g_i) = m_0} c_i g_i + \sum_{\mathbf{lm}(c_i g_i) \prec m_0} c_i g_i =$$

$$= \sum_{\mathbf{lm}(c_i g_i) = m_0} \mathbf{lt}(c_i)\, g_i + \sum_{\mathbf{lm}(c_i g_i) = m_0} (c_i - \mathbf{lt}(c_i)) g_i + \sum_{\mathbf{lm}(c_i g_i) \prec m_0} c_i g_i$$

donde es evidente que la segunda y tercera suma tienen monomio principal menor que m_0, y como $\mathbf{lm}(p) \prec m_0$ también la primera suma debe tener monomio principal menor que m_0. Así, la primera suma cumple con las condiciones del lema anterior, y por lo tanto debe poder escribirse como combinación lineal de los S-polinomios $S(\boldsymbol{x}^{\boldsymbol{\alpha}_i} g_i, \boldsymbol{x}^{\boldsymbol{\alpha}_j} g_j)$ donde $a_i \boldsymbol{x}^{\boldsymbol{\alpha}_i} = \mathbf{lt}(c_i)$ con $a_i \in \mathbb{K}$. Como $mcm(\boldsymbol{x}^{\boldsymbol{\alpha}_i} g_i, \boldsymbol{x}^{\boldsymbol{\alpha}_j} g_j) = mcm(m_0, m_0) = m_0$, entonces se tiene

$$S(\boldsymbol{x}^{\boldsymbol{\alpha}_i} g_i, \boldsymbol{x}^{\boldsymbol{\alpha}_j} g_j) = \frac{m_0}{\boldsymbol{x}^{\boldsymbol{\alpha}_i}\mathbf{lt}(g_i)} \boldsymbol{x}^{\boldsymbol{\alpha}_i} g_i - \frac{m_0}{\boldsymbol{x}^{\boldsymbol{\alpha}_j}\mathbf{lt}(g_j)} \boldsymbol{x}^{\boldsymbol{\alpha}_i} g_j =$$

$$= \frac{m_0}{m_{ij}} \left(\frac{m_{ij}}{\mathbf{lt}(g_i)} g_i - \frac{m_{ij}}{\mathbf{lt}(g_j)} g_j \right) = \frac{m_0}{m_{ij}} S(g_i, g_j)$$

donde $m_{ij} = mcm(g_i, g_j) \preceq m_0$. Como por hipótesis los S-polinomios se reducen a cero, $S(g_i, g_j) \xrightarrow{G} 0$ entonces existen polinomios h_{ijk} tales que

$$S(\boldsymbol{x}^{\boldsymbol{\alpha}_i} g_i, \boldsymbol{x}^{\boldsymbol{\alpha}_j} g_j) = \frac{m_0}{m_{ij}} S(g_i, g_j) = \sum_k \frac{m_0}{m_{ij}} h_{ijk} g_k$$

con

$$\mathbf{lm}\left(\frac{m_0}{m_{ij}} h_{ijk} g_k\right) = \mathbf{lm}\left(\frac{m_0}{m_{ij}}\right) \mathbf{lm}(h_{ijk} g_k) \preceq \mathbf{lm}\left(\frac{m_0}{m_{ij}}\right) \mathbf{lm}(S(g_i, g_j)) = \mathbf{lm}(S(\boldsymbol{x}^{\boldsymbol{\alpha}_i} g_i, \boldsymbol{x}^{\boldsymbol{\alpha}_j} g_j)) \prec$$

$$\prec \text{máx}(\boldsymbol{x}^{\boldsymbol{\alpha}_i} g_i, \boldsymbol{x}^{\boldsymbol{\alpha}_j} g_j) = m_0 \text{ para todo } k.$$

Por lo que cada S-polinomio $S(\boldsymbol{x}^{\boldsymbol{\alpha}_i} g_i, \boldsymbol{x}^{\boldsymbol{\alpha}_j} g_j)$ se puede escribir como una combinación lineal de los g_k's donde cada sumando tiene monomio principal menor a m_0. Uniendo esto con el hecho que $\sum \mathbf{lt}(c_i)\, g_i$ es una $\mathbb{K}$-combinación lineal de estos y la fórmula deducida para p, se obtiene que

$$p = \sum_i c_i' g_i$$

con $\text{máx}(\mathbf{lm}(c_i' g_i)) \prec m_0$ y $\text{máx}(\mathbf{lm}(c_j' g_j)) \in \mathfrak{M}$ $(\rightarrow\leftarrow)$. Por lo tanto $\mathbf{lm}(p) = m_0 = \text{máx}(\mathbf{lm}(c_i g_i)) \Rightarrow$ $\mathbf{lm}(p) = \mathbf{lm}(c_i g_i)$ para algún i, lo que significa que $\mathbf{lm}(p) \in \mathbf{lt}(G)$. Por la arbitrariedad de p, $\mathbf{lt}(\langle G \rangle) \subseteq \mathbf{lt}(G) \Rightarrow \mathbf{lt}(\langle G \rangle) = \mathbf{lt}(G) \Rightarrow G$ es una base de Gröbner. ∎

Aunque el criterio requiere que $S(g_i, g_j) \xrightarrow{G} 0$, para implementar el criterio como un mecanismo de prueba, basta con calcular los residuos $\mathbf{rem}(S(g_i, g_j); G)$. Si $\mathbf{rem}(S(g_i, g_j); G) = 0$ evidente-

mente $S(g_i, g_j) \xrightarrow{G} 0$ y G es base de Gröbner; si en cambio para algún S-polinomio se obtuviera $\mathbf{rem}(S(g_i, g_j); G) \neq 0$ en realidad se tendría que G no es una base de Gröbner, pues de serlo el residuo debería ser único y, por el criterio de Buchberger, tendría que ser cero.

Corolario 6.4. Si $G \subseteq \mathbb{K}[x_1, \cdots, x_n]$ es un conjunto finito de monomios, entonces G es una base de Gröbner.

Demostración. Sean $a\boldsymbol{x}^{\boldsymbol{\alpha}}, b\boldsymbol{x}^{\boldsymbol{\beta}} \in G$ dos elementos cualesquiera distintos de G, entonces

$$\begin{aligned} S(a\boldsymbol{x}^{\boldsymbol{\alpha}}, b\boldsymbol{x}^{\boldsymbol{\beta}}) &= \frac{mcm\big(a\boldsymbol{x}^{\boldsymbol{\alpha}}, b\boldsymbol{x}^{\boldsymbol{\beta}}\big)}{\mathbf{lt}(a\boldsymbol{x}^{\boldsymbol{\alpha}})} a\boldsymbol{x}^{\boldsymbol{\alpha}} - \frac{mcm\big(a\boldsymbol{x}^{\boldsymbol{\alpha}}, b\boldsymbol{x}^{\boldsymbol{\beta}}\big)}{\mathbf{lt}(b\boldsymbol{x}^{\boldsymbol{\beta}})} b\boldsymbol{x}^{\boldsymbol{\beta}} = \\ &= \frac{mcm\big(a\boldsymbol{x}^{\boldsymbol{\alpha}}, b\boldsymbol{x}^{\boldsymbol{\beta}}\big)}{a\boldsymbol{x}^{\boldsymbol{\alpha}}} a\boldsymbol{x}^{\boldsymbol{\alpha}} - \frac{mcm\big(a\boldsymbol{x}^{\boldsymbol{\alpha}}, b\boldsymbol{x}^{\boldsymbol{\beta}}\big)}{b\boldsymbol{x}^{\boldsymbol{\beta}}} b\boldsymbol{x}^{\boldsymbol{\beta}} = \\ &= mcm\big(a\boldsymbol{x}^{\boldsymbol{\alpha}}, b\boldsymbol{x}^{\boldsymbol{\beta}}\big) - mcm\big(a\boldsymbol{x}^{\boldsymbol{\alpha}}, b\boldsymbol{x}^{\boldsymbol{\beta}}\big) = 0 \end{aligned}$$

por lo que G es una base de Gröbner. ∎

Este resultado permite construir bases de Gröbner para algunos ideales sencillos, lo que permite a continuación dar un ejemplo para el teorema 5.12, el cual había quedado pendiente por falta de un método para construir dicha base.

***Ejemplo* 6.2**

A manera de ejemplo considérese el anillo de polinomios $\mathbb{R}[x, y, z]$ con el orden de monomios DegLex y el ideal $I = \langle x^2 + y + z, y^3 - y^2, xyz + y, y^2 + z\rangle$. El conjunto $G = \{x^2, y, z\}$ es una base de Gröbner de I, por lo que $R = \mathbb{R}[x, y, z]/I = \mathbb{R}[x, y, z]/\langle x^2, y, z\rangle$ y, según el teorema anterior, tiene como base de espacio vectorial al conjunto de monomios que no pertenecen a $\mathbf{lt}(G) = \langle x^2, y, z\rangle$, sin embargo los únicos monomios de $\mathbb{R}[x, y, z]$ que no son divisibles por x^2, y, z son 1 y x, entonces el anillo cociente R es en realidad el espacio vectorial $\{\alpha + \beta x | \alpha, \beta \in \mathbb{R}\} \simeq \mathbb{R}^2$ de dimensión 2.

6.2. El algoritmo de Buchberger

El critero de Buchberger provee un mecanismo para determinar si un conjunto es una base de Gröbner o no en una cantidad finita de pasos. Si se está examinando un conjunto G y el S-polinomio de dos de sus elementos no se reduce a cero, G no es una base de Gröbner, pero se puede intentar solventar este problema agregando dicho S-polinomio al conjunto, forzándolo así a reducirse a cero. El algoritmo de Buchberger explota esta idea y la genialidad de la tesis de Buchberger de 1965 es que en ella se demuestra que este procedimiento en efecto llega a producir una base de Gröbner a partir de cualquier conjunto G.

Algoritmo 6.1 (Algoritmo de Buchberger)

Input: $F \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un conjunto finito de polinomios no cero

Output: $G \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un conjunto finito de polinomios tal que G es una base de Gröbner de $\langle F \rangle$

begin
- $G \leftarrow F$;
- $B \leftarrow \{\{g_i, g_j\} \,|g_i, g_j \in G\ , i < j\}$;
- **while** $B \neq \emptyset$ **do**
  - seleccione $\{g_i, g_j\} \in B$;
  - $B \leftarrow B \setminus \{\{g_i, g_j\}\}$;
  - $h \leftarrow S(g_i, g_j)$;
  - $r \leftarrow \mathbf{rem}(h; G)$;
  - **if** $r \neq 0$ **then**
    - $B \leftarrow B \cup \{\{g, r\} \mid g \in G\}$;
    - $G \leftarrow G \cup \{r\}$
  - **end**
- **end**

end

Demostración.

- **Finitud:** Al iniciar el algoritmo se hace $G \leftarrow F$. Como G es un conjunto finito, es evidente que inicialmente el conjunto B de todas las parejas de elementos de G es también finito. Obsérvese que dentro del ciclo **while** se agranda el conjunto G cada vez que $r \neq 0$. Sean $G^{(0)}, G^{(1)}, G^{(2)}, \cdots$ los valores que toma G en las iteraciones del ciclo **while** en las que en efecto se le agrega un elemento. Como por hipótesis para agregar r a $G^{(t)}$ se debe dar $r = \mathbf{rem}\big(h; G^{(t)}\big) \neq 0$ entonces $r \notin G^{(t)}$, por lo que $G^{(t+1)} = G^{(t)} \cup \{r\}$ es estrictamente mayor. Más aún, como r es reducida respecto a $G^{(t)}$, ninguno de sus monomios es divisible dentro de $\mathbf{lt}(g)$ para $g \in G^{(t)}$, por lo que $\mathbf{lt}(r) \notin \mathbf{lt}\big(G^{(t)}\big)$ y $\mathbf{lt}\big(G^{(t)}\big) \subsetneq \mathbf{lt}\big(G^{(t+1)}\big)$ para todo $1 \leq t$. Así $\mathbf{lt}\big(G^{(0)}\big) \subsetneq \mathbf{lt}\big(G^{(1)}\big) \subsetneq \mathbf{lt}\big(G^{(2)}\big) \subsetneq \cdots$ es una cadena estrictamente ascendente de ideales en $\mathbb{K}[x_1, \cdots, x_n]$ que es un anillo Noetheriano, por lo que la condición de la cadena ascendente asegura que es una secuencia estacionaria y eventualmente el conjunto G deja de crecer en el ciclo **while**. Cuando esto sucede, quiere decir que $r = 0$ en cada iteración, por lo que no se entra en el condicional **if** y tampoco se agregan elementos al conjunto B, únicamente se elimina un elemento de B por cada iteración del ciclo **while**; como B es finito, eventualmente $B = \emptyset$ y se finaliza el algoritmo.

- **Correctitud:** Sea G' el resultado del conjunto G al finalizar el algoritmo. Claramente el conjunto G' contiene a los elementos de F junto con los residuos $r \neq 0$ que se hayan agregado

durante el ciclo **while** y $G \subseteq G'$ durante todo el algoritmo. Se mostrará ahora que para cualquier pareja de elementos $\{g_1, g_2\}$ que haya estado en el conjunto B en algún momento del algoritmo, $\mathbf{rem}(S(g_1, g_2); G') = 0$. Si en un momento de la ejecución del algoritmo $\{g_1, g_2\} \in B$, como el algoritmo se termina cuando $B = \emptyset$ entonces dicha pareja debe eliminarse de B en algún momento más adelante. Durante esa iteración del ciclo **while**, se elimina la pareja de B y se calculan $h = S(g_1, g_2)$ y $r = \mathbf{rem}(h; G)$. Si $r = 0$ evidentemente h también se reduce a cero respecto a G' (pues $G \subseteq G'$); si por el contrario $r \neq 0$ entonces r se introduce al conjunto G, por lo que $r \in G'$ y entonces h se deberá reducir a cero respecto G' nuevamente.

Ahora, para poder usar el criterio de Buchberger y demostrar que G' es una base de Gröbner, basta probar que todas las parejas de elementos de G' fueron elementos de B en algún momento de la ejecución, pues el argumento anterior demuestra que su S-polinomio se reduce a cero respecto a G'. Sean entonces g_1, g_2 dos elementos distintos cualesquiera de G', se puede dar uno de tres casos:

1. $g_i, g_j \in F$. Como en la inicialización $G \leftarrow F$ y luego se construye B con todas las parejas de elementos de G entonces $\{g_i, g_j\} \in B$ en ese momento.
2. Sin pérdida de la generalidad, $g_i \in F$ y g_j se agregó más adelante. Para que g_j se haya agregado a G, quiere decir que en algún momento se ejecutó la condición **if** con $g_j = r \neq 0$, y en ese momento se agregaron también a B todas las parejas de la forma $\{g, r\} = \{g, g_j\}$ para $g \in G$, sin embargo, como $g_i \in F \subseteq G$ entonces particularmente se agregó la pareja $\{g_i, g_j\}$ a B.
3. Tanto g_i como g_j se agregaron a G' luego de la inicialización. Sin pérdida de la generalidad supóngase que g_i se agregó antes que g_j, entonces en el momento en el que se agregó g_j, ya se tenía $g_i \in G$ y al agregar las parejas a B de la forma $\{g, r\} = \{g, g_j\}$ para $g \in G$ se agregó también $\{g_i, g_j\}$, por lo que $\{g_i, g_j\} \in B$ en algún instante del algoritmo.

Por lo tanto, G' es una base de Gröbner. Más aún, como $F \subseteq G'$ entonces $\langle F \rangle \subseteq \langle G' \rangle$, sin embargo cada elemento que se agregó a G era la reducción respecto a G de un S-polinomio de elementos de G, *i.e.* era una combinación lineal de elementos de G y, en última instancia, de F, por lo que $G' \subseteq \langle F \rangle$ y entonces $\langle G' \rangle = \langle F \rangle$ como se buscaba. ∎

***Ejemplo* 6.3**

Considere el anillo $\mathbb{Q}[x, y]$ con el orden DegLex y sea $I = \langle f_1, f_2 \rangle = \langle x^3 - 2xy, x^2y - 2y^2 + x \rangle$. La Tabla 6.1 muestra la ejecución del algoritmo de Buchberger para encontrar una base de Gröbner G del ideal I.

El algoritmo devuelve como resultado $G = \{x^3 - 2xy, x^2y - 2y^2 + x, -x^2, -2xy, -4y^2 + 2x\}$, una base de Gröbner de $\langle I \rangle$. Esto pues, cómo se hizo en la demostración de la correctitud del

Tabla 6.1: Ejemplo algoritmo de Buchberger (algoritmo 6.1)

Paso	h	r	G	**Obs.**
0			$f_1 = x^3 - 2xy, f_2 = x^2y - 2y^2 + x$	Inicio
1	$S(f_1, f_2) = -x^2$	$-x^2$	$f_1 = x^3 - 2xy, f_2 = x^2y - 2y^2 + x$	$r \neq 0$
2	$S(f_1, f_3) = -2xy$	$-2xy$	$f_1 = x^3 - 2xy, f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2$	$r \neq 0$
3	$S(f_1, f_4) =$ $-4y^2 + 2x$	$-4y^2 + 2x$	$f_1 = x^3 - 2xy, f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2, f_4 = -2xy$	$r \neq 0$
4	$S(f_1, f_5) =$ $2x^4$–$8xy^3$	0	$f_1 = x^3 - 2xy, f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2, f_4 = -2xy, f_5 = -4y^2 + 2x$	$r = 0$
5	$S(f_2, f_3) =$ $-2y^2 + x$	0	$f_1 = x^3 - 2xy, f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2, f_4 = -2xy, f_5 = -4y^2 + 2x$	$r = 0$
6	$S(f_2, f_4) =$ $-4y^2 + 2x$	0	$f_1 = x^3 - 2xy, f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2, f_4 = -2xy, f_5 = -4y^2 + 2x$	$r = 0$
7	$S(f_2, f_5) =$ $2x^3 - 8y^3 + 4xy$	0	$f_1 = x^3 - 2xy, f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2, f_4 = -2xy, f_5 = -4y^2 + 2x$	$r = 0$
8	$S(f_3, f_4) = 0$	0	$f_1 = x^3 - 2xy, f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2, f_4 = -2xy, f_5 = -4y^2 + 2x$	$r = 0$
9	$S(f_3, f_5) = 2x^3$	0	$f_1 = x^3 - 2xy, f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2, f_4 = -2xy, f_5 = -4y^2 + 2x$	$r = 0$
10	$S(f_4, f_5) = 2x^2$	0	$f_1 = x^3 - 2xy, f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2, f_4 = -2xy, f_5 = -4y^2 + 2x$	$r = 0$

algoritmo de Buchberger, todos los S-polinomios de elementos de G se calculan en algún momento, y en ese instante o se reducen a cero o se agregan (reducidos módulo G) al conjunto G.

6.3. Mejoras al algoritmo de Buchberger

Como se verá en la próxima sección al discutir la complejidad del algoritmo de Buchberger, este es un algoritmo computacionalmente muy costoso, por lo que para poder implementarlo es deseable hacerle tantas mejoras como sea posible. El algoritmo 6.1 es una de las primeras versiones que surge históricamente del mismo (Becker, 1993) por lo cual calcula una gran cantidad de reducciones respecto a G, algo computacionalmente costoso, sin considerar que se pueden estar realizando cálculos superfluos si se cae en uno de dos casos: 1) calcular la reducción de un S-polinomio que era sencillo determinar que iba a ser cero; o 2) calcular la reducción de un S-polinomio mediante el algoritmo de reducción obteniendo un valor distinto a cero cuando utilizando los divisores en otro orden hubiera sido posible reducirlo a cero; por lo que se agrega un elemento a G y muchas parejas a B innecesariamente. Los criterios que se verán a continuación, que para efectos de referencia se llamarán Segundo y Tercer criterio de Buchberger, junto con la integración del algoritmo 5.1 que se discutía en el capítulo anterior, permitirán evitar estos cálculos innecesarios para ahorrar tiempo y recursos.

Definición 6.5. *Dados dos polinomios $p, q \in \mathbb{K}[x_1, \cdots, x_n]$ distintos de cero, se dice que p **y** q **son disjuntos** si $\mathbf{lm}(p)$ y $\mathbf{lm}(q)$ no tienen ninguna variable en común, i.e. si mcd$(\mathbf{lm}(p), \mathbf{lm}(q)) = 1$*

(o equivalentemente $mcm(\mathbf{lm}(p)\,,\mathbf{lm}(q)) = \mathbf{lm}(p)\,\mathbf{lm}(q) = \mathbf{lm}(pq)$*).*

Teorema 6.6 (Segundo Criterio de Buchberger)**.** Sean $p, q \in \mathbb{K}[x_1, \cdots, x_n]$ dos polinomios disjuntos distintos de cero, entonces $S(p,q) \xrightarrow{\{p,q\}} 0$.

Demostración. Considérese sin pérdida de la generalidad que p y q son mónicos, pues es evidente que si $S(p,q) \xrightarrow{\{p,q\}} 0$ entonces también se debe cumplir $S(ap, bq) \xrightarrow{\{ap,bq\}} 0$. Sean entonces $p = \mathbf{lm}(p) + p_1$ y $q = \mathbf{lm}(q) + q_1$ con $\mathbf{lm}(p_1) \prec \mathbf{lm}(p)$ y $\mathbf{lm}(q_1) \prec \mathbf{lm}(p)$. Como p y q son disjuntos se tiene que $mcm(\mathbf{lm}(p)\,,\mathbf{lm}(q)) = \mathbf{lm}(p)\,\mathbf{lm}(q)$ y entonces

$$\begin{aligned} S(p,q) &= \frac{\mathbf{lm}(p)\,\mathbf{lm}(q)}{\mathbf{lt}(p)} \cdot p - \frac{\mathbf{lm}(p)\,\mathbf{lm}(q)}{\mathbf{lt}(q)} \cdot q = \mathbf{lm}(q)\,p - \mathbf{lm}(p)\,q = \\ &= (q - q_1)p - (p - p_1)q = qp - q_1p - pq + p_1q = \\ &= -q_1p + p_1q = p_1q - q_1p \end{aligned}$$

Por lo que $S(p,q)$ es combinación lineal de p y q, y únicamente resta probar que $\mathbf{lm}(S(p,q)) = \text{máx}(\mathbf{lm}(p_1q)\,,\mathbf{lm}(q_1p))$. Para que este no fuera el caso, como $S(p,q) = p_1q - q_1p$ se tendrían que cancelar los términos principales de los sumandos, pero para ello sería necesario que $\mathbf{lm}(p_1q) = \mathbf{lm}(q_1p)$ o equivalentemente, $\mathbf{lm}(p_1)\,\mathbf{lm}(q) = \mathbf{lm}(q_1)\,\mathbf{lm}(p)$; entonces $\mathbf{lm}(p_1)\,\mathbf{lm}(q)$ sería divisible por $\mathbf{lm}(q)$ y $\mathbf{lm}(p)$, y además como $\mathbf{lm}(p_1) \prec \mathbf{lm}(p) \Rightarrow \mathbf{lm}(p_1)\,\mathbf{lm}(q) \prec \mathbf{lm}(p)\,\mathbf{lm}(q) = mcm(\mathbf{lm}(p)\,,\mathbf{lm}(q))$ lo cual es una contradicción $(\rightarrow\leftarrow)$. Por lo tanto, $S(p,q) \xrightarrow{\{p,q\}} 0$. ∎

Definición 6.7. *Sean* $p \in \mathbb{K}[x_1, \cdots, x_n]$*,* G *un subconjunto finito no vacío de* $\mathbb{K}[x_1, \cdots, x_n]$ *y* $m \in \mathcal{M}^n$*. La representación de* p *como*

$$p = \sum_{g_i \in G} c_i g_i$$

con $c_i \in \mathbb{K}[x_1, \cdots, x_n]$*, es una* m***-representación de*** p ***respecto a*** G *si, y sólo si,para cada* i *se cumple* $c_i = 0$ *ó* $\mathbf{lm}(c_ig_i) \preceq m$*.*

Teorema 6.8 (Tercer Criterio de Buchberger)**.** Sean $G = \{g_1, \cdots, g_s\}$ un subconjunto finito de $\mathbb{K}[x_1, \cdots, x_n]$ y $g_1, p, g_2 \in \mathbb{K}[x_1, \cdots, x_n]$ polinomios distintos de cero que cumplen con:

I) $\mathbf{lm}(p)\,|mcm(\mathbf{lm}(g_1)\,,\mathbf{lm}(g_2))$, y

II) Para $i = 1, 2$, $S(g_i, p)$ tiene una m_i-representación respecto a G, para algún $m_i \in \mathcal{M}^n$ tal que $m_i \preceq mcm(\mathbf{lm}(g_i)\,,\mathbf{lm}(p))$.

Entonces el S-polinomio $S(g_1, g_2)$ tiene una m-representación respecto a G, para algún $m \in \mathcal{M}^n$ tal que $m \preceq mcm(\mathbf{lm}(g_1)\,,\mathbf{lm}(g_2))$.

Demostración. Por la condición II) existen representaciones

$$S(g_1,p)=\sum_{i=1}^{s}c_ig_i,\ \text{con } c_i=0 \text{ ó } \mathbf{lm}(c_ig_i)\preceq m_1,$$
$$S(g_2,p)=\sum_{i=1}^{s}c_i'g_i,\ \text{con } c_i'=0 \text{ ó } \mathbf{lm}(c_i'g_i)\preceq m_2,$$

con $m_1\preceq mcm(\mathbf{lm}(g_1),\mathbf{lm}(p))$ y $m_2\preceq mcm(\mathbf{lm}(g_2),\mathbf{lm}(p))$ para polinomios $c_i,c_i'\in\mathbb{K}[x_1,\cdots,x_n]$ para $i=1,\cdots,s$.

Por la condición I), $\mathbf{lm}(p)\,|mcm(\mathbf{lm}(g_1),\mathbf{lm}(g_2))$. Además, se sabe que $\mathbf{lm}(g_1)\,|mcm(\mathbf{lm}(g_1),\mathbf{lm}(g_2))$ y de la propiedad 4.17 se tiene que

$$mcm(\mathbf{lm}(p),\mathbf{lm}(g_1))\,|mcm(\mathbf{lm}(g_1),\mathbf{lm}(g_2)).$$

Análogamente, se obtiene también que

$$mcm(\mathbf{lm}(p),\mathbf{lm}(g_2))\,|mcm(\mathbf{lm}(g_1),\mathbf{lm}(g_2)),$$

por lo que deben existir $s_1,s_2\in\mathcal{M}^n$ tales que

$$s_1mcm(\mathbf{lm}(g_1),\mathbf{lm}(p))=mcm(\mathbf{lm}(g_1),\mathbf{lm}(g_2)),$$
$$s_2mcm(\mathbf{lm}(g_2),\mathbf{lm}(p))=mcm(\mathbf{lm}(g_1),\mathbf{lm}(g_2)).$$

Sean ahora $u_1,u_2,v_1,v_2\in\mathcal{M}^n$ tales que

$$mcm(\mathbf{lm}(g_1),\mathbf{lm}(p))=u_1\mathbf{lm}(g_1)=v_1\mathbf{lm}(p)$$
$$mcm(\mathbf{lm}(g_2),\mathbf{lm}(p))=u_2\mathbf{lm}(g_2)=v_2\mathbf{lm}(p)$$

gracias a la propiedad 4.17. Entonces

$$s_1v_1\mathbf{lm}(p)=s_1mcm(\mathbf{lm}(g_1),\mathbf{lm}(p))=mcm(\mathbf{lm}(g_1),\mathbf{lm}(g_2))=$$
$$=s_2mcm(\mathbf{lm}(g_2),\mathbf{lm}(p))=s_2v_2\mathbf{lm}(p)$$

$\Rightarrow s_1v_1=s_2v_2$. Nótese entonces que

$$S(g_1,g_2)=\frac{mcm(\mathbf{lm}(g_1),\mathbf{lm}(g_2))}{\mathbf{lt}(g_1)}\cdot g_1-\frac{mcm(\mathbf{lm}(g_1),\mathbf{lm}(g_2))}{\mathbf{lt}(g_2)}\cdot g_2=$$
$$=\frac{s_1mcm(\mathbf{lm}(g_1),\mathbf{lm}(p))}{\mathbf{lt}(g_1)}\cdot g_1-\frac{s_2mcm(\mathbf{lm}(g_2),\mathbf{lm}(p))}{\mathbf{lt}(g_2)}\cdot g_2=$$

$$=\frac{s_1mcm(\mathbf{lm}(g_1),\mathbf{lm}(p))}{\mathbf{lt}(g_1)}\cdot g_1+\left(-\frac{s_1v_1\mathbf{lm}(p)}{\mathbf{lt}(p)}\cdot p+\frac{s_2v_2\mathbf{lm}(p)}{\mathbf{lt}(p)}\cdot p\right)-\frac{s_2mcm(\mathbf{lm}(g_2),\mathbf{lm}(p))}{\mathbf{lt}(g_2)}\cdot g_2=$$

$$= \left(\frac{s_1 mcm(\mathbf{lm}(g_1), \mathbf{lm}(p))}{\mathbf{lt}(g_1)} \cdot g_1 - \frac{s_1 v_1 \mathbf{lm}(p)}{\mathbf{lt}(p)} \cdot p \right) + \\ + \left(\frac{s_2 v_2 \mathbf{lm}(p)}{\mathbf{lt}(p)} \cdot p - \frac{s_2 mcm(\mathbf{lm}(g_2), \mathbf{lm}(p))}{\mathbf{lt}(g_2)} \cdot g_2 \right) =$$

$$= \left(\frac{s_1 mcm(\mathbf{lm}(g_1), \mathbf{lm}(p))}{\mathbf{lt}(g_1)} \cdot g_1 - \frac{s_1 mcm(\mathbf{lm}(g_1), \mathbf{lm}(p))}{\mathbf{lt}(p)} \cdot p \right) + \\ + \left(\frac{s_2 mcm(\mathbf{lm}(g_2), \mathbf{lm}(p))}{\mathbf{lt}(p)} \cdot p - \frac{s_2 mcm(\mathbf{lm}(g_2), \mathbf{lm}(p))}{\mathbf{lt}(g_2)} \cdot g_2 \right) =$$

$$= s_1 S(g_1, p) + s_2 S(p, g_2).$$

Luego,

$$S(g_1, g_2) = s_1 S(g_1, p) + s_2 S(p, g_1) = \\ = s_1 \sum_{i=1}^{s} c_i g_i + s_2 \sum_{i=1}^{s} c_i' g_i = \sum_{i=1}^{s} (s_1 c_i + s_2 c_i) g_i.$$

Finalmente, si $s_1 c_i + s_2 c_i' \neq 0$ entonces

$$\mathbf{lm}((s_1 c_i + s_2 c_i') g_i) \preceq \text{máx}(\mathbf{lm}(s_1 c_i g_i), \mathbf{lm}(s_2 c_i' g_i)) = \text{máx}(s_1 \mathbf{lm}(c_i g_i), s_2 \mathbf{lm}(c_i' g_i) \\ \preceq \text{máx}(s_1 m_1, s_2 m_2) \preceq \text{máx}(s_1 mcm(\mathbf{lm}(g_1), \mathbf{lm}(p)), s_2 mcm(\mathbf{lm}(g_2), \mathbf{lm}(p))) \\ \preceq \text{máx}(mcm(\mathbf{lm}(g_1), \mathbf{lm}(g_2)), mcm(\mathbf{lm}(g_1), \mathbf{lm}(g_2))) = mcm(\mathbf{lm}(g_1), \mathbf{lm}(g_2))$$

Así, se tiene una m-representación de $S(g_1, g_2)$ respecto a G, para algún $m \preceq mcm(\mathbf{lm}(g_1), \mathbf{lm}(g_2))$, lo que a su vez implica que $S(g_1, g_2) \in \langle G \rangle$ ∎

El siguiente algoritmo presenta una mejora al algoritmo de Buchberger 6.1 al incorporar el segundo y tercer criterio de Buchberger. Como menciona Becker (1993), existen múltiples formas de incorporar el tercer criterio, el algoritmo que se muestra incluye la llamada **estrategia normal**. La clave del algoritmo es que guarda registro de las parejas de polinomios $\{g_1, g_2\} \subseteq G$ que se seleccionan en el ciclo **while** marcándolas como *tratadas* en una matriz booleana. Si los polinomios g_1 y g_2 son disjuntos, el algoritmo ni siquiera agrega la pareja al conjunto B, pero sí la marca como tratada. Luego, durante la ejecución del ciclo **while**, las parejas de B se seleccionan tomando primero las que tienen monomio principal mínimo. Cuando esto sucede, se examina primero si es posible encontrar $p \in G$ tal que $\mathbf{lm}(p) \,|mcm(\mathbf{lm}(g_1), \mathbf{lm}(g_2))$ y que las parejas $\{g_1, p\}$ y $\{p, g_2\}$ estén marcadas como *tratadas*. Si este es el caso, por el tercer criterio de Buchberger se sabe que $S(g_1, g_2) \xrightarrow{G} 0$ por lo que el algoritmo marca directamente la pareja como *tratada*. En caso contrario, la pareja es tratada tal y como se hacía en el algoritmo de Buchberger original (algoritmo 6.1) y luego de esto ya es marcada como *tratada*.

Algoritmo 6.2 (Algoritmo de Buchberger mejorado)

Input: $F \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un conjunto finito de polinomios no cero

Output: $G \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un conjunto finito de polinomios tal que G es una base de Gröbner de $\langle F \rangle$

begin

`/* INICIALIZACIÓN */`

$G \leftarrow$ **RED**(F), donde **RED** es el algoritmo 5.1 para bases de Gröbner Reducidas;

$B \leftarrow \{\{g_i, g_j\} \,|g_i, g_j \in G$ no disjuntos y con $i < j\}$;

$M \leftarrow$ una matriz booleana con entradas para cada pareja $\{g_i, g_j\} \in G$ con $i < j$;

foreach $\{g_i, g_j\} \subseteq G$ $i < j$ **do**
  if $\{g_i, g_j\} \in B$ **then** $M(g_i, g_j) \leftarrow$**false**; **else** $M(g_i, g_j) \leftarrow$**true**;

end

`/* CUERPO DEL ALGORITMO */`

while $B \neq \emptyset$ **do**

seleccione $\{g_i, g_j\} \in B$ con $mcm(\mathbf{lm}(g_i), \mathbf{lm}(g_j))$ mínimo entre las parejas de B;

$B \leftarrow B \setminus \{\{g_i, g_j\}\}$;

$M(g_i, g_j) \leftarrow$**true**;

if *No (existe* $p \in G$ *tal que:* $\mathbf{lm}(p) \,|mcm(\mathbf{lm}(g_i), \mathbf{lm}(g_j))$ *y*

$M(g_i, p) = M(p, g_j) =$ **true***)* **then**

$h \leftarrow S(g_i, g_j)$;

$r \leftarrow$ **rem**$(h; G)$;

if $r \neq 0$ **then**

foreach $g \in G$ **do**

agregar a M una entrada para $\{r, g\}$;

if g *y* r *son disjuntos* **then**
  $M(r, g) \leftarrow$**true**;

else

$B \leftarrow B \cup \{\{r, g\}\}$;

$M(r, g) \leftarrow$**false**;

end

end

$G \leftarrow G \cup \{r\}$

end

end

end

end

Demostración.

- **Finitud:** Al igual que en el algoritmo de Buchberger 6.1, el algoritmo finaliza cuando $B = \emptyset$. En ambas versiones, por cada iteración del ciclo **while** en el que el que el residuo r del S-polinomio $S(g_i, g_j)$ no es cero, se agregan a B las parejas de la forma $\{r, g\}$ para $g \in G$; con la salvedad de las parejas disjuntas, pues el teorema 6.6 asegura que se reducen a cero sin necesidad de calcular sus residuos. De esta cuenta, el algoritmo 6.2 debería tener menos repeticiones del ciclo **while** que el algoritmo 6.1, y dado que este último es finito, el primero también debe serlo.

- **Correctitud:**

  Sea G' el estado del conjunto G al finalizar el algoritmo. En la inicialización $G \leftarrow \mathbf{RED}(F)$, por lo que el algoritmo 5.1 asegura que $\langle G \rangle = \langle F \rangle$. Luego, G' contiene los elementos que G tenía en la inicialización junto con los residuos r que se hayan agregado durante la iteración del ciclo **while**. Como cada residuo r es una combinación lineal de los elementos de G, $\langle G \rangle$ es un invariante del ciclo **while**, por lo que $\langle F \rangle = \langle G \rangle = \langle G' \rangle$ durante toda la ejecución del algoritmo. Para probar ahora que G' es una base de Gröbner, se utiliza nuevamente el criterio de Buchberger (teorema 6.3) de forma similar a como se hizo en el algoritmo 6.1. Para ello, note que, al igual que en el algoritmo 6.1, todas las parejas $\{g_i, g_j\} \subseteq G'$ forman parte de B en algún momento de la ejecución del algoritmo, excepto por aquellas parejas de polinomios disjuntos que son marcadas como *tratadas* en M directamente. En dichos casos, el teorema 6.6 asegura que $S(g_i, g_j) \xrightarrow{\{g_i, g_j\}} 0$, por lo que la propiedad 4.29 garantiza que $S(g_i, g_j) \xrightarrow{G'} 0$.

  Notése ahora que si $M(g_i, g_j) = \mathbf{true}$ entonces $S(g_i, g_j) \xrightarrow{G'} 0$. Luego de la inicialización, las únicas entradas de M que son **true** son aquellas que corresponden a parejas de polinomios disjuntos, por lo que, como se acaba de ver, $S(g_i, g_j) \xrightarrow{G'} 0$ y se cumple la propiedad. Procediendo ahora inductivamente, supóngase que luego de t iteraciones del ciclo **while** todas las entradas **true** de M correspondan a S-polinomios que se reducen a cero respecto a G'. Sea $\{g_i, g_j\}$ la pareja que se selecciona durante la iteración $t + 1$ del ciclo **while**, entonces se hace $M(g_i, g_j) \leftarrow \mathbf{true}$. Supóngase que existe $p \in G$ tal que: $\mathbf{lm}(p) \,|mcm(\mathbf{lm}(g_i), \mathbf{lm}(g_j))$ y $M(g_i, p) = M(p, g_j) = \mathbf{true}$. Por la hipótesis de inducción, $S(g_i, p) \xrightarrow{G'} 0$ y $S(p, g_j) \xrightarrow{G'} 0$ lo que, por la definición 4.28, significa que existen m_1 y m_2 representaciones para $S(g_i, p)$ y $S(g_j, p)$ respecto a G respectivamente, donde

  $$m_i \preceq \mathbf{lm}(S(g_i, p)) \preceq mcm(\mathbf{lm}(g_i), \mathbf{lm}(p))$$

  por lo que p cumple con las condiciones del Tercer Criterio de Buchberger (teorema 6.8) y se tendrá que existe una representación respecto a G' para $S(g_i, g_j)$, lo que implica que $S(g_i, g_j) \xrightarrow{G'} 0$.

Si en cambio no existe dicho p, se ingresa al condicional **if** y se ejecuta el ciclo **foreach**, en el cual se agregan entradas **true** a la matriz M sólo si r, g son disjuntos, con lo que se sigue cumpliendo la propiedad. Así, durante toda la ejecución del algoritmo, si $M(g_i, g_j) =$ **true** entonces $S(g_i, g_j) \xrightarrow{G'} 0$. Finalmente, es evidente notar que cada entrada **false** de la matriz M corresponde a una pareja $\{g_i, g_j\} \in B$. Como el algoritmo finaliza cuando $B = \emptyset$ y cada vez que se elimina una pareja de B se hace $M(g_i, g_j) =$**true**, quiere decir que al final del algoritmo por cada pareja de elementos de G' hay una entrada en la matriz M y dicha entrada es **true**. De aquí se deduce que $S(g_i, g_j) \xrightarrow{G'} 0$ para todo $g_i, g_j \in G'$, con $g_i \neq g_j$ y G' debe ser una base de Gröbner de $\langle G' \rangle = \langle F \rangle$. ■

***Ejemplo* 6.4**

Considere el mismo ideal estudiado en el ejemplo 6.3: $I = \langle F \rangle = \langle f_1, f_2 \rangle = \langle x^3 - 2xy, x^2y - 2y^2 + x \rangle$. La Tabla 6.3 muestra el estado de las variables h, r, G, y B conforme se itera sobre el ciclo **while** del algoritmo 6.2. Como el algoritmo 6.2 selecciona la pareja con el menor mínimo común múltiplo de sus monomios principales, en la tabla se incluyen estos valores para hacer más sencillo el seguimiento del algoritmo.

Nótese que en este caso el conjunto F contiene polinomios reducidos, pues son mónicos y ninguno de sus monomios principales divide a algún término del otro polinomio. De esta cuenta, $G \leftarrow$ **RED**(F) resulta ser el mismo conjunto, y únicamente se realiza una llamada al algoritmo de reducción 4.1 dentro de **RED**. Por lo tanto, el conjunto G en el paso 0 es el mismo F.

Así, el algoritmo 6.2 devuelve la base de Gröbner

$$G = \{x^3 - 2xy, x^2y - 2y^2 + x, -x^2, -2y^2 + x, -2xy\}$$

que prácticamente es la misma base obtenida en el ejemplo 6.3 utilizando el algoritmo 6.1, con la única diferencia de que el polinomio $-2y^2 + x$ es la mitad del que se obtuvo en el ejemplo anterior. Como se mencionó al principio de la sección, los pasos más tardados del algoritmo de Buchberger son las reducciones, por lo que el algoritmo de Buchberger mejorado resulta mejor en este ejemplo, realizando únicamente 6 reducciones (contando la realizada dentro del algoritmo **RED** en la inicialización) comparado con las 9 que realiza el algoritmo de Buchberger original. Más aún, concentrándose en las reducciones a cero, el nuevo algoritmo realizó únicamente tres reducciones a cero (de nuevo sumando la realizada en la inicialización) contra las siete realizadas por el algoritmo original (Tabla 6.2). Al igual que se ve en este pequeño ejemplo, las mejoras realizadas al algoritmo son importantes en los ideales que resultan de las aplicaciones, como aseguran Becker (1993) y Cox *et al.* (2007).

Tabla 6.2: Comparación algoritmos 6.1 y 6.2 para el ideal $I = \langle x^3 - 2xy, x^2y - 2y^2 + x \rangle$

	Algoritmo de Buchberger (algoritmo 6.1)	Algoritmo de Buchberger mejorado (algoritmo 6.2)
# Iteraciones	10	8
# Reducciones	9	6
# Reducciones a cero	7	3

6.4. Complejidad del algoritmo de Buchcberger

Esta sección pretende mostrar algunos resultados conocidos sobre la complejidad del algoritmo de Buchberger, así como incluir bibliografía para el lector interesado. No se pretende dar una exposición ni completa ni detallada del tema, por lo que se asume que el lector está familiarizado con los conceptos de la *complejidad* de un algoritmo/problema, problemas P, problemas NP y problemas EXPSPACE. Para una exposición completa de estos temas se recomienda revisar Dasgupta, Papadimitriou, y Vazirani (2008) y Arora (2009). El lector que no desee explorar estos temas puede saltarse esta sección sin ningún inconveniente.

Cuando se habla de la complejidad de un algoritmo, se habla de estimar cuántas veces el algoritmo realizará una tarea específica (por lo general la más costosa del mismo) en función del tamaño de su entrada. Al ejecutar el algoritmo de Buchberger, resulta evidente que el paso más costoso es la reducción de cada S-polinomio respecto a G utilizando el algoritmo 4.1. A su vez, el "tamaño" de la entrada del algoritmo, el conjunto $F \subseteq \mathbb{K}[x_1, \cdots, x_m]$, se puede interpretar tanto como $n = |F|$ el número de polinomios de F ó d el máximo grado de los polinomios de F. Aunque, por lo general, los estudios de la complejidad del algoritmo de Buchberger pretenden estimar el número de reducciones en el algoritmo como función de d, *i.e.* $O(f(d))$. Otros autores consideran que el número de reducciones será proporcional al grado de los polinomios que en algún momento se agreguen a G, por lo que más bien intentan estimar cotas para este (Cox *et al.*, 2007).

La complejidad del algoritmo de Buchberger sigue siendo un área activa de investigación y aún no existen resultados definitivos en esta dirección (Cox *et al.*, 2007). Aunque la versión del algoritmo 6.2 está lejos de ser la mejor versión actualmente conocida del algoritmo de Buchberger, su complejidad no es distinta a la de las mejores versiones modernas, pues sigue siendo sencillo generar ejemplos de ideales para los cuales el cálculo de una base de Gröbner toma mucho tiempo o consume una gran cantidad de memoria. Para una recopilación de resultados en esta dirección veáse Mayr (1997).

En realidad, como menciona Bardet (2005), la complejidad del algoritmo de Buchberger no puede ser pequeña, pues resuelve el problema de calcular bases de Gröbner, que a su vez se puede utilizar para resolver muchos otros problemas de complejidad alta conocida. Por ejemplo, como se verá en el próximo capítulo, las bases de Gröbner se pueden utilizar para resolver sistemas de ecuaciones polinomiales, las cuales a su vez resuelven problemas como *el problema de la mochila* y *el*

Tabla 6.3: Ejemplo algoritmo de Buchberger mejorado (algoritmo 6.2)

Paso	$\{g_1, g_2\}$	h	r	G	B $(\{g_1, g_2\}, mcm(\mathbf{lm}(g_1), \mathbf{lm}(g_2))$	**Comentario**
0				$f_1 = x^3 - 2xy,$ $f_2 = x^2y - 2y^2 + x$	$(\{f_1, f_2\}, x^3y)$	Inicialización
1	$\{f_1, f_2\}$	$S(f_1, f_2) = -x^2$	$-x^2$	$f_1 = x^3 - 2xy,$ $f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2$	$(\{f_1, f_3\}, x^3),$ $(\{f_2, f_3\}, x^2)$	$r \neq 0$
2	$\{f_2, f_3\}$	$S(f_2, f_3) = -2y^2 + x$	$-2y^2 + x$	$f_1 = x^3 - 2xy,$ $f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2,$ $f_4 = -2y^2 + x$	$(\{f_1, f_3\}, x^3),$ $(\{f_2, f_4\}, x^2y^2)$	$r \neq 0$, son disjuntos: $\{f_1, f_4\}, \{f_3, f_4\}$
3	$\{f_1, f_3\}$	$S(f_1, f_3) = -2xy$	$-2xy$	$f_1 = x^3 - 2xy,$ $f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2,$ $f_4 = -2y^2 + x,$ $f_5 = -2xy$	$(\{f_2, f_4\}, x^2y^2),$ $(\{f_1, f_5\}, x^3y),$ $(\{f_2, f_5\}, x^2y),$ $(\{f_3, f_5\}, x^2y),$ $(\{f_4, f_5\}, xy^2)$	$r \neq 0$
4	$\{f_4, f_5\}$	$S(f_4, f_5) = -\frac{1}{2}x^2$	0	$f_1 = x^3 - 2xy,$ $f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2,$ $f_4 = -2y^2 + x,$ $f_5 = -2xy$	$(\{f_2, f_4\}, x^2y^2),$ $(\{f_1, f_5\}, x^3y),$ $(\{f_2, f_5\}, x^2y),$ $(\{f_3, f_5\}, x^2y)$	$r = 0$
5	$\{f_3, f_5\}$	$S(f_3, f_5) = 0$	0	$f_1 = x^3 - 2xy,$ $f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2,$ $f_4 = -2y^2 + x,$ $f_5 = -2xy$	$(\{f_2, f_4\}, x^2y^2),$ $(\{f_1, f_5\}, x^3y),$ $(\{f_2, f_5\}, x^2y)$	$r = 0$
6	$\{f_2, f_5\}$			$f_1 = x^3 - 2xy,$ $f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2,$ $f_4 = -2y^2 + x,$ $f_5 = -2xy$	$(\{f_2, f_4\}, x^2y^2),$ $(\{f_1, f_5\}, x^3y)$	$\mathbf{lm}(f_3) \mid mcm(\mathbf{lm}(f_2), \mathbf{lm}(f_5))$ y $\{f_2, f_3\}$ y $\{f_3, f_5\}$ ya fueron tratados
7	$\{f_2, f_4\}$			$f_1 = x^3 - 2xy,$ $f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2,$ $f_4 = -2y^2 + x,$ $f_5 = -2xy$	$(\{f_1, f_5\}, x^3y)$	$\mathbf{lm}(f_3) \mid mcm(\mathbf{lm}(f_2), \mathbf{lm}(f_4))$ y $\{f_2, f_3\}$ y $\{f_3, f_4\}$ ya fueron tratados
8	$\{f_1, f_5\}$			$f_1 = x^3 - 2xy,$ $f_2 = x^2y - 2y^2 + x$ $f_3 = -x^2,$ $f_4 = -2y^2 + x,$ $f_5 = -2xy$		$\mathbf{lm}(f_3) \mid mcm(\mathbf{lm}(f_1), \mathbf{lm}(f_5))$ y $\{f_1, f_3\}$ y $\{f_3, f_5\}$ ya fueron tratados

problema de satisfacibilidad booleana (SAT), ambos conocidos problemas NP-completos (Bardet, 2005). Los citados problemas se pueden traducir en sistemas de ecuaciones como se muestra a continuación.

Problema de la mochila. Dados $n+1$ enteros naturales $(b_1, \cdots, b_n, c)$, el problema de resolver el siguiente sistema sobredeterminado

$$\sum_{i=1}^{n} x_i b_i = c, \quad x_i(1-x_i) = 0, \ i = 1, \cdots n$$

es conocido como el problema de la mochila 0-1 (*0-1 Knapsack problem*), que fue demostrado ser NP-completo por Karp en 1972.

Problema de satisfacibilidad booleana (3-SAT). Dadas n variables booleanas X_i y un número finito cláusulas booleanas, cada cláusula con tres literales de la forma:

$$Y_j \vee Y_k \vee Y_\ell, \quad Y_j, Y_k, Y_\ell \in \{X_1, \cdots, X_n, \neg X_1, \cdots, \neg X_n\}$$

se debe determinar si existe una asignación de las variables booleanas X_i tal que todas las cláusulas sean verdaderas simultáneamente. Este problema se puede convertir en un sistema de ecuaciones polinomiales al asignarle variables x_i a las variables booleanas X_i que cumplan con $x_i(1-x_i) = 0$ y traduciendo los operadores lógicos como

$$\neg X \longleftrightarrow 1 - x_i$$

$$X \vee Y = \textbf{true} \longleftrightarrow x + y - xy - 1 = 0$$

La demostración de que el problema 3-SAT es NP-completo fue dada por Cook en 1971.

El problema de calcular bases de Gröbner resulta ser mucho peor que NP-completo, pues resuelve también el problema de *pertenencia de ideales* que es un problema EXPSPACE, es decir que es del orden $2^{2^{O(d)}}$ Bardet (2005). Como menciona Cox *et al.* (2007), al ejecutar el algoritmo de Buchberger con polinomios de grado no mayor a d, se pueden llegar a calcular polinomios de grado proporcional a 2^{2^d}. Incluso utilizando el orden GrevLex (que generalmente encuentra las bases de Gröbner más pequeñas, véase Stoutemyer, 2012), es sencillo encontrar polinomios que producen polinomios de grado gigantesco durante la ejecución del algoritmo de Buchberger. Por ejemplo, Cox *et al.* (2007) muestra que si

$$F = \left\{x^{n+1} - yz^{n-1}w, xy^{n-1} - z^n, x^n z - y^n w\right\}$$

con el orden GrevLex con $x \succ y \succ z \succ w$, entonces su base de Gröbner reducida contiene al polinomio

$$z^{n^2+1} - y^{n^2} w.$$

A pesar de que la complejidad de los *peores casos* no es esperanzadora, como menciona Cox *et al.*

(2007), los ideales necesarios para resolver problemas *promedio* de aplicación práctica (incluyendo problemas geométricos) resultan en tiempos de ejecución y tamaño de memoria mucho más manejables para el algoritmo de Buchberger. Por esta razón el método sigue siendo viable para la resolución de problemas en la vida real.

6.5. Pertenencia a ideales y forma normal en $\mathbb{K}[x_1, \cdots, x_n]$

Finalmente, luego de estudiar bases de Gröbner y el algoritmo de Buchberger para generar tales bases para cualquier ideal en $\mathbb{K}[x_1, \cdots, x_n]$, es posible dar la solución a los dos problemas que se plantearon al inicio del Capítulo 4, generalizando lo que se hizo en el Capítulo 3 para polinomios de una sola variable.

Problema 3. Dados un ideal $I = \langle p_1, p_2, \cdots, p_m \rangle \subseteq \mathbb{K}[x_1, \cdots, x_n]$ y un polinomio no cero $p \in \mathbb{K}[x_1, \cdots x_n]$. Determinar si p es un elemento de I o no.

Solución.

1. Elegir un orden de monomios $\prec$ cualquiera para $\mathcal{M}^n \subseteq \mathbb{K}[x_1, \cdots, x_n]$ (según la definición 4.7).

2. Calcular $G' \subseteq I$ una base de Gröbner para I utilizando el algoritmo de Buchberger mejorado (algoritmo 6.2) o cualquier otra implementación del mismo.

3. Opcionalmente, si en el paso anterior no se obtuvo una base de Gröbner reducida, calcular la base de Gröbner reducida G a partir de G' utilizando el algoritmo 5.1.

4. Determinar $r = \mathbf{rem}(p; G)$ utilizando el algoritmo de reducción (algoritmo 4.1).

5. Finalmente, p es elemento de I si y sólo si $r = 0$.

Problema 4. Dado un ideal $I = \langle p_1, p_2, \cdots, p_m \rangle \subseteq \mathbb{K}[x_1, \cdots, x_n]$. Caracterizar el anillo cociente $\mathbb{K}[x_1, \cdots, x_n]/I$, *i.e.* determinar la forma de sus elementos y determinar la forma de la suma y el producto.

Solución.

1. Elegir un orden de monomios $\prec$ cualquiera para $\mathcal{M}^n \subseteq \mathbb{K}[x_1, \cdots, x_n]$ (según la definición 4.7).

2. Calcular $G' \subseteq I$ una base de Gröbner para I utilizando el algoritmo de Buchberger mejorado (algoritmo 6.2) o cualquier otra implementación del mismo.

3. Opcionalmente, si en el paso anterior no se obtuvo una base de Gröbner reducida, calcular la base de Gröbner reducida G a partir de G' utilizando el algoritmo 5.1.

4. Entonces, para $p+I, q+I \in \mathbb{K}[x_1, \cdots, x_n]/I$, por el teorema 5.12 se tiene:

   - **Forma Normal:** $p+I$ y $q+I$ se pueden representar de forma única como r_1+I y r_2+I respectivamente, donde $r_1 = \mathbf{rem}(p; G)$, $r_2 = \mathbf{rem}(q; G)$ utilizando el algoritmo 4.1.
   - **Suma:** $(p+I)+(q+I) = (r_1+I)+(r_2+I) = (r_1+r_2)+I$, sumando r_1+r_2 por términos como se hace normalmente. Como r_1 y r_2 están reducidos respecto a G, ninguno de sus términos pertenece a $\mathbf{lt}(G)$, por lo que al sumarlos, ningún término de r_1+r_2 pertenece tampoco a $\mathbf{lt}(G)$ significando que ya está reducido respecto a G, por lo que ya es la forma normal buscada.
   - **Producto:** $(p+I)(q+I) = (r_1+I)(r_2+I) = (r_1r_2+I)$, como no necesariamente r_1r_2 está ya reducido respecto a G, por lo que es necesario calcular su forma normal $r_3 = \mathbf{rem}(r_1r_2; G)$, por lo que el producto es r_3+I.

7 TEMAS DE GEOMETRÍA ALGEBRAICA

En este capítulo se pretende mostrar algunos conceptos básicos de geometría algebraica, especialmente aquellos para los cuales resultan importantes las bases de Gröbner, y que serán de utilidad en el próximo capítulo para determinar una metodología para demostrar teoremas de geometría euclideana automáticamente. Para una introducción más completa de los temas de geometría algebraica se recomienda revisar el clásico texto de Kunz (1985).

Antes de comenzar el estudio de los temas de este capítulo, vale la pena discutir dónde está la intersección de la geometría y el álgebra, y cuál es el campo de estudio de la *geometría algebraica*.

Como menciona Stillwell (2010) en su libro sobre la historia de las matemáticas, la geometría y el álgebra han estado conectadas desde sus inicios en la civilización Griega, aunque bajo un enfoque muy distinto. Por lo general, las ecuaciones algebraicas eran vistas como *propiedades* de las curvas geométricas, y se utilizaba la geometría para hacer deducciones sobre las ecuaciones. Esta estrategia para trabajar con las ecuaciones no es de extrañar, pues era más formal trabajar la geometría que axiomatizó y trabajó Euclides, a tratar de hacer deducciones a partir de ecuaciones que se escribían con palabras y que fácilmente podían ocupar páginas completas. Un nuevo enfoque en la unión de la geometría y el álgebra no fue posible hasta el siglo XV, cuando el lenguaje y la notación de las ecuaciones hubo evolucionado lo suficiente para que ahora fuera la geometría la que se beneficiara de las ecuaciones. Esta unión se hizo posible con la *geometría analítica* que Descartes y Fermat desarrollaron hacia la década de 1630, luego de darse cuenta que los problemas geométricos pueden ser traducidos a álgebra mediante coordenadas. La geometría analítica permitió entonces resolver muchos de los problemas geométricos de forma rutinaria a través de la manipulación algebraica (Stillwell, 2010).

En la misma línea, la geometría algebraica identifica a las curvas en el plano con ecuaciones polinomiales que deben satisfacer sus coordenadas, pero aplica técnicas más abstractas de álgebra conmutativa a ideales de estos polinomios. En las palabras de Dummit (2004),

$\ll$ *el objeto de estudio de la geometría algebraica comienza donde la resolución de ecuaciones termina* $\gg$

es decir, que es más importante comprender las propiedades intrínsecas del conjunto de soluciones de un sistema de ecuaciones como un todo, a encontrar una solución numérica. La geometría algebraica se encarga de estudiar los ceros de ecuaciones polinomiales con las técnicas del álgebra conmutativa pero con el lenguaje y los problemas de la geometría (Dummit, 2004). La aplicación

del álgebra conmutativa para generalizar la geometría algebraica es de inicios de siglo XX, y los resultados más importantes en esta dirección se deben a Hilbert, siendo estos el teorema de la base (corolario 5.4) y teorema de los ceros (*Nullstellensatz*, teorema 7.20), mismos que se estudian en el presente trabajo, siendo más asequibles gracias a la simpleza de sus demostraciones utilizando bases de Gröbner.

7.1. Variedades algebraicas afines

Uno de los objetos de estudio de la geometría algebraica son las *variedades algebraicas afines* (o simplemente *variedades afines*), las cuales llevan la idea de identificar curvas con polinomios que satisfacen sus coordenadas a un nivel algebraico más abstracto, específicamente ideales de polinomios.

Para hablar de "curvas cuyas coordenadas *satisfacen* un polinomio", es necesario definir formalmente el concepto de *función polinomial*, similar a como se hizo en la definición 3.21, pero ahora para polinomios multivariados. Se seguirá la misma convención de notación que en el Capítulo 3. Para hablar del valor que toma la función del polinomio $p \in \mathbb{K}[x_1, \cdots, x_n]$ en un punto $\vec{a} = (a_1, \cdots, a_n) \in \mathbb{K}^n$ se denotará por $p(\vec{a}) = p(a_1, \cdots, a_n)$, mientras que para hablar del polinomio en sí se denotará simplemente por p, o como $p(\boldsymbol{x}) = p(x_1, \cdots, x_n)$. También será posible *valuar* p en puntos de la forma $(a_1, \cdots, a_k, x_{k+1}, \cdots, x_n)$ que tengan algunas coordenadas fijas y otras variables, esto se interpretará como el polinomio $p(a_1, \cdots, a_k, x_{k+1}, \cdots, x_n) \in \mathbb{K}[x_{k+1}, \cdots, x_n]$ resultante de sustituir las variables fijadas con los valores a_i.

Definición 7.1. *Dado un polinomio $p \in \mathbb{K}[x_1, \cdots, x_n]$ con $p = \sum_{i=0}^{k} a_i \boldsymbol{x^i}$, este induce una función $p : \mathbb{K}^n \to \mathbb{K}$ tal que $\vec{a} \mapsto p(a)$, donde para cada valor $\vec{a} = (a_1, \cdots, a_n) \in \mathbb{K}^n$ se tiene*

$$p(\vec{a}) = p(a_1, \cdots, a_n) = \sum_{i=0}^{k} a_i \vec{a}^{\boldsymbol{i}} = \sum_{i=0}^{k} a_i (a_1^{i_1} \cdot a_2^{i_2} \cdots a_n^{i_n}).$$

Definición 7.2. *Dado un polinomio $p \in \mathbb{K}[x_1, \cdots, x_n] \setminus \{0\}$, se dice que $\vec{a_0} \in \mathbb{K}^n$ es **una raíz o cero de** p, si $p(\vec{a_0}) = 0$.*

Proposición 7.3. Dados dos polinomios $p, q \in \mathbb{K}[x_1, \cdots, x_n]$, si $p = q$ (como polinomios) entonces $p(\vec{a}) = q(\vec{a})$ $\forall \vec{a} \in \mathbb{K}^n$ (como funciones).

Demostración. La demostración es análoga a la del lema 3.23 para polinomios de una sola variable. ∎

Definición 7.4. *Sea $X \subseteq \mathbb{K}^n$ un subconjunto cualquiera. Se define el **ideal de** X como el conjunto*

$$\mathcal{I}(X) = \{p \in \mathbb{K}[x_1, \cdots, x_n] \mid p(\vec{c}) = 0 \ \forall \vec{c} \in X\}.$$

Proposición 7.5. Sea $X \subseteq \mathbb{K}^n$ un subconjunto cualquiera. El ideal de X, $\mathcal{I}(X)$ es un ideal de $\mathbb{K}[x_1, \cdots, x_n]$.

Demostración. Por la propiedad 2.10:

I) Sean $p, q \in \mathcal{I}(X)$. Entonces $p(\vec{c}) = q(\vec{c}) = 0 \ \forall \ \vec{c} \in X$. Luego, $(p-q)(\vec{c}) = p(\vec{c}) - q(\vec{c}) = 0 - 0 = 0, \ \forall \vec{c} \in X$ por lo que $p - q \in \mathcal{I}(X)$.

II) Sean $p \in \mathcal{I}(X)$ y $r \in \mathbb{K}[x_1, \cdots, x_n]$ un polinomio cualquiera. Entonces $(pr)(\vec{c}) = (rp)(\vec{c}) = r(\vec{c})p(\vec{c}) = r(\vec{c}) \cdot 0 = 0, \ \forall \ \vec{c} \in X$, por lo que $pr, rp \in \mathcal{I}(X)$.

Por lo tanto, $\mathcal{I}(X)$ es un ideal. ■

Observación. Particularmente, $\mathcal{I}(\emptyset) = \mathbb{K}[x_1, \cdots, x_n]$ pues no existe ningún polinomio que no se haga cero para algún valor de $\emptyset$.

***Ejemplo* 7.1**

En el caso particular en el que $\mathbb{K}$ es un campo de característica cero, $\mathcal{I}(\mathbb{K}^n) = \{0\}$. Para ver por qué esto es cierto en el caso particular en el que $\mathbb{K} = \mathbb{C}$, supóngase por el absurdo que existe $p \in \mathcal{I}(\mathbb{C}^n)$ con $p \neq 0$. Entonces $p(\vec{c}) = 0 \ \forall \vec{c} \in \mathbb{C}^n$, particularmente $p(c, 0, \cdots, 0) = 0 \ \forall \ c \in \mathbb{C}$. *i.e.* c es raíz de $p(x_1, 0, \cdots, 0) \in \mathbb{C}[x_1] \ \forall c \in \mathbb{C}$, pero esto se contradice con el corolario 3.4 al teorema fundamental del álgebra que asegura que todo polinomio no cero en $\mathbb{C}[x_1]$ debe tener un número finito de raíces. Por lo tanto $\mathcal{I}(\mathbb{C}^n) = \{0\}$.

Definición 7.6. *Sea $S \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un subconjunto de polinomios. Se define el conjunto $\mathcal{V}(S)$ como*

$$\mathcal{V}(S) = \{(c_1, \cdots, c_n) \in \mathbb{K}^n | \ p(c_1, \cdots, c_n) = 0, \ \forall \ p \in S\}.$$

Definición 7.7 (Variedad algebraica afín)**.** *Sea $X \subseteq \mathbb{K}^n$ un subconjunto cualquiera. Se dice que X es una* ***variedad algebraica afín*** *(o simplemente* ***variedad algebraica****) si existe algún conjunto $S \subseteq \mathbb{K}[x_1, \cdots, x_n]$ tal que $X = \mathcal{V}(S)$.*

***Ejemplo* 7.2** (Hipersuperficies)

Las hipersuperficies son definidas por un sólo polinomio $p \in \mathbb{K}[x_1, \cdots, x_n]$ no constante, así $X = \mathcal{V}(\{p\})$ contiene a todos los puntos que hacen cero al polinomio. Cuando $n = 2$ a las hipersuperficies se les llama *curvas*, y cuando $n = 3$, *superficies*. La Figura 7.1 presenta algunas curvas y superficies.

Lema 7.8.

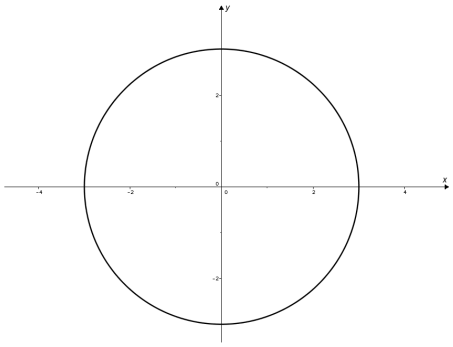
I) Sean $X, Y \subseteq \mathbb{K}^n$ tales que $X \subseteq Y$, entonces $\mathcal{I}(Y) \subseteq \mathcal{I}(X)$.
Por otro lado, si $S, T \subseteq \mathbb{K}[x_1, \cdots, x_n]$ tales que $S \subseteq T$, entonces $\mathcal{V}(T) \subseteq \mathcal{V}(S)$.

II) Sea X una variedad algebraica de $\mathbb{K}^n$. Entonces existe un ideal $I \subseteq \mathbb{K}[x_1, \cdots, x_n]$ tal que $X = \mathcal{V}(I)$.

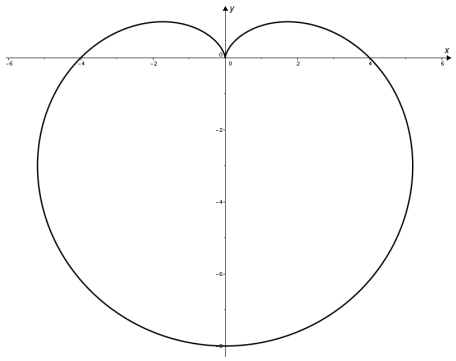
III) Si $X \subseteq \mathbb{K}^n$ es un subconjunto cualquiera, entonces $X \subseteq \mathcal{V}(\mathcal{I}(X))$, con igualdad si y sólo si X es una variedad algebraica.

Figura 7.1: Ejemplos de variedades algebraicas: curvas y superficies

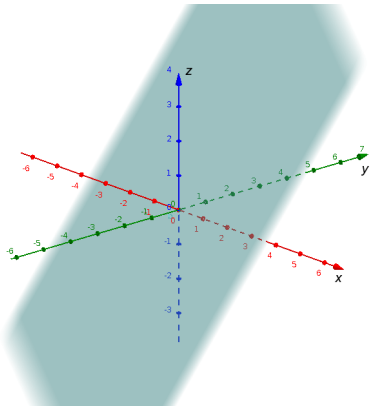
(a) Curva $x^2 + y^2 - 9 = 0$



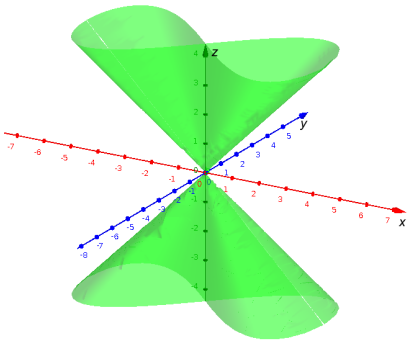
(b) Curva $(x^2 + y^2 + 4y)^2 - 16(x^2 + y^2) = 0$



(c) Superficie $2x + 3y - 2z = 0$



(d) Superficie $x^4 + (y^2 - x^2)z^2 = 0$



IV) Si $I \subseteq \mathbb{K}[x_1, \cdots, x_n]$ es un ideal, entonces $I \subseteq \mathcal{I}(\mathcal{V}(I))$.

V) $\mathcal{V}(\langle 0 \rangle) = \mathbb{K}^n$ y $\mathcal{V}(\mathbb{K}[x_1, \cdots, x_n]) = \emptyset$.

VI) Sean I y J ideales de $\mathbb{K}[x_1, \cdots, x_n]$, entonces:

$$\mathcal{V}(I \cap J) = \mathcal{V}(I) \cup \mathcal{V}(J),$$
$$\mathcal{V}(IJ) = \mathcal{V}(I) \cup \mathcal{V}(J),$$
$$\mathcal{V}(I + J) = \mathcal{V}(I) \cap \mathcal{V}(J).$$

Demostración.

I) Sea $p \in \mathcal{I}(Y)$. Entonces $p(\vec{c}) = 0, \ \forall \ \vec{c} \in Y$. Particularmente, si $\vec{c} \in X \subseteq Y$, $p(\vec{c}) = 0$. Por lo que $p(\vec{c}) = 0, \ \forall \ \vec{c} \in X$ y $p \in \mathcal{I}(X)$. Por lo tanto $\mathcal{I}(Y) \subseteq \mathcal{I}(X)$.

Sea ahora $\vec{c} \in \mathcal{V}(T)$. Entonces $p(\vec{c}) = 0, \ \forall \ p \in T$. Particularmente, si $p \in S \in T$, $p(\vec{c}) = 0$. Por lo que $p(\vec{c}) = 0, \ \forall \ p \in S$ y $\vec{c} \in \mathcal{V}(S)$. Por lo tanto $\mathcal{V}(T) \subseteq \mathcal{V}(S)$.

II) Como X es una variedad algebraica, existe un conjunto $S \subseteq \mathbb{K}[x_1, \cdots, x_n]$ tal que $X = \mathcal{V}(S)$. Se probará que para el ideal $\langle S \rangle$, $X = \mathcal{V}(\langle S \rangle)$. Como $S \subseteq \langle S \rangle$, la propiedad anterior asegura

que $\mathcal{V}(\langle S\rangle) \subseteq \mathcal{V}(S) = X$ por lo que basta con probar que $X = \mathcal{V}(S) \subseteq \mathcal{V}(\langle S\rangle)$.

Sea $\vec{c} \in \mathcal{V}(S)$. Entonces $p(\vec{c}) = 0,\ \forall\ p \in S$. Para un elemento cualquiera $q \in \langle S\rangle$, se tiene que $q = \sum_{i=0}^{k} a_i p_i$ con los $a_i \in \mathbb{K}[x_1, \cdots, x_n]$ y los $p_i \in S$. Entonces

$$q(\vec{c}) = \sum_{i=0}^{k} a_i(\vec{c})\cancelto{0}{p_i(\vec{c})} = 0,$$

por lo que $q(\vec{c}) = 0\ \forall\ q \in \langle S\rangle$ y entonces $\vec{c} \in \mathcal{V}(\langle S\rangle)$. De donde $X \subseteq \mathcal{V}(\langle S\rangle)$, lo que demuestra la afirmación.

III) Si $\vec{c} \in X$ entonces $p(\vec{c}) = 0,\ \forall\ p \in \mathcal{I}(X)$ por definición, pero entonces $\vec{c} \in \mathcal{V}(\mathcal{I}(X))$. Si $X = \mathcal{V}(\mathcal{I}(X))$ entonces evidentemente X es una variedad algebraica. Por otro lado, si X es una variedad algebraica, por la propiedad anterior debe existir un ideal I tal que $X = \mathcal{V}(I)$. Esto quiere decir que X consiste de todos los puntos que hacen cero a todos los polinomios de I, pero el conjunto $\mathcal{I}(X)$ contiene a todos los polinomios que se hacen cero con todos los puntos de X, por lo que $I \subseteq \mathcal{I}(X)$. Entonces $X \subseteq \mathcal{V}(\mathcal{I}(X)) \subseteq \mathcal{V}(I) = X$, por lo tanto $X = \mathcal{V}(\mathcal{I}(X))$.

IV) Si $p \in I$ entonces $p(\vec{c}) = 0,\ \forall\ \vec{c} \in \mathcal{V}(I)$, por definición de $\mathcal{V}(I)$. Pero entonces $p \in \mathcal{I}(\mathcal{V}(I))$ por definición. Por lo tanto, $I \subseteq \mathcal{I}(\mathcal{V}(I))$.

V) Claramente $0(\vec{c}) = 0$ para todo $\vec{c} \in \mathbb{K}^n$, por lo que $\mathcal{V}(\langle 0\rangle) = \mathbb{K}^n$. Por otro lado, si existiera $\vec{c} = (c_1, \cdots, c_n) \in \mathcal{V}(\mathbb{K}[x_1, \cdots, x_n])$, siempre es posible construir el polinomio $p \in \mathbb{K}[x_1, \cdots, x_n]$ con $p = (x_1 - c_1)(x_2 - c_2)\cdots(x_n - c_n) - 1$ tal que $p(\vec{c}) = -1 \neq 0$ lo que es una contradicción, por lo que $\mathcal{V}(\mathbb{K}[x_1, \cdots, x_n]) = \emptyset$.

VI)

- $I \cap J \subseteq I \Rightarrow \mathcal{V}(I) \subseteq \mathcal{V}(I \cap J)$. $I \cap J \subseteq J \Rightarrow \mathcal{V}(J) \subseteq \mathcal{V}(I \cap J)$. De allí que, $\mathcal{V}(I) \cup \mathcal{V}(J) \subseteq \mathcal{V}(I \cap J)$. Por otro lado, si $\vec{c} \notin \mathcal{V}(I) \cup \mathcal{V}(J) \Rightarrow \vec{c} \notin \mathcal{V}(I)$ ni $\vec{c} \in \mathcal{V}(J) \Rightarrow$ existen $p_1 \in I$ y $p_2 \in J$ tales que $p_1(\vec{c}) \neq 0$ y $p_2(\vec{c}) \neq 0$. Como I y J son ideales, entonces $p_1 p_2 \in I$ y $p_1 p_2 \in J$ por lo que $p_1 p_2 \in I \cap J$, pero $p_1 p_2(\vec{c}) = p_1(\vec{c}) p_2(\vec{c}) \neq 0$ entonces $\vec{c} \notin \mathcal{V}(I \cap J)$. Por lo que, su contrapositiva $\mathcal{V}(I \cap J) \subseteq \mathcal{V}(I) \cup \mathcal{V}(J)$.
- $IJ = \left\{\sum p_i q_i | p_i \in I,\ q_i \in J\right\} \subseteq \left\{\sum p_i q_i | p_i \in I,\ q_i \in \mathbb{K}[x_1, \cdots, x_n]\right\} = \langle I\rangle = I$. Análogamente, $IJ \subseteq J$, por lo que $\mathcal{V}(I) \subseteq \mathcal{V}(IJ)$ y $\mathcal{V}(J) \subseteq \mathcal{V}(IJ)$. Por lo que $\mathcal{V}(I) \cup \mathcal{V}(J) \subseteq \mathcal{V}(IJ)$. Por otro lado, si $\vec{c} \notin \mathcal{V}(I) \cup \mathcal{V}(J) \Rightarrow \vec{c} \notin \mathcal{V}(I)$ ni $\vec{c} \in \mathcal{V}(J) \Rightarrow$ existen $p_1 \in I$ y $p_2 \in J$ tales que $p_1(\vec{c}) \neq 0$ y $p_2(\vec{c}) \neq 0$. Como $p_1 p_2 \in IJ$, pero $p_1 p_2(\vec{c}) = p_1(\vec{c}) p_2(\vec{c}) \neq 0$ entonces $\vec{c} \notin \mathcal{V}(IJ)$. Por lo que, por contrapuesta, $\mathcal{V}(IJ) \subseteq \mathcal{V}(I) \cup \mathcal{V}(J)$.
- $\vec{c} \in \mathcal{V}(I) \cap \mathcal{V}(J) \Leftrightarrow$ para cualquier $p \in I, p(\vec{c}) = 0$ y cualquier $p \in J, p(\vec{c}) = 0 \Leftrightarrow$ para $p \in I$ ó $p \in J$, $p(\vec{c}) = 0 \Leftrightarrow$ para cualquier $p \in I \cup J$, $p(\vec{c}) = 0 \Leftrightarrow \vec{c} \in \mathcal{V}(I \cup J)$. Por lo tanto, $\mathcal{V}(I \cup J) = \mathcal{V}(I) \cap \mathcal{V}(J)$.

■

Observación. Es importante notar que en la propiedad IV, la inclusión puede en efecto ser estricta. Como ejemplo considérese el ideal $I = \langle x^2 + 1 \rangle$ en $\mathbb{R}[x]$. Entonces $\mathcal{V}(I) = \emptyset$, pues $x^2 + 1 = 0$ no tiene ceros reales, y entonces $\mathcal{I}(\mathcal{V}(I)) = \mathcal{I}(\emptyset) = \mathbb{R}[x]$, por lo que la inclusión es evidentemente estricta $I = \langle x^2 + 1 \rangle \subsetneq \mathbb{R}[x] = \mathcal{I}(\mathcal{V}(I))$.

Definición 7.9. *Una variedad algebraica X es **irreducible** si para cualesquiera variedades algebraicas X_1 y X_2 tales que $X = X_1 \cup X_2$ se tiene que $X = X_1$ ó $X = X_2$.*

Proposición 7.10. Sea $X \subseteq \mathbb{K}^n$ una variedad algebraica. X es irreducible si, y sólo si, $\mathcal{I}(X) \subseteq \mathbb{K}[x_1, \cdots, x_n]$ es un ideal primo.

Demostración. Supóngase que $\mathcal{I}(X)$ no es un ideal primo. Entonces existen $p, q \in \mathbb{K}[x_1, \cdots, x_n]$ tales que $pq \in \mathcal{I}(X)$ pero $p, q \notin \mathcal{I}(X)$. Sean $X_1 = \mathcal{V}(\mathcal{I}(X) + \langle p \rangle)$ y $X_2 = \mathcal{V}(\mathcal{I}(X) + \langle q \rangle)$. Claramente X_1, X_2 son variedades algebraicas, como $\mathcal{I}(X) \subsetneq \mathcal{I}(X) + \langle p \rangle$ y $\mathcal{I}(X) \subsetneq \mathcal{I}(X) + \langle q \rangle$ entonces $X = \mathcal{V}(\mathcal{I}(X)) \supsetneq \mathcal{V}(\mathcal{I}(X) + \langle p \rangle) = X_1$ y $X = \mathcal{V}(\mathcal{I}(X)) \supsetneq \mathcal{V}(\mathcal{I}(X) + \langle q \rangle) = X_2$, por lo que $X \supset X_1 \cup X_2$. Por otro lado, sea $\vec{c} \in X \Rightarrow pq(\vec{c}) = 0 \Rightarrow \ p(\vec{c}) = 0$ ó $q(\vec{c}) = 0$, entonces $\vec{c} \in \mathcal{V}(\mathcal{I}(X) + \langle p \rangle)$ ó $\vec{c} \in \mathcal{V}(\mathcal{I}(X) + \langle q \rangle)$, por lo que $\vec{c} \in X_1 \cup X_2$. Por lo tanto, $X = X_1 \cup X_2$ y no es irreducible.

Supóngase ahora que X no es irreducible. Entonces existen variedades algebraicas X_1 y X_2 tales que $X = X_1 \cup X_2$, con $X \neq X_1, X_2$. Si $\mathcal{I}(X) = \mathcal{I}(X_1) \Rightarrow \ X = \mathcal{V}(\mathcal{I}(X)) = \mathcal{V}(\mathcal{I}(X_1)) = X_1$ $(\rightarrow\leftarrow)$ por lo que $\mathcal{I}(X) \neq \mathcal{I}(X_1)$, análogamente $\mathcal{I}(X) \neq \mathcal{I}(X_2) \Rightarrow \mathcal{I}(X_i) \setminus I(X) \neq \emptyset$ para $i = 1, 2$. Sean $p \in \mathcal{I}(X_1) \setminus \mathcal{I}(X)$ y $q \in \mathcal{I}(X_2) \setminus \mathcal{I}(X)$, como $pq \in \mathcal{I}(X_1)$ y $pq \in \mathcal{I}(X_2) \Rightarrow \ pq \in \mathcal{I}(X_1) \cap \mathcal{I}(X_2)$ entonces $pq(\vec{c}) = 0, \ \forall \ \vec{c} \in X_1$ y $pq(\vec{c}) = 0, \ \forall \ \vec{c} \in X_2$, por lo que $pq(\vec{c}) = 0, \ \forall \ \vec{c} \in X_1 \cup X_2 = X \Rightarrow$ $pq \in \mathcal{I}(X)$, de donde $\mathcal{I}(X)$ no es primo. ■

Teorema 7.11. Sea $X_1 \supset X_2 \supset \cdots$ una secuencia estrictamente descendiente de variedades algebraicas de $\mathbb{K}^n$. Entonces la secuencia debe ser finita.

Demostración. Por el lema anterior, se puede construir la cadena de ideales $\mathcal{I}(X_i)$ en $\mathbb{K}[x_1, \cdots, x_n]$, siendo esta una cadena ascendente:

$$\mathcal{I}(X_1) \subseteq \mathcal{I}(X_2) \subseteq \cdots.$$

Note que esta cadena es más bien estrictamente ascendente, pues si $\mathcal{I}(X_k) = \mathcal{I}(X_{k+1})$ para algún k, se tendría que $X_k = \mathcal{V}(\mathcal{I}(X_k)) = \mathcal{V}(\mathcal{I}(X_{k+1}) = X_{k+1}$ lo que no puede ser, pues la cadena de X_i's es estrictamente descendiente. Entonces, por la propiedad de cadenas ascendentes debe ser finita. Por lo tanto, la secuencia estrictamente descendiente de variedades algebraicas, debe ser también finita. ■

Teorema 7.12. Sea $X \subseteq \mathbb{K}^n$ una variedad algebraica. Entonces X es una unión finita de variedades algebraicas irreducibles. Más aún, si esta descomposición se hace irredundante (*i.e.* $X_i \not\subseteq X_j$ para $i \neq j$) entonces los componentes irreducibles son únicos.

Demostración. Si X es irreducible, se tiene lo deseado. En otro caso, $X = Y_1 \cup Z_2$, $X \neq Y, Z$. Si Y_1 y Z_2 fueran irreducibles, se obtiene lo deseado, supóngase entonces sin pérdida de la generalidad que Y_1 no es irreducible, entonces $Y_1 = Y_2 \cup Z_2$, $Y_1 \neq Y_2, Z_2$. Siguiendo con este proceso, X se debe poder escribir como la unión de un número finito de variedades irreducibles, siendo esto falso únicamente cuando el proceso descrito no pueda terminarse, *i.e.* siempre que Y_i no es irreducible, $Y_i = Y_{i+1} \cup Z_{i+1}$, sin pérdida de la generalidad, con Y_{i+1} no irreducible. Pero para que este fuera el caso, se tendría una secuencia $Y_1 \supsetneq Y_2 \supsetneq \cdots$ infinita de variedades algebraicas, lo cual sería una contradicción por el teorema anterior.

Para probar la unicidad, supóngase que $X = X_1 \cup \cdots \cup X_m$ y $X = Y_1 \cup \cdots \cup Y_k$ son dos descomposiciones irredundantes en irreducibles. Entonces $X_1 = X_1 \cap X = X_1 \cap (Y_1 \cup \cdots \cup Y_k) = (X_1 \cap Y_1) \cup \cdots \cup (X_1 \cap Y_k)$. Como X_1 es irreducible, se debe tener $X_1 = X_1 \cap Y_j$ para algún j entonces $Y_j \subseteq X_1$. De la misma forma, debe existir algún i tal que $X_i \subseteq Y_j \subseteq X_1$, como la descomposición es irredundante entonces $i = 1$ y $X_1 = Y_1$. Continuando con el proceso se obtiene que $m = k$ y que las descomposiciones tienen los mismos componentes. ∎

7.2. *Nullstellensatz*-Teorema de los ceros de Hilbert

El teorema de los ceros de Hilbert (o *Nullstellensatz* en alemán) es uno de los teoremas fundamentales de la geometría algebraica. A manera de generalización del teorema fundamental del álgebra, para campos $\mathbb{K}$ algebraicamente cerrados, asegura que todo sistema de ecuaciones polinomiales debe tener algún cero en $\mathbb{K}^n$. Más aún, en su forma fuerte, permite hacer una identificación biunívoca entre las variedades algebraicas de $\mathbb{K}^n$ y un subconjunto de los ideales de $\mathbb{K}[x_1, \cdots, x_n]$. Por simplicidad, en la discusión que sigue se tomará $\mathbb{K} = \mathbb{C}$, aunque los argumentos son fácilmente extrapolables a cualquier campo algebraicamente cerrado.

Dado un sistema de ecuaciones polinomiales

$$\begin{aligned} p_1(x_1, \cdots, x_n) &= 0 \\ p_2(x_1, \cdots, x_n) &= 0 \\ &\vdots \\ p_k(x_1, \cdots, x_n) &= 0 \end{aligned}$$

se sabe que cualquier solución del sistema debe ser un elemento de la variedad algebraica $\mathcal{V}(\{p_1, \cdots, p_k\})$ que coincide con $\mathcal{V}(\langle p_1, \cdots, p_k \rangle)$ gracias a la propiedad II del lema 7.8. La versión débil del *Nullstellensatz* de Hilbert asegura que para cualquier ideal I propio de $\mathbb{C}[x_1, \cdots, x_n]$, esta variedad es no

vacía, teniendo así que el sistema de ecuaciones debe tener solución. La demostración al siguiente teorema sigue la idea de Lars Svensson mostrada por Fröberg (1997).

Lema 7.13. Sea $I \subseteq \mathbb{C}[x_1, \cdots, x_n]$ un ideal tal que $I \cap \mathbb{C}[x_1] = \emptyset$. Entonces existe $\alpha \in \mathbb{C}$ tal que el ideal $I_\alpha \neq \mathbb{C}[x_2, \cdots, x_n]$, donde $I_\alpha = \{q(\alpha, x_2, \cdots, x_n) | q \in I\}$.

Demostración. Considérese el anillo $\mathbb{C}(x_1)[x_2, \cdots, x_n]$ de polinomios en variables $x_2, \cdots, x_n$ y con coeficientes funciones racionales en x_1. Para cada polinomio $p \in \mathbb{C}(x_1)[x_2, \cdots, x_n]$ puede construirse el polinomio $q \in \mathbb{C}[x_1]$ mínimo común múltiplo de los denominadores de los coeficientes en $\mathbb{C}(x_1)$ de p, por lo que el polinomio $p' = qp \in \mathbb{C}[x_1, \cdots, x_n]$. Sea $J \subseteq \mathbb{C}(x_1)[x_2, \cdots, x_n]$ el ideal generado por I en este anillo, *i.e.*

$$J = \left\{ \sum c_i p_i \middle| \ c_i \in \mathbb{C}(x_1)[x_2, \cdots, x_n], p_i \in I \right\}.$$

Sea $G = \{g_1, \cdots, g_m\}$ una base de Gröbner finita para J. Por el criterio de Buchberger, para todo $i \neq j$ debe darse $\mathbf{rem}(S(g_i, g_j); G) = 0$, por lo que existen polinomios $h_{ij}^r \in \mathbb{C}(x_1)[x_2, \cdots, x_n]$ tales que

$$S(g_i, g_j) = h_{ij}^1 g_1 + h_{ij}^2 g_2 + \cdots + h_{ij}^m g_m, \ \forall \ i \neq j \tag{7.1}$$

Dado que hay un número finito de parejas $1 \leq i, j \leq k$, debe haber un número finito de coeficientes $c_i(x_1) = \frac{a_i(x_1)}{b_i(x_1)} \in \mathbb{C}(x_1)$ involucrado en las expresiones 7.1, por lo que se puede construir el polinomio

$$P(x_1) = \prod_i a_i(x_1) \prod_i b_i(x_1) \in \mathbb{C}[x_1] \setminus \{0\}$$

el cual, por el teorema fundamental del álgebra, debe tener un número finito de ceros (o ninguno si fuera constante). Por lo tanto existe $\alpha \in \mathbb{C}$ que no es raíz de P, por lo que no anula ni indefine ninguno de los coeficientes c_i, $i = 1, 2, \cdots, g$.

Sea ahora $Q = \prod_i \frac{b_i}{\mathbf{lc}(b_i)}$ el producto de todos los denominadores de los coeficientes $c_i(x_1) \in \mathbb{C}(x_1)$, normalizado para que sea mónico. Sean $g_i' = Qg_i \in \mathbb{C}[x_1, \cdots, x_n]$, entonces:

$$\begin{aligned}
S(g_i', g_j') = S(Qg_i, Qg_j) &= \frac{mcm(\mathbf{lm}(Qg_i), \mathbf{lm}(Qg_j))}{\mathbf{lt}(Qg_i)} \cdot Qg_i - \frac{mcm(\mathbf{lm}(Qg_i), \mathbf{lm}(Qg_j))}{\mathbf{lt}(Qg_j)} \cdot Qg_j = \\
&= \frac{\cancel{\mathbf{lm}(Q)} mcm(\mathbf{lm}(g_i), \mathbf{lm}(g_j))}{\cancel{\mathbf{lt}(Q)}\mathbf{lt}(g_i)} \cdot Qg_i - \frac{\cancel{\mathbf{lm}(Q)} mcm(\mathbf{lm}(g_i), \mathbf{lm}(g_j))}{\cancel{\mathbf{lt}(Q)}\mathbf{lt}(g_j)} \cdot Qg_j = \\
&= QS(g_i, g_j) = Q\left(h_{ij}^1 g_1 + h_{ij}^2 g_2 + \cdots + h_{ij}^m g_m\right) = \\
&= h_{ij}^1 (Qg_1) + h_{ij}^2 (Qg_2) + \cdots + h_{ij}^m (Qg_m) = \\
&= h_{ij}^1 g_1' + h_{ij}^2 g_2' + \cdots + h_{ij}^m g_m'
\end{aligned} \tag{7.2}$$

Si ahora se valúa la expresión 7.2 en $x_1 = \alpha$ se obtiene la siguiente igualdad en $\mathbb{C}[x_2, \cdots, x_n]$,

denotando $h_{ij}^{'r} = h_{ij}^{r}|_{x_1=\alpha}$:

$$S(g_i'(\alpha, x_2, \cdots, x_n), g_j'(\alpha, x_2, \cdots, x_n)) = \\ = h_{ij}^{'1} g_1'(\alpha, x_2, \cdots, x_n) + h_{ij}^{'2} g_2'(\alpha, x_2, \cdots, x_n) + \cdots + h_{ij}^{'m} g_m'(\alpha, x_2, \cdots, x_n) \tag{7.3}$$

para todo $i \neq j$, por lo que por el criterio de Buchberger, $G_\alpha = \{g_1'(\alpha, x_2, \cdots, x_n), \cdots, g_m'(\alpha, x_2, \cdots, x_n)\}$ es una base de Gröbner para $\langle G_\alpha \rangle$.

Supóngase ahora por el absurdo que $\langle G_\alpha \rangle = \mathbb{C}[x_2, \cdots, x_n]$, entonces $1 \in \langle G_\alpha \rangle$ y se puede escribir como combinación lineal de los $g_i'(\alpha, x_2, \cdots, x_n)$, por lo que para algún i, $\mathbf{lm}(g_i'(\alpha, x_2, \cdots, x_n)) \,|1 \Rightarrow \mathbf{lm}(g_i'(\alpha, x_2, \cdots, x_n)) = 1 \Rightarrow g_i'(\alpha, x_2, \cdots, x_n) \in \mathbb{C} \Rightarrow g_i'(x_1, x_2, \cdots, x_n) \in \mathbb{C}[x_1]$. Pero entonces $g_i'(x_1, x_2, \cdots, x_n) = Q(x_1) g_i(x_2, \cdots, x_n) \in J$ por lo que existen polinomios $h_j \in \mathbb{C}(x_1)[x_2, \cdots, x_n]$ tales que

$$g_i'(x_1, x_2, \cdots, x_n) = Q(x_1) g_i(x_2, \cdots, x_n) = \sum_{j=1}^{r} h_j p_j, \quad p_j \in I, \tag{7.4}$$

donde nuevamente hay involucrada una cantidad finita de coeficientes $c_i \in \mathbb{C}(x_1)$. Haciendo $R(x_1) \in \mathbb{C}[x_1]$ el mínimo común múltiplo de todos los denominadores involucrados y multiplicando la expresión 7.4 por $R(x_1)$, se obtiene

$$R(x_1) g_i'(x_1, x_2, \cdots, x_n) = R(x_1) Q(x_1) g_i(x_2, \cdots, x_n) = \sum_{j=1}^{r} (R(x_1) h_j) p_j$$

Donde todos los polinomios pertenecen a $\mathbb{C}[x_1, \cdots, x_n]$, probando que $R(x_1) g_i'(x_1, \cdots, x_n) \in I \cap \mathbb{C}[x_1] = \emptyset$ $(\rightarrow\leftarrow)$. Por lo tanto $\langle G_\alpha \rangle \neq \langle 1 \rangle$.

Finalmente, observe que si $q(\alpha, x_2, \cdots, x_n) \in I_\alpha$ entonces

$$q(\alpha, x_2, \cdots, x_n) = \sum_{i=1}^{m} q_i|_{x_1=\alpha} \cdot g_i|_{x_1=\alpha} = \sum_{i=1}^{m} \frac{1}{Q(\alpha)} q_i|_{x_1=\alpha} \cdot g_i'(\alpha, x_2, \cdots, x_n) \in \langle G_\alpha \rangle ,$$

por lo que $I_\alpha \subseteq \langle G_\alpha \rangle \neq \mathbb{C}[x_2, \cdots, x_n]$, probando así el enunciado. ∎

Lema 7.14. Sea $I \subseteq \mathbb{C}[x_1, \cdots, x_n]$ un ideal tal que $I \cap \mathbb{C}[x_i] = \langle p_i \rangle$ y $\mathbf{deg}(p_i) > 0$ para todo $1 \leq i \leq n$. Entonces existe $\vec{\alpha} = (\alpha_1, \cdots, \alpha_n) \in \mathbb{C}$ tal que $I \subseteq \langle x_1 - \alpha_1, x_2 - \alpha_2, \cdots, x_n - \alpha_n \rangle$.

Demostración. Para cada $1 \leq i \leq n$, $\langle p_i \rangle \subseteq \mathbb{C}[x_i]$, por lo que por el teorema fundamental del álgebra existe $\alpha_i \in \mathbb{C}$ tal que $p_i(\alpha_i) = 0$. Entonces, por el teorema del factor, $p_i(x_i) = (x_i - \alpha_i) p_i'(x_i)$ con $\mathbf{deg}(p_i'(x_1)) < \mathbf{deg}(p_i(x_i))$. Nótese que $1 \notin I + \langle x_i - \alpha_i \rangle$ pues de lo contrario existirían polinomios $q_1 \in I$ y $q_2 \in \mathbb{C}[x_1, \cdots, x_n]$ tales que $1 = q_1 + q_2(x_i - \alpha_i) \Rightarrow p_i' = q_1 p_i' + q_2 p_i'(x_i - \alpha_i) = q_1 p_i' + q_2 p_i \in I \Rightarrow p_i' \in I \cap \mathbb{C}[x_i] = \langle p_i \rangle \Rightarrow p_i | p_i' \Rightarrow \mathbf{deg}(p_i) \leq \mathbf{deg}(p_i')$ llegando así a una contradicción.

Sea $J = (I + \langle x_1 - \alpha_1 \rangle) + \cdots + (I + \langle x_n - \alpha_n \rangle) = I + \langle x_1 - \alpha_1, \cdots, x_n - \alpha_n \rangle$. Nótese que $1 \notin J$, pues de lo contrario $1 \in I + \langle x_i - \alpha_i \rangle$ para algún i $(\rightarrow\leftarrow)$. Por lo que $J \neq \mathbb{C}[x_1, \cdots, x_n]$.

Se prueba ahora que $\langle x_1 - \alpha_1, \cdots, x_n - \alpha_n \rangle$ es un ideal maximal. Supóngase por el absur-

do que existe un ideal I' tal que $\langle x_1-\alpha_1,\cdots,x_n-\alpha_n\rangle \subsetneq I' \subsetneq \mathbb{C}[x_1,\cdots,x_n]$. Sea $p \in I' \neq \langle x_1-\alpha_1,\cdots,x_n-\alpha_n\rangle$ y sea G una base de Gröbner para $\langle x_1-\alpha_1,\cdots,x_n-\alpha_n\rangle$, entonces $\mathbf{rem}(p;G) = r \neq 0$, con r reducido respecto a G, *i.e.* que ningún término de r es divisible por los monomios principales de los elementos de G, lo que significa que ningún término de r es elemento de $\mathbf{lt}(G) = \mathbf{lt}(x_1-\alpha_1,\cdots,x_n-\alpha_n) = \langle \mathcal{M}^n \setminus \{1\}\rangle$, por lo tanto $r \in \mathbb{C}$ es constante y es posible encontrar una combinación lineal de los $x_i-\alpha_i$ y p que dé como resultado 1, o sea que $1 \in I' = \mathbb{C}[x_1,\cdots,x_n]$ ($\rightarrow\leftarrow$). Por lo tanto, $\langle x_1-\alpha_1,\cdots,x_n-\alpha_n\rangle$ debe ser maximal.

Finalmente, como $I+\langle x_1-\alpha_1,\cdots,x_n-\alpha_n\rangle \subsetneq \mathbb{C}[x_1,\cdots,x_n]$ entonces $I+\langle x_1-\alpha_1,\cdots,x_n-\alpha_n\rangle = \langle x_1-\alpha_1,\cdots,x_n-\alpha_n\rangle \Rightarrow I \subseteq \langle x_1-\alpha_1,\cdots,x_n-\alpha_n\rangle$.

■

Teorema 7.15 (*Nullstellensatz* débil)**.** Sea I un ideal de $\mathbb{C}[x_1,\cdots,x_n]$ con $I \neq \mathbb{C}[x_1,\cdots,x_n]$. Entonces $\mathcal{V}(I) \neq \emptyset$.

Demostración. Por inducción sobre n, el número de variables.

1. Para $n = 1$. I es un ideal propio de $\mathbb{C}[x_1]$, que es un anillo de ideales principales por el teorema 3.13. Entonces existe $p \in \mathbb{C}[x_1]$ tal que $I = \langle p\rangle$. Más aún, como $\langle p\rangle = I \neq \mathbb{C}[x_1]$, p no puede ser constante, por lo que el teorema fundamental del álgebra asegura que existe $c \in \mathbb{C}$ tal que $p(c) = 0 \Rightarrow c \in \mathcal{V}(I) \neq \emptyset$.

2. Sea ahora $n > 1$ y supóngase que el teorema es verdadero para todo ideal $I \subseteq \mathbb{C}[x_1,\cdots,x_m]$ para $m < n$. Considérense dos casos:

Caso 1: $I \cap \mathbb{C}[x_i] = \langle 0\rangle$ para algún i. Sin pérdida de la generalidad supóngase que $i = 1$. Por el lema 7.13 existe $\alpha \in \mathbb{C}$ tal que $I_\alpha \subsetneq \mathbb{C}[x_2,\cdots,x_n]$. Por la hipótesis de inducción, $\mathcal{V}(I_\alpha) \neq \emptyset$ entonces existe $\vec{c} = (c_2,\cdots,c_n) \in \mathbb{C}^{n-1}$ tal que $p(\alpha,c_2,\cdots,c_n) = 0$ para todo $p(\alpha,x_2,\cdots,x_n) \in I_\alpha$. Que es equivalente a que $p(\alpha,c_2,\cdots,c_n) = 0$ para todo $p \in I$, por lo que $(\alpha,c_2,\cdots,x_n) \in \mathcal{V}(I) \neq \emptyset$.

Caso 2: $I \cap \mathbb{C}[x_i] \neq \langle 0\rangle$ para todo i. Claramente $1 \notin I \cap \mathbb{C}[x_i]$, pues esto implicaría que $1 \in I$ haciendo $I = \mathbb{C}[x_1,\cdots,x_n]$, entonces $I \cap \mathbb{C}[x_i]$ es un ideal propio de $\mathbb{C}[x_i]$. Como $\mathbb{C}[x_i]$ es un anillo de ideales principales, debe existir $p_i \in \mathbb{C}[x_i]$ con $\mathbf{deg}(p_i) > 0$ tal que $I \cap \mathbb{C}[x_i] = \langle p_i\rangle$. Por el lema 7.14, existe $\vec{\alpha} = (\alpha_1,\cdots,\alpha_n) \in \mathbb{C}^n$ tal que $I \subseteq \langle x_1-\alpha_1,\cdots,x_n-\alpha_n\rangle$. Entonces $\vec{\alpha} \in \mathcal{V}(\langle x_1-\alpha_1,\cdots,x_n-\alpha_n\rangle) \subseteq \mathcal{V}(I) \neq \emptyset$.

■

Definición 7.16. *Dados un anillo conmutativo con unidad R y un ideal $I \subseteq R$, se define el* ***ideal radical de*** *I como*

$$\sqrt{I} = \left\{ r \in R \mid r^N \in I \text{ para algún } N \in \mathbb{Z}^+ \right\}$$

Definición 7.17. *Dados un anillo conmutativo con unidad R y un ideal $I \subseteq R$, se dice que I* ***es un ideal radical*** *si y sólo si $I = \sqrt{I}$.*

Proposición 7.18. Dados un anillo conmutativo con unidad R y un ideal $I \subseteq R$, el ideal radical de I, $\sqrt{I}$ es un ideal de R que contiene a I. Más aún, $\sqrt{I}$ es un ideal radical.

Demostración. Evidentemente, para todo $r \in I$, $r \in \sqrt{I}$, pues $r^1 \in I \Rightarrow I \subseteq \sqrt{I}$. Ahora se prueba que $\sqrt{I}$ es un ideal utilizando la caracterización 2.10:

I) Sean $a, b \in \sqrt{I}$ entonces existen $N, M \in \mathbb{Z}^+$ tales que $a^N, b^M \in I$. Por el teorema del binomio, se tiene que
$$(a-b)^{N+M-1} = \sum_{i=0}^{N+M-1} \binom{N+M-1}{i} (-1)^i a^{N+M-1-i} b^i$$
donde es claro que si $0 \leq i \leq M-1$ la potencia de a es mayor o igual que N, y para $M \leq i \leq N+M-1$ la potencia de b es mayor o igual que M, por lo que $a^{N+M-1-i}b^i \in \sqrt{I}$ para todo $0 \leq i \leq N+M-1$ y $(a-b)^{N+M-1} \in I$. Entonces $a - b \in \sqrt{I}$.

II) Sean $a \in \sqrt{I}$ y $r \in R$. Entonces existe $N \in \mathbb{Z}^+$ tal que $a^N \in I \Rightarrow (ar)^N = (ra)^N = r^N a^N \in I$ $\Rightarrow ar, ra \in \sqrt{I}$

Por lo que $\sqrt{I}$ es un ideal. Ahora bien, se sabe que $\sqrt{I} \subseteq \sqrt{\sqrt{I}}$. Sea $a \in \sqrt{\sqrt{I}}$, entonces existe $N \in \mathbb{Z}^+$ tal que $a^N \in \sqrt{I}$, entonces existe $M \in \mathbb{Z}^+$ tal que $(a^N)^M = a^{NM} \in I$, por lo que $a \in \sqrt{I}$ y se tiene que $\sqrt{I} = \sqrt{\sqrt{I}} \Rightarrow \sqrt{I}$ es un ideal radical. ∎

Lema 7.19. Sea $I \subseteq \mathbb{K}[x_1, \cdots, x_n]$ un ideal. Entonces $\mathcal{V}(I) = \mathcal{V}(\sqrt{I})$.

Demostración. Como $I \subseteq \sqrt{I}$ entonces $\mathcal{V}(\sqrt{I}) \subseteq \mathcal{V}(I)$. Por otro lado, sea $\vec{c} \in \mathcal{V}(I)$ entonces $p(\vec{c}) = 0 \ \forall \ p \in I$. Luego, para cualquier $p \in \sqrt{I}$ existe $N \in \mathbb{Z}^+$ tal que $p^N \in I$, por lo que $p^N(\vec{c}) = 0 \ \Leftrightarrow \ p(\vec{c}) = 0$ de donde $\vec{c} \in \mathcal{V}(\sqrt{I})$. Por lo que $\mathcal{V}(I) \subseteq \mathcal{V}(\sqrt{I})$. ∎

Teorema 7.20 (*Nullstellensatz* fuerte)**.** Para cada ideal $I \subseteq \mathbb{C}[x_1, \cdots, x_n]$ se cumple $\mathcal{I}(\mathcal{V}(I)) = \sqrt{I}$.

Demostración. Por los lemas 7.8 y 7.19, se sabe que $\sqrt{I} \subseteq \mathcal{I}(\mathcal{V}(\sqrt{I})) = \mathcal{I}(\mathcal{V}(I))$. Sea $p \in \mathcal{I}(\mathcal{V}(I))$. Considérese el ideal $J = I + \langle py - 1 \rangle$ en el anillo $\mathbb{C}[x_1, \cdots, x_n, y]$ con una nueva variable y. Obsérvese que si $(c_1, \cdots, c_n, c_{n+1}) \in \mathcal{V}(J) \subseteq \mathbb{C}^{n+1} \Rightarrow (c_1, \cdots, c_n) \in \mathcal{V}(I) \subseteq \mathbb{C}^n$. Como $p \in \mathcal{I}(\mathcal{V}(I))$ entonces $p(c_1, \cdots, c_n) = 0$, pero eso implica que $p(c_1, \cdots, c_n) \cdot c_{n+1} - 1 = -1$ $(\rightarrow\leftarrow)$. Por lo que $\mathcal{V}(J) = \emptyset$, y la versión débil del *Nullstellensatz* (teorema 7.15) implica que $J = \mathbb{C}[x_1, \cdots, x_n, y]$. Entonces existen $g_i \in \mathbb{C}[x_1, \cdots, x_n, y]$ tales que

$$1 = g_1 p_1 + \cdots + g_m p_m + g_{m+1}(py - 1) \ (\star) \ \text{ con los } p_i \in I \tag{7.5}$$

Como $\mathbb{C}[x_1, \cdots, x_n, y] \subseteq \mathbb{C}(x_1, \cdots, x_n)[y]$, la expresión 7.5 puede verse como una igualdad de polinomios en el anillo $\mathbb{C}(x_1, \cdots, x_n)[y]$, por lo que la igualdad se conserva al valuar en cualquier $y \in \mathbb{C}(x_1, \cdots, x_n)$. Particularmente hágase $y = \frac{1}{p} \in \mathbb{C}(x_1, \cdots, x_n)$, entonces 7.5 se convierte en:

$$1 = g_1' p_1 + \cdots + g_m' p_m' + g_{m+1}\cancelto{0}{\left(p\frac{1}{p} - 1\right)} \tag{7.6}$$

donde $g_i' = g_i\left(x_1, \cdots, x_n, \frac{1}{p}\right)$. Es claro que en la igualdad 7.6 todos los denominadores presentes son potencias de p, por lo que si N es la potencia más grande de p en un denominador, se multiplica p^N por la igualdad 7.6 y se obtiene:

$$p^N = g_1'' p_1 + \cdots + g_m'' p_m,$$

donde $g_i'' = p^N g_i' \in \mathbb{C}[x_1, \cdots, x_n]$, por lo que $p^N \in I$ y entonces $p \in \sqrt{I}$. Así, $\mathcal{I}(\mathcal{V}(I)) \subseteq \sqrt{I}$. Por lo tanto $\mathcal{I}(\mathcal{V}(I)) = \sqrt{I}$. ∎

El teorema de los ceros de Hilbert en su versión fuerte, permite hacer una identificación biunívoca entre las variedades algebraicas de $\mathbb{C}^n$ y los ideales radicales de $\mathbb{C}[x_1, \cdots, x_n]$, de tal forma que los mapas $\mathcal{I}$ y $\mathcal{V}$ son inversos. Es decir, si $X \subseteq \mathbb{C}^n$ es una variedad algebraica, entonces existe un ideal $I \subseteq \mathbb{C}[x_1, \cdots, x_n]$ tal que $X = \mathcal{V}(I) = \mathcal{V}(\sqrt{I})$, por lo que se identifica la variedad X con el ideal radical $\sqrt{I}$, así:

$$\mathcal{I}(X) = \mathcal{I}(\mathcal{V}(\sqrt{I})) = \sqrt{\sqrt{I}} = \sqrt{I} \quad \text{y}$$
$$\mathcal{V}(\sqrt{I}) = \mathcal{V}(I) = X$$

La demostración del *Nullstellensatz* versión fuerte también incluye otra herramienta útil: un mecanismo para determinar si un polinomio p pertenece o no al ideal $\sqrt{I}$. El siguiente corolario muestra el funcionamiento de este criterio.

Corolario 7.21. Sean $p \in \mathbb{C}[x_1, \cdots, x_n]$ e $I \subseteq \mathbb{C}[x_1, \cdots, x_n]$ un ideal. Entonces $p \in \sqrt{I}$ si y sólo si $I + \langle py - 1\rangle = \mathbb{C}[x_1, \cdots, x_n, y]$.

Demostración. ($\Rightarrow$) Supóngase $p \in \sqrt{I} = \mathcal{I}(\mathcal{V}(I))$. Si existiera $(c_1, \cdots, c_n, c) \in \mathcal{V}(I + \langle py - 1\rangle)$ entonces $(c_1, \cdots, c_n) \in \mathcal{V}(I)$, por lo que $p(c_1, \cdots, c_n) = 0$, pero entonces $p(c_1, \cdots, c_n)y - 1 = -1$ ($\rightarrow\leftarrow$) entonces $\mathcal{V}(I + \langle py - 1\rangle) = \emptyset$. Por lo tanto, el *Nullstellensatz* versión débil asegura que $I + \langle py - 1\rangle = \mathbb{C}[x_1, \cdots, x_n, y]$.

($\Leftarrow$) Supóngase $I + \langle py - 1\rangle = \mathbb{C}[x_1, \cdots, x_n]$. Supóngase por reducción al absurdo que $p \notin \sqrt{I} = \mathcal{I}(\mathcal{V}(I))$, entonces existe $(c_1, \cdots, c_n) \in \mathcal{V}(I) \subseteq \mathbb{C}^n$ tal que $p(c_1, \cdots, c_n) \neq 0$. Sea ahora $c = \frac{1}{p(c_1, \cdots, c_n)} \in \mathbb{C}$, el cuál está bien definido pues el denominador es distinto a cero. Como $(c_1, \cdots, c_n) \in \mathcal{V}(I)$, el vector $(c_1, \cdots, c_n, c)$ anula a todos los polinomios de I, más aún también anula a $py - 1$, pues

$$p(c_1, \cdots, c_n, c) \cdot c - 1 = p(c_1, \cdots, c_n) \cdot \frac{1}{p(c_1, \cdots, c_n)} - 1 = 0,$$

por lo que $(c_1, \cdots, c_n, c) \in \mathcal{V}(I + \langle py - 1\rangle) \neq \emptyset$ y $I + \langle py - 1\rangle \neq \mathbb{C}[x_1, \cdots, x_n, y]$ llegando así a una contradicción. Por lo tanto, $p \in \sqrt{I}$. ∎

Observación. Si $I = \langle p_1, \cdots, p_n\rangle$ es el ideal en cuestión, $J = I + \langle py - 1\rangle = \langle p_1, \cdots, p_n, py - 1\rangle$. Para determinar si $J = \mathbb{C}[x_1, \cdots, x_n, y]$ o no, basta con determinar la base de Gröbner reducida G de J, dándose la igualdad si y sólo si $G = \{1\}$. En otras palabras, $p \in \sqrt{\langle p_1, \cdots, p_n\rangle}$ si y sólo sí la base de Gröbner reducida de $\langle p_1, \cdots, p_n, py - 1\rangle$ es $\{1\}$.

7.3. Solución de sistemas cero-dimensionales

Como se ha mencionado a lo largo del trabajo, una de las aplicaciones más importante de las bases de Gröbner es la resolución de sistemas de ecuaciones. Dado un número finito de ecuaciones polinomiales, $p_i(x_1, \cdots, x_n) = 0$, $i = 1, \cdots, m$, resolver tal sistema equivale a determinar el conjunto $\mathcal{V}(\langle p_1, \cdots, p_m\rangle) \subseteq \mathbb{K}^n$. El teorema de los ceros de Hilbert, asegura que si $\mathbb{K}$ es un campo algebraicamente cerrado y si el ideal $\langle p_1, \cdots, p_m\rangle \neq \langle 1\rangle$ entonces el sistema tiene al menos una solución. En esta sección, se verá cómo encontrar estas soluciones explicitamente, cuando el sistema tiene un número finito de soluciones. Por simplicidad se utilizará $\mathbb{K} = \mathbb{C}$, aunque de nuevo se puede extrapolar las ideas mostradas a cualquier campo algebraicamente cerrado. Para más información sobre este tema, véase (Fröberg, 1997).

Definición 7.22. *Sea $I \subseteq \mathbb{C}[x_1, \cdots, x_n]$ un ideal no cero. Se dice que I es **cero-dimensional** si la dimensión del $\mathbb{C}$-espacio vectorial $\mathbb{C}[x_1, \cdots, x_n]/I$ es finita, i.e. $\dim_{\mathbb{C}} \mathbb{C}[x_1, \cdots, x_n]/I < \infty$. A su vez, se dice que un sistema de ecuaciones $\{p_i = 0 | i = 1, \cdots, m\}$ es **un sistema cero-dimensional** si el ideal $\langle p_1, \cdots, p_m\rangle$ es cero-dimensional.*

Lema 7.23. Sea $I \subseteq \mathbb{C}[x_1, \cdots, x_n]$ un ideal de monomios. Entonces

$$\dim_{\mathbb{C}} \mathbb{C}[x_1, \cdots, x_n]/I < \infty$$

si, y sólo si, existe $m_i \in \mathbb{Z}^+$ tal que $x_i^{m_i} \in I$ para cada $i = 1, \cdots, n$.

Demostración. $\Rightarrow$ Supóngase que $\dim_{\mathbb{C}} \mathbb{C}[x_1, \cdots, x_n]/I < \infty$. Si existiera algún i tal que ninguna potencia de x_i pertenezca a I, entonces $x_i^m \in \mathcal{M}^n \setminus I \Rightarrow x_i \in \mathcal{M}^n \setminus \mathbf{lt}(I)$ (claramente $I = \mathbf{lt}(I)$ para cualquier orden de monomios, pues I es un ideal de monomios). Por el teorema 5.12, $x_i^m + I$ debe entonces estar en la base del espacio vectorial $\mathbb{C}[x_1, \cdots, x_n]/I$ para todo m, por lo que su dimensión no podría ser finita.

$\Leftarrow$ Supóngase que para cada i existe m_i tal que $x_i^{m_i} \in I$ y sea $N = \text{máx}_i(m_i)$. Entonces todo monomio de grado mayor a $n(N-1)$ debe pertenecer a I. Nótese que este es el caso, pues para que un monomio $\boldsymbol{x^\alpha}$ tenga grado mayor a $n(N-1)$, al menos uno de sus exponentes α_i debe ser mayor que N, entonces $\alpha_i > m_i \Rightarrow\ x_i^{m_i} | x_i^{\alpha_i} \Rightarrow\ x_i^{m_i} | \boldsymbol{x^\alpha}\ \Rightarrow\ \boldsymbol{x^\alpha} \in I$. Así, sólo un número finito de monomios no pertenece a $I = \mathbf{lt}(I)$, y por el teorema 5.12, estos constituyen una base finita del espacio vectorial $\mathbb{C}[x_1, \cdots, x_n]/I$. ∎

Lema 7.24. Sea $I \subseteq \mathbb{C}[x_1, \cdots, x_n]$ un ideal no cero. Entonces I es cero-dimensional si, y sólo si, $I \cap \mathbb{C}[x_i] \neq \langle 0 \rangle$ para todo $i = 1, \cdots, n$.

Demostración. $\Rightarrow$ Supóngase que I es cero-dimensional. Tómese i fijo, y sea $\prec$ el orden de monomios Lex con $x_i \prec x_{i+1} \prec \cdots \prec x_n \prec x_1 \prec \cdots \prec x_{i-1}$. En este orden, todas las potencias de x_i son menores que cualquier otro monomio que incluya alguna variable x_j, $j \neq i$. Por el lema anterior, existe m_i tal que $x_i^{m_i} \in \mathbf{lt}(I)$, lo que implica que $x_i^{m_i}$ debe ser el término principal de algún polinomio $p \in I$. Como los demás términos de p son menores que $x_i^{m_i}$ en el orden que se está utilizando, todos deben ser potencias de x_i y evidentemente $p \in I \cap \mathbb{C}[x_i] \neq \langle 0 \rangle$.

$\Leftarrow$ Supóngase ahora que $I \cap \mathbb{C}[x_i] \neq \langle 0 \rangle$ para todo $i = 1, \cdots, n$. Sean entonces $P_i \in \mathbb{C}[x_i] \cap I$ polinomios distintos de cero para cada i y sea V el $\mathbb{C}-$espacio vectorial generado por la base

$$\left\{ \boldsymbol{x^i} |\ \boldsymbol{i} = (i_1, \cdots, i_n) \text{ con } i_j < \mathbf{deg}(P_j) \right\}$$

Sea entonces $\phi : V \to \mathbb{C}[x_1, \cdots, x_n]/I$ el mapa lineal definido por $\phi(\boldsymbol{x^i}) \mapsto [\boldsymbol{x^i}]$, que mapea elementos de V a sus respectivas clases módulo I. Si $[p]$ es una clase lateral cualquiera de $\mathbb{C}[x_1, \cdots, x_n]/I \Rightarrow p \in \mathbb{C}[x_1, \cdots, x_n]$. Sea $r = \mathbf{rem}(p; P_1, \cdots, P_n)$. Como r está reducido respecto a los P_i's, ninguno de sus monomios es divisible dentro de $\mathbf{lm}(P_i) = x_i^{\mathbf{deg}(P_i)}$, lo que significa que el exponente de x_i debe ser menor que $\mathbf{deg}(P_i)$, por lo que claramente $r \in V$. Por otro lado, $p - r \in \langle P_1, \cdots, P_n \rangle \subseteq I \Rightarrow [p] = [r]$. Entonces, existe $r \in V$ tal que $\phi(r) = [r] = [p]$, por lo que ϕ es sobreyectivo. Por lo tanto $\dim_{\mathbb{C}} \mathbb{C}[x_1, \cdots, x_n]/I \leq \dim_{\mathbb{C}} V < \infty$. ∎

El siguiente teorema demuestra que efectivamente la definición dada para ideales de dimensión cero es una caracterización del hecho de que tengan una cantidad finita de soluciones.

Teorema 7.25. Sea $I \subseteq \mathbb{C}[x_1, \cdots, x_n]$ un ideal. I es cero-dimensional si, y sólo si, $\mathcal{V}(I)$ es un conjunto finito.

Demostración. $\Rightarrow$ Supóngase que I es un ideal cero-dimensional. Por el lema anterior, para cada i existe $P_i(x_i) \in I \cap \mathbb{C}[x_i]$. Luego, se tiene que $\mathcal{V}(I) \subseteq \{(c_1, \cdots, c_n) \in \mathbb{C}^n | P_i(c_i) = 0 \text{ para } i = 1, \cdots, n\}$, pues los P_i's deben ser anulados por $\mathcal{V}(I)$. Como cada P_i tiene una cantidad finita de soluciones en $\mathbb{C}$ (por el corolario 3.26), el conjunto que se acaba de construir debe ser finito y $\mathcal{V}(I)$ es finito.

$\Leftarrow$ Supóngase que $\mathcal{V}(I)$ es un conjunto finito. Si $\mathcal{V}(I) = \emptyset$ entonces por la versión débil de *Nullstellensatz* (teorema 7.15) $I = \langle 1 \rangle = \mathbb{C}[x_1, \cdots, x_n] \Rightarrow \dim_{\mathbb{C}} \mathbb{C}[x_1, \cdots, x_n]/I = 0$. Si $\mathcal{V}(I) \neq \emptyset$ entonces $\mathcal{V}(I) = \{\vec{c}_i, \cdots, \vec{c}_N\}$ con

$$\vec{c}_j = (c_{1,j}, c_{2,j}, \cdots, c_{n,j}) \ \text{ con } c_{i,j} \in \mathbb{C}.$$

Sean entonces $H_i = (x_i - a_{i,1})(x_i - a_{i,2}) \cdots (x_i - a_{i,N})$, $i = 1, \cdots, n$. Luego, evidentemente los H_i's se anulan para todos los $\vec{c}_j$, por lo que $H_i \in \mathcal{I}(\mathcal{V}(I))$. Por el *Nullstellensatz* versión fuerte (teorema 7.20), $\mathcal{I}(\mathcal{V}(I)) = \sqrt{I}$, por lo que $H_i \in \sqrt{I} \cap \mathbb{C}[x_i] \Rightarrow H_i^{n_i} \in I \cap \mathbb{C}[x_i] \neq \langle 0 \rangle$, para toda I. Por lo tanto, por el lema anterior, I debe ser un ideal cero-dimensional. ∎

El teorema anterior asegura que si un ideal I es cero-dimensional, entonces existen polinomios únicamente en la variable x_i que pertenzcan a I, para toda i. El siguiente corolario permite de cierta forma extender dicha propiedad a una base de Gröbner de I calculada respecto del orden lexicográfico, dando una *forma triangular* para el ideal. Es esta forma triangular la que permitirá resolver los sistemas de ecuaciones polinomiales.

Corolario 7.26 (Forma Triangular)**.** Sea $I \subseteq \mathbb{C}[x_1, \cdots, x_n]$ un ideal cero-dimensional distinto de cero. Sea $G = \{g_1, \cdots, g_n\}$ una base de Gröbner para I respecto al orden Lex con $x_n \prec x_{n-1} \prec \cdots \prec x_2 \prec x_1$, ordenada tal que $\mathbf{lm}(g_i) \succ \mathbf{lm}(g_{i+1})$. Entonces, para cada $i = 1, \cdots, n$, existe j tal que $\mathbf{lm}(g_j) = x_i^{d_i}$, para algún $d_i > 0$ y $g_j \in \mathbb{C}[x_i, x_{i+1}, \cdots, x_n]$.

Demostración. Como I es un ideal cero-dimensional, el lema 7.24 asegura que para cada i, existe un polinomio P_i en $I \cap \mathbb{C}[x_i]$. Como $\mathbf{lt}(G) = \mathbf{lt}(I)$, debe haber un polinomio $g_j \in G$ tal que $\mathbf{lm}(g_j) \,|\mathbf{lm}(P_i)$, pero como $P_i \in \mathbb{C}[x_i]$, claramente $\mathbf{lm}(g_j) = x_i^{d_i}$ para algún $d_i > 0$. Como $x_i^{d_i}$ es el monomio principal de g_j con el orden Lex, todos los demás monomios deben ser menores que $x_i^{d_i}$, lo que significa que deben incluir únicamente las variables $x_i, x_{i+1}, \cdots, x_n$, por lo que $g_j \in \mathbb{C}[x_i, x_{i+1}, \cdots, x_n]$. ∎

El corolario anterior conduce directamente a un mecanismo para resolver sistemas cero-dimensionales, pues la forma triangular de $g_1, \cdots, g_s$ es particularmente conveniente para encontrar $\mathcal{V}(I)$. El último polinomio de la base de Gröbner debe ser elemento de $\mathbb{C}[x_n]$, *i.e.* $g_s = g_s(x_n)$. Se encuentran las raíces de g_s y se sustituyen una por una en los demás polinomios de G. Luego de sustituir, al menos uno de los polinomios anteriores debe ser elemento de $\mathbb{C}[x_{n-1}]$, por lo que se puede repetir el proceso, continuando así hasta resolver polinomios en $\mathbb{C}[x_1]$. Este método, conocido como *backsolving*, permite ver que la recíproca del corolario también es cierta. Si la base de Gröbner respecto al orden lexicográfico de un ideal I está en forma triangular, se puede obtener $\mathcal{V}(I)$ por *backsolving* y este es entonces un conjunto finito, por lo que el sistema debe ser cero-dimensional. El método se ilustra mejor en el siguiente ejemplo.

***Ejemplo* 7.3**

Se desea encontrar las soluciones $(x, y, z) \in \mathbb{R}^3$ para el sistema de ecuaciones:

$$\begin{aligned} x^2 + y^2 + z^2 &= 4, \\ x^2 + 2y^2 &= 5, \\ xz &= 1. \end{aligned}$$

Para ello considérese el sistema equivalente

$$\begin{aligned} x^2+y^2+z^2-4&=0,\\ x^2+2y^2-5&=0,\\ xz-1&=0. \end{aligned} \tag{7.7}$$

Nótese entonces que resolver el sistema 7.7 corresponde a encontrar $\mathcal{V}(I)\cap\mathbb{R}$ para el ideal

$$I=\left\langle x^2+y^2+z^2-4, x^2+2y^2-5, xz-1\right\rangle \subseteq \mathbb{C}[x,y,z].$$

Se obtiene entonces la base de Gröbner reducida para I, respecto al orden lexicográfico con $z \prec y \prec x$:

$$\begin{aligned} G=\{&g_1 = x+2z^3-3z,\\ &g_2 = y^2-z^2-1,\\ &g_3 = z^4-\tfrac{3}{2}z^2+\tfrac{1}{2}\} \end{aligned}$$

Se resuelve entonces g_3 para z:

$$\begin{aligned} &z^4-\tfrac{3}{2}z^2+\tfrac{1}{2}=0\\ &z^2=\frac{\frac{3}{2}\pm\sqrt{\frac{9}{4}-2}}{2}=\tfrac{3}{4}\pm\tfrac{1}{4}=1,\ \tfrac{1}{2}\\ &z=\pm 1,\ \pm\tfrac{\sqrt{2}}{2} \end{aligned}$$

Por lo que las soluciones del sistema deben ser de la forma $(x,y,\pm1)$, $(x,y,\pm\frac{\sqrt{2}}{2})$. Luego, al sustituir z por cada raíz de g_3, g_2 se convierte en un polinomio únicamente en la variable y, por lo que se encuentran sus raíces.

• Si $z=\pm1$

$g_2(x,y,\pm1)=y^2-1-1=0$

$y=\pm\sqrt{2}$

• Si $z=\pm\frac{\sqrt{2}}{2}$

$g_2(x,y,\pm\sqrt{2}/2)=y^2-\dfrac{1}{2}-1=0$

$y=\pm\dfrac{\sqrt{2}\sqrt{3}}{2}$

Por lo que ahora se tienen ocho posibles combinaciones para los valores de y y z, teniendo soluciones de la forma $\left(x,\pm1,\pm\sqrt{2}\right)$ y $\left(x,\pm\frac{\sqrt{2}}{2},\pm\frac{\sqrt{2}\sqrt{3}}{2}\right)$. Finalmente se sustituyen en el polinomio g_1 que ahora se convierte univariado en x y se resuelve para x:

$$g_1(x,y,z)=x+2z^3-3z=0 \ \Rightarrow\ x=3z-2z^3$$

• Si $z=1$

$x=3\cdot1-2(1)^3=1$

• Si $z=-1$

$x=3\cdot(-1)-2(-1)^3=-1$

• Si $z=\frac{\sqrt{2}}{2}$

$x=3\cdot(\frac{\sqrt{2}}{2})-2(\frac{\sqrt{2}}{2})^3=-1$

• Si $z=-\frac{\sqrt{2}}{2}$

$x=3\cdot(-\frac{\sqrt{2}}{2})-2(-\frac{\sqrt{2}}{2})^3=-1$

Por lo tanto, se obtienen 8 soluciones al sistema de ecuaciones. En este ejemplo en particular,

todas son de números reales, por lo que

$$\mathcal{V}(I)\cap\mathbb{R}=\mathcal{V}(I)=\left\{\left(-\sqrt{2},-\tfrac{\sqrt{2}\sqrt{3}}{2},-\tfrac{\sqrt{2}}{2}\right),\left(-\sqrt{2},\tfrac{\sqrt{2}\sqrt{3}}{2},-\tfrac{\sqrt{2}}{2}\right),\left(\sqrt{2},-\tfrac{\sqrt{2}\sqrt{3}}{2},\tfrac{\sqrt{2}}{2}\right),\left(\sqrt{2},\tfrac{\sqrt{2}\sqrt{3}}{2},\tfrac{\sqrt{2}}{2}\right),\right.$$
$$\left.\left(-1,-\sqrt{2},-1\right),\left(-1,\sqrt{2},-1\right),\left(1,-\sqrt{2},1\right),\left(1,\sqrt{2},1\right)\right\}.$$

8 DEMOSTRACIÓN AUTOMATIZADA DE TEOREMAS GEOMÉTRICOS

En este capítulo se mostrará la aplicación de las bases de Gröbner a la demostración automotizada de teoremas geométricos. Como ilustración, al final del capítulo se presenta su aplicación a los casos no degenerados del teorema del hexagrama de Pascal.

El creciente interés por construir demostradores automáticos de teoremas, que dieran los pasos formales de las demostraciones, o al menos demostradores automatizados, que determinaran la validez de los resultados aunque no dieran los pasos de la demostración, se hizo evidente en la década de los 60's. Según Ferro y Gallo (1988), en 1959 fue desarrollado por H. Gelernter el primer programa capaz de producir demostraciones formales de geometría elemental, en el centro de investigaciones de la IBM en Nueva York. Aunque este programa contaba con una base de datos de axiomas y de teoremas previamente demostrados que utilizaba como un operador de reducción, en muchas ocasiones necesitaba de la guía humana para sustituir razonamientos complejos por otros más sencillos. A pesar de la importancia histórica de este método, muchos de los investigadores comenzaron a concentrarse en los llamados demostradores algebraicos, pues estos superaban varias de las deficiencias del mismo (Ferro y Gallo, 1988).

Entre 1977 y 1978, Wu Wentsün propuso una metodología algebraica para demostrar teoremas de geometría de forma automatizada, basada en un algoritmo de eliminación descubierto por J. F. Ritt en 1950. El método de Wu fue implementado por Chou en Austin, Texas, creando así la mayor colección de teoremas probados mecánicamente hasta la fecha (Ferro y Gallo, 1988). Estimulados por el éxito de Wu, múltiples investigadores se han dedicado a estudiar el uso de bases de Gröbner a la misma problemática (Wu, 1997). Aunque es usual que el método con bases de Gröbner sea computacionalmente más costoso que el algoritmo de Wu, ambas metodologías son escencialmente equivalentes y pueden resolver el mismo tipo de problemas. Por otro lado, son mucho más comunes los paquetes algebraicos con implementaciones del algoritmo de Buchberger que las implementaciones del algoritmo de Wu (Stokes y Bulmer, 2001). Además, las bases de Gröbner se pueden generalizar fácilmente a álgebras no conmutativas y conservan información algebraica que se pierde con el método de Wu (Buchberger, 2013).

Aunque en la actualidad se investigan diferentes técnicas utilizando bases de Gröbner, se presenta a continuación la metodología mostrada por Cox *et al.* (2007) y J. Wu (1997).

8.1. Proposiciones geométricas y polinomios

El primer paso para utilizar demostradores algebraicos es trasladar las proposiciones geométricas en ecuaciones polinomiales. Para ello es necesario enfocarse en una geometría y hacer explícito el campo $\mathbb{K}$ de los coeficientes y variables de los polinomios (Ferro y Gallo, 1988). Así, en lo que sigue se utilizará un sistema de coordenadas cartesianas en $\mathbb{R}^2$, que permitirá traducir tanto las hipótesis como la conclusión de teoremas de geometría Euclideana en ecuaciones polinomiales. El problema entonces, será determinar si la conclusión es verdadera para todos aquellos puntos que satisfagan las hipótesis. En otras palabras, los teoremas que se podrán tratar con esta metodología son aquellos de la forma:

$$(\forall \vec{c} \in \mathbb{R}^{m+n})[h_1(\vec{c}) = 0 \wedge h_2(\vec{c}) = 0 \wedge \cdots \wedge h_k(\vec{c}) = 0 \;\Rightarrow\; g(\vec{c}) = 0]$$

donde $h_i \in \mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n]$ son las hipótesis y $g \in \mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n]$ es la conclusión.

Observación. Como regla general, se pensará en las variables u_i como *parámetros* que identifican coordenadas independientes. Las variables x_i denotarán las coordenadas que queden total o al menos parcialmente determinadas por las hipótesis del teorema y los puntos definidos previamente. Evidentemente esta elección de coordenadas no es única por cada teorema, pues depende de la construcción que se haga del mismo.

Es relevante notar que no todas las proposiciones geométricas son expresables de esta manera, la siguiente proposición enumera algunas de las proposiciones geométricas más comunes que pueden traducirse en ecuaciones polinomiales.

Proposición 8.1. Sean A, B, C, D, E, F puntos en el plano. Cada una de las siguientes proposiciones geométricas puede ser expresada por una o más ecuaciones polinomiales.

I) AB es paralela a CD.

II) AB es perpendicular a CD.

III) A, B y C son colineales.

IV) La medida de los segmentos $\overline{AB}$ y $\overline{CD}$ son iguales, *i.e.* $\overline{AB} = \overline{CD}$.

V) C se encuentra sobre una circunferencia con centro A y radio AB

VI) C es el punto medio de $\overline{AB}$.

VII) Los ángulos agudos $\angle ABC$ y $\angle DEF$ son iguales, *i.e.* $\angle ABC = \angle DEF$.

Demostración. Considérese el plano con un sistema de coordenadas cartesianas $\mathbb{R}^2$, por lo que se tiene

$$A = (u_1, u_2), \quad B = (u_3, u_4), \quad C = (u_5, u_6)$$

$$D = (u_7, u_8), \quad E = (u_9, u_{10}), \quad F = (u_{11}, u_{12})$$

I) Dos rectas son paralelas si, y sólo si, sus pendientes son iguales. Entonces

$$AB \parallel CD \Leftrightarrow \frac{u_4 - u_2}{u_3 - u_1} = \frac{u_8 - u_6}{u_7 - u_5} \Leftrightarrow (u_4 - u_2)(u_7 - u_5) = (u_8 - u_6)(u_3 - u_1)$$
$$\Leftrightarrow (u_4 - u_2)(u_7 - u_5) - (u_8 - u_6)(u_3 - u_1) = 0$$

donde además la expresión sigue siendo verdadera en el caso cuando las rectas son verticales.

II) Se aplica de nuevo el concepto de pendiente, notando ahora que dos rectas son perpendiculares si, y sólo si, el producto de sus pendientes es -1. Así,

$$AB \perp CD \Leftrightarrow \frac{u_4 - u_2}{u_3 - u_1} \cdot \frac{u_8 - u_6}{u_7 - u_5} = -1 \Leftrightarrow \frac{u_4 - u_2}{u_3 - u_1} = -\frac{u_7 - u_5}{u_8 - u_6} \Leftrightarrow$$
$$\Leftrightarrow (u_4 - u_2)(u_8 - u_6) = -(u_7 - u_5)(u_3 - u_1) \Leftrightarrow$$
$$\Leftrightarrow (u_4 - u_2)(u_8 - u_6) + (u_7 - u_5)(u_3 - u_1) = 0.$$

Nótese que de nuevo se conserva la igualdad en el caso que alguna de las rectas sea vertical.

III) Una forma de expresar esta proposición es requiriendo que C pertenezca a la recta AB. La recta que pasa por A y B se puede expresar como:

$$y = \frac{u_4 - u_2}{u_3 - u_1}(x - u_1) + u_2 \Leftrightarrow (y - u_2)(u_3 - u_1) - (u_4 - u_2)(x - u_1) = 0.$$

Nuevamente el caso cuando AB es vertical está considerado. Luego $C \in AB$ si, y sólo si, cuando $x = u_5$ y $y = u_6$ se cumple la identidad, es decir

$$(u_6 - u_2)(u_3 - u_1) - (u_4 - u_2)(u_5 - u_1) = 0$$

IV)
$$\overline{AB} = \overline{CD} \Leftrightarrow \sqrt{(u_4 - u_2)^2 + (u_3 - u_1)^2} = \sqrt{(u_8 - u_6)^2 + (u_7 - u_5)^2}$$
$$\Leftrightarrow (u_4 - u_2)^2 + (u_3 - u_1)^2 - (u_8 - u_6)^2 - (u_7 - u_5)^2 = 0.$$

V) Basta con que C cumpla con la ecuación de una circunferencia centrada en A y radio AB,

$$(x - u_1)^2 + (y - u_2)^2 = (u_4 - u_2)^2 + (u_3 - u_1)^2 \Leftrightarrow$$
$$\Leftrightarrow (x - u_1)^2 + (y - u_2)^2 - (u_4 - u_2)^2 - (u_3 - u_1)^2 = 0$$

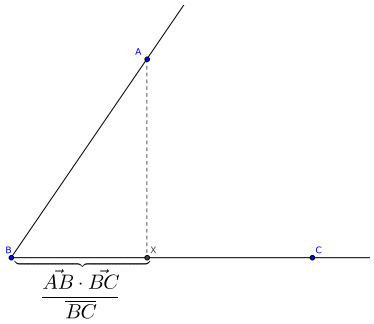
Por lo que C está en la circunferencia si y sólo si

$$(u_5 - u_1)^2 + (u_6 - u_2)^2 - (u_4 - u_2)^2 - (u_3 - u_1)^2 = 0$$

VI) Que C sea el punto medio de AB es equivalente a que las distancias $\overline{AC} = \overline{CB}$ y que C esté sobre la recta AB, por lo que se reduce a las ecuaciones de III y IV.

VII) Sean $\angle ABC$ y $\angle DEF$ ángulos agudos. Sean X y Y puntos sobre BC y EF tales que $AX \perp BC$ y $DY \perp EF$. El hecho que los ángulos $\angle ABC$ y $\angle DEF$ sean agudos asegura la existencia y unicidad de los puntos X y Y. Entonces, $\angle ABC = \angle DEF$ si, y sólo si, los triángulos

Figura 8.1: Proyección de $\overline{AB}$ sobre BC



rectángulos $\triangle ABX$ y $\triangle DEY$ son semejantes. Como son triángulos rectángulos, su semejanza se traduce en la igualdad de razones

$$\frac{\overline{AX}}{\overline{BX}} = \frac{\overline{DY}}{\overline{EY}}$$

Como $\overline{BX}$ es la proyección de $\overline{BA}$ sobre BC (ver Figura 8.1) y $\triangle ABX$ es rectángulo se tiene

$$\left(\frac{\overline{AX}}{\overline{BX}}\right)^2 = \frac{\overline{AX}^2}{\overline{BX}^2} = \frac{\overline{AB}^2 - \overline{BX}^2}{\overline{BX}^2} = \frac{\overline{AB}^2}{\overline{BX}^2} - 1 = \frac{\overline{AB}^2\overline{BC}^2}{(\vec{AB}\cdot\vec{BC})^2} - 1 =$$
$$= \frac{((u_3-u_1)^2+(u_4-u_2)^2)((u_5-u_3)^2+(u_6-u_4)^2)}{((u_3-u_1)(u_5-u_3)+(u_4-u_2)(u_6-u_4))^2} - 1$$

Análogamente

$$\frac{\overline{DY}^2}{\overline{EY}^2} = \frac{((u_9-u_7)^2+(u_{10}-u_8)^2)((u_{11}-u_9)^2+(u_{12}-u_{10})^2)}{((u_9-u_7)(u_{11}-u_9)+(u_{10}-u_8)(u_{12}-u_{10}))^2} - 1$$

Por lo tanto, los ángulos agudos $\angle ABC$ y $\angle DEF$ son iguales si y sólo si

$$\frac{\overline{AX}}{\overline{BX}} = \frac{\overline{DY}}{\overline{EY}} \Leftrightarrow$$
$$\frac{((u_3-u_1)^2+(u_4-u_2)^2)((u_5-u_3)^2+(u_6-u_4)^2)}{((u_3-u_1)(u_5-u_3)+(u_4-u_2)(u_6-u_4))^2} = \frac{((u_9-u_7)^2+(u_{10}-u_8)^2)((u_{11}-u_9)^2+(u_{12}-u_{10})^2)}{((u_9-u_7)(u_{11}-u_9)+(u_{10}-u_8)(u_{12}-u_{10}))^2} \Leftrightarrow$$
$$[((u_3-u_1)^2+(u_4-u_2)^2)((u_5-u_3)^2+(u_6-u_4)^2)][(u_9-u_7)(u_{11}-u_9)+(u_{10}-u_8)(u_{12}-u_{10})]^2 -$$
$$[((u_9-u_7)^2+(u_{10}-u_8)^2)((u_{11}-u_9)^2(u_{12}-u_{10})^2)][(u_3-u_1)(u_5-u_3)+(u_4-u_2)(u_6-u_4)]^2 = 0.\blacksquare$$

A continuación se ejemplifica cómo traducir un teorema de geometría euclideana a un sistema de ecuaciones polinomiales, utilizando la proposición 8.1.

***Ejemplo* 8.1**

Teorema. Sea $ABCD$ un paralelogramo en el plano ($AB \parallel CD$ y $AC \parallel BD$).

Entonces las diagonales $\overline{AD}$ y $\overline{BC}$ se intersectan en su punto medio.

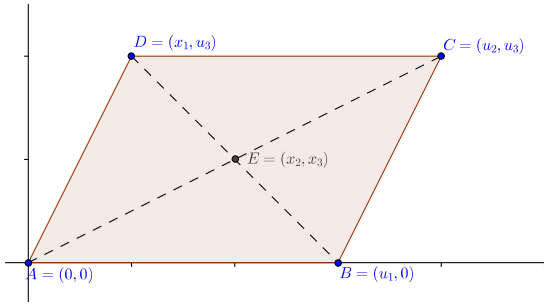
Es claro que la veracidad del teorema es independiente al sistema de coordenadas que se elija, por lo que por simplicidad se situará el origen en A, y el eje de las abscisas sobre la recta AB. De esta

forma los puntos quedan como

$$A = (0,0),\ B = (u_1,0),\ C = (u_2,u_3)$$

A diferencia de los puntos A, B y C que son independientes, D queda determinado por la elección de los otros tres, por lo que, por convención, en lugar de denotar sus coordenadas por variables u_i se denotarán por variables x_i. Más aún, podrían introducirse dos nuevas variables $D = (x_1, x_2)$ pero sería necesario incluir la condición $AB \parallel CD$. Es posible resumir toda esta información requiriendo de una vez que la coordenada en y de D sea la misma que C, *i.e.* $D = (x_1, u_3)$ reduciendo así el número de variables que aparecen en el teorema (Figura 8.2).

Figura 8.2: Paralelogramo sobre el plano $\mathbb{R}^2$ para traducir teorema en polinomios



Por la elección de los puntos A, B, C y D ya se tiene que $AB \parallel CD$. Para asegurar que $ABCD$ sea un paralelogramo se debe tener que $AD \parallel BC$:

$$(u_3 - 0)(u_2 - u_1) - (u_3 - 0)(x_1 - 0) = 0$$
$$h_1 = u_3u_2 - u_3u_1 - u_3x_1 = 0$$

Sea $E = (x_2, x_3)$ el punto de intersección de $\overline{AC}$ y $\overline{BD}$. Esto se puede interpretar como $E \in \overline{AC}$ y $E \in \overline{BD}$, por lo que se puede reescribir en ecuaciones polinomiales según el inciso III) de la proposición 8.1:

$$(x_3 - 0)(u_2 - 0) - (u_3 - 0)(x_2 - 0) = 0$$
$$h_2 = x_3u_2 - u_3x_2 = 0$$
$$(x_3 - 0)(x_1 - 0) - (u_3 - 0)(x_2 - u_1) = 0$$
$$h_3 = x_3x_1 - u_3x_2 + u_3u_1 = 0$$

Las conclusiones del teorema, que E es el punto medio de AC y BD se pueden escribir como:

$$\overline{AE} = \overline{EC}: \quad g_1 = x_2^2 + x_3^2 - (u_2 - x_2)^2 - (u_3 - x_3)^2 = 0$$
$$\overline{BE} = \overline{ED}: \quad g_2 = (x_2 - u_1)^2 + x_3^2 - (x_1 - x_2)^2 - (u_3 - x_3)^2 = 0$$

Así, se obtienen las 3 hipótesis h_1, h_2 y h_3 y las conclusiones g_1, g_2, polinomios en $\mathbb{Q}[u_1, u_2, u_3, x_1, x_2, x_3]$, aunque $u_1, u_2, u_3, x_1, x_2, x_3$ podrían ser cualquier valor de $\mathbb{R}$, por lo que el teorema se traduce en

las proposiciones:

$$(\forall\ u_1, u_2, u_3, x_1, x_2, x_3 \in \mathbb{R})[h_1 = 0 \wedge h_2 = 0 \wedge h_3 = 0 \Rightarrow g_1 = 0]$$

$$(\forall\ u_1, u_2, u_3, x_1, x_2, x_3 \in \mathbb{R})[h_1 = 0 \wedge h_2 = 0 \wedge h_3 = 0 \Rightarrow g_2 = 0]$$

Lo cual significa que para que el teorema sea verdadero, basta verificar que $\mathcal{V}(g_i) \supset \mathcal{V}_{\mathbb{R}}(\langle h_1, h_2, h_3\rangle)$ para $i = 1, 2$.

Como se ve en el ejemplo anterior, para que la conclusión g de un teorema sea verdadera dadas las hipótesis $h_1, \cdots, h_k$, es suficiente con que g sea cero siempre que lo sean los h_i, *i.e.* $\mathcal{V}_{\mathbb{R}}(g) \supset \mathcal{V}_{\mathbb{R}}(\langle h_1, \cdots, h_k\rangle)$. La siguiente definición formaliza esta idea.

Definición 8.2. *Sean $h_1, h_2, \cdots, h_k \in \mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n]$ las hipótesis de un teorema geométrico. Y sea $g \in \mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n]$ otro polinomio no cero. Se dice que g se* ***deduce estrictamente*** *de las hipótesis $h_1, \cdots, h_k$ si $g \in \mathcal{I}(V)$, con $V = \mathcal{V}(\langle h_1, \cdots, h_k\rangle)$.*

Observación. Note que por el lema 7.8 se tiene que

$$g \in \mathcal{I}(V) \Rightarrow\ \langle g\rangle \in \mathcal{I}(V) \Rightarrow\ \mathcal{V}(g) = \mathcal{V}(\langle g\rangle) \supset \mathcal{V}(\mathcal{I}(V)) \supset V$$

con $V = \mathcal{V}(\langle h_1, \cdots, h_k\rangle) = \mathcal{V}(\{h_1, \cdots, h_k\})$, por lo que

$$\mathcal{V}(g) \supset \mathcal{V}(\{h_1, \cdots, h_k\})$$

Así, g siempre se anula cuando los h_i's se anulan.

Proposición 8.3. Sean $h_1, \cdots, h_k, g$ polinomios en $\mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n]$. Si $g \in \sqrt{\langle h_1, \cdots, h_k\rangle}$ entonces g se deduce estrictamente de $h_1, \cdots, h_k$.

Demostración. Si $g \in \sqrt{\langle h_1, \cdots, h_k\rangle}$ entonces existe $r \in \mathbb{Z}^+$ tal que $g^r \in \langle h_1, \cdots, h_k\rangle$. Entonces $g^r = \sum c_i h_i$ para algunos $c_i \in \mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n]$, por lo que g^r, y, por ende, g se anulan cuando los h_i's se anulan, *i.e.* $g \in \mathcal{I}(\mathcal{V}(\langle h_1, \cdots, h_k\rangle))$. ∎

Nótese que por la observación al lema 7.8, es posible que

$$\sqrt{\langle h_1, \cdots, h_k\rangle} \subsetneq \mathcal{I}(\mathcal{V}(\sqrt{\langle h_1, \cdots, h_k\rangle})) = \mathcal{I}(\mathcal{V}(\langle h_1, \cdots, h_k\rangle)).$$

Por esto, aún cuando $g \notin \sqrt{\langle h_1, \cdots, h_k\rangle}$, podría darse el caso que g sí se deduzca estrictamente de $h_1, \cdots, h_k$. Considérese en cambio el ideal

$$I_{\mathbb{C}} = \langle h_1, \cdots, h_k\rangle \subseteq \mathbb{C}[u_1, \cdots, u_m, x_1, \cdots, x_n],$$

La versión fuerte del *Nullstellensatz* (teorema 7.20) asegura que $\mathcal{I}(\mathcal{V}(I_{\mathbb{C}})) = \sqrt{I_{\mathbb{C}}}$, para $\mathcal{V}(I_{\mathbb{C}}) \subseteq \mathbb{C}^{m+n}$. Resulta evidente que si $g \in \sqrt{I_C}$ entonces $g \in \sqrt{\langle h_1, \cdots, h_k\rangle_{\mathbb{R}}}$, pero también su recíproca es cierta (véase Cox *et al.*, 2007), por lo que

$$g \in \sqrt{\langle h_1, \cdots, h_k\rangle_{\mathbb{R}}}\ \Leftrightarrow\ g \in \sqrt{I_{\mathbb{C}}}.$$

Utilizando la metodología del corolario 7.21, se tiene

$$g \in \sqrt{\langle h_1, \cdots, h_k \rangle_{\mathbb{R}}} \Leftrightarrow \text{la base de Gröbner reducida de } \langle h_1, \cdots, h_k, 1 - yg \rangle \text{ es } \{1\}.$$

Esto significa que la propiedad 8.3 permite determinar si el teorema es verdadero para todo $\mathbb{C}^{m+n}$ o no. De esta forma, es sencillo probar teoremas en $\mathbb{C}$, que evidentemente se cumplirán también en $\mathbb{R}$, mas no es sencillo refutar teoremas en $\mathbb{R}$, pues podría darse el caso que el teorema se cumpla en $\mathbb{R}$ pero no su generalización en $\mathbb{C}$.

Además de la inconveniencia de no estar trabajando en un campo algebraicamente cerrado (en donde el Nullstellenzats fuerte aseguraría el recíproco de la propiedad 8.3), resulta ser que la definición 8.2 es muy fuerte para los teoremas que generalmente se estudian en geometría. Como menciona Cox *et al.* (2007), es usual enunciar teoremas en geometría que se cumplen siempre y cuando las hipótesis no estén en un caso degenerado (por ejemplo el segmento $\overline{AB}$ se degenera en un punto cuando $A = B$). El trabajo a continuación va encaminado a dar una definición de *deducción* que permita demostrar este tipo de teoremas también.

8.2. Demostraciones geométricas con casos degenerados

Considérese un teorema de geometría de la forma

$$(\forall \vec{c} \in \mathbb{R}^{m+n})[h_1(\vec{c}) = 0 \wedge \cdots \wedge h_k(\vec{c}) = 0 \Rightarrow g(\vec{c}) = 0]$$

con $h_1, \cdots, h_k, g \in \mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n]$. Si existe un polinomio $p \in \langle h_1, \cdots, h_k \rangle \cap \mathbb{R}[u_1, \cdots, u_m]$, entonces los u_i's no son totalmente independientes como se suponía, pues existen valores fijos para los u_i's que permiten que las coordenadas x_i's sean arbitrarias, contrario a lo que se suponía que eran dependientes de los u_i's. En estos casos en los que las hipótesis se *degeneran*, es irrelevante lo que suceda con la conclusión g, pues únicamente es necesario que g se cumpla cuando las hipótesis no están degeneradas.

Sea $V = \mathcal{V}(\langle h_1, \cdots, h_k \rangle) \subseteq \mathbb{R}^{m+n}$, el teorema 7.12 asegura que existe una descomposición irredundante única en variedades algebraicas irreducibles como

$$V = V_1 \cup V_2 \cup \cdots \cup V_r.$$

Si se diera el caso en que existan polinomios sólo sobre los u_i's que se anulen en algunos de los componentes V_j de V,el problema se reduciría a determinar si la conclusión sigue en el resto de componentes, pues como se mencionó anteriormente, es irrelevante qué suceda con g en los V_j's con casos degenerados. En otras palabras, se desea determinar si $g \in \mathcal{I}(V')$, con $V' \subseteq V$ la unión de los componentes de V en los que no se anula ningún polinomio de $\mathbb{R}[u_1, \cdots, u_k]$.

Definición 8.4. *Sea W una variedad algebraica irreducible en $\mathbb{R}^{m+n}$ con coordenadas $u_1, \cdots, u_m$, $x_1, \cdots, x_n$. Se dice que las variables $u_1, \cdots, u_m$ son* ***algebraicamente independientes en*** *W si $\mathbb{R}[u_1, \cdots, u_m] \cap \mathcal{I}(W) = \{0\}$, i.e. no existe ningún polinomio únicamente sobre las variables*

$u_1, \cdots, u_m$ que se anule sobre W.

Definición 8.5. *Sean $h_1, h_2, \cdots, h_k \in \mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n]$ las hipótesis de un teorema geométrico. Y sea $g \in \mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n]$ otro polinomio no cero. Se dice que g se* ***deduce genéricamente*** *de las hipótesis $h_1, \cdots, h_k$ si $g \in \mathcal{I}(V')$, donde V' es la unión de los componentes irreducibles de $V = \mathcal{V}(\langle h_1, \cdots, h_k \rangle)$ en los que los u_i's son algebraicamente independientes.*

Proposición 8.6. Sean $h_1, \cdots, h_k, g$ polinomios en $\mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n]$. Si existe un polinomio distinto de cero $c(u_1, \cdots, u_m) \in \mathbb{R}[u_1, \cdots, u_m]$ tal que $c \cdot g \in \sqrt{H}$, con $H = \langle h_1, \cdots, h_k \rangle \subseteq \mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n]$. Entonces g se deduce genéricamente de $h_1, \cdots, h_k$.

Demostración. Sea V_i una de las componentes irreducibles de V' como en la definición 8.5. Como $cg \in \sqrt{H}$, $cg \in \mathcal{I}(V') \subseteq \mathcal{I}(V_i)$. Por la proposición 7.10, como V_i es una variedad irreducible, $\mathcal{I}(V_i)$ debe ser un ideal primo. Luego, c no puede estar en $\mathcal{I}(V_i)$, pues los u_i's son algebraicamente independientes sobre V_i, entonces $g \in \mathcal{I}(V_i)$, para todos los componentes de V'. Por lo tanto

$$g \in \bigcap \mathcal{I}(V_i) = \mathcal{I}\left(\bigcup V_i\right) = \mathcal{I}(V')$$

y g se deduce genéricamente de $h_1, \cdots, h_k$. ∎

Teorema 8.7. Sean $h_1, \cdots, h_k, g$ polinomios en $\mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n]$. Sea H el ideal generado por los h_i's en $\mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n]$. Y sea $\widetilde{H}$ el ideal generado por los h_i's en el anillo $\mathbb{R}(u_1, \cdots, u_m)[x_1, \cdots, x_n]$. Entonces los siguientes enunciados son equivalentes:

1. Existe un polinomio distinto de cero $c(u_1, \cdots, u_m) \in \mathbb{R}[u_1, \cdots, u_m]$ tal que $cg \in \sqrt{H}$.
2. $g \in \sqrt{\widetilde{H}}$.
3. $\{1\}$ es la base de Gröbner reducida del ideal

$$\langle h_1, \cdots, h_k, 1 - yg \rangle \subseteq \mathbb{R}(u_1, \cdots, u_m)[x_1, \cdots, x_n, y].$$

Demostración.

- (1⇒2): Supóngase que existe $c \in \mathbb{R}[u_1, \cdots, u_m]$ distinto de cero tal que $c \cdot g \in \sqrt{H}$ entonces existe una potencia r tal que $(c \cdot g)^r = c^r g^r \in H$. Entonces

$$c^r g^r = \sum_{i=1}^{k} p_i h_i, \text{ con } p_i \in \mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n] \Rightarrow$$
$$g^r = \sum_{i=1}^{k} \frac{p_i}{c^r} h_i,$$

  lo cual es posible hacer pues $c^r \neq 0$. Evidentemente ahora los $\frac{p_i}{c^r} \in \mathbb{R}(u_1, \cdots, u_m)[x_1, \cdots, x_n]$, pues su denominador únicamente contiene a las variables $u_1, \cdots, u_m$. Por lo tanto, $g^r \in \widetilde{H} \Rightarrow g \in \sqrt{\widetilde{H}}$.

- (2⇒ 1): Si $g \in \sqrt{\widetilde{H}}$ entonces existe $r \in \mathbb{Z}^+$ tal que $g^r \in \widetilde{H} \Rightarrow$

$$g^r = \sum_{i=1}^{k} p_i h_i,$$

donde los $p_i \in \mathbb{R}(u_1, \cdots, u_m)[x_1, \cdots, x_n]$ y los $h_i \in \mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n]$. Sea c el producto de los k denominadores de los p_i's (pudiendo estos ser 1). Claramente $c \in \mathbb{R}[u_1, \cdots, u_m]$, $c \neq 0$ y $c^r p_i \in \mathbb{R}[u_1, \cdots, u_m, x_1, \cdots, x_n]$ $\forall\ i = 1, \cdots, k$. Multiplicando por c^r la expresión anterior se obtiene

$$(c \cdot g)^r = \sum_{i=1}^{k} (c^r p_i) h_i,$$

por lo que $(c \cdot g)^r \in H \Rightarrow\ c \cdot g \in \sqrt{H}$.

- (2⇒ 3): $g \in \sqrt{\widetilde{H}}$ entonces $g^r \in \widetilde{H}$ y evidentemente $y^r g^r \in \widetilde{H}$. Entonces existen $p_i \in \mathbb{R}(u_1, \cdots, u_m)[x_1, \cdots, x_n]$, $i = 1, \cdots, k$ tales que

$$y^r g^r = p_1 h_1 + \cdots + p_k h_k.$$

Luego se obtiene

$$\begin{aligned} 1 &= y^r g^r + (1 - y^r g^r) = y^r g^r + (1 + yg + y^2 g^2 + \cdots + y^{r-1} g^{r-1})(1 - yg) \\ &= p_1 h_1 + p_2 h_2 + \cdots + p_k h_k + (1 + yg + y^2 g^2 + \cdots + y^{r-1} g^{r-1})(1 - yg) \\ &\in \langle h_1, h_2, \cdots, h_k, 1 - yh \rangle \subseteq \mathbb{R}(u_1, \cdots, u_m)[x_1, \cdots, x_n, y] \end{aligned}$$

Lo anterior implica que $\langle h_1, h_2, \cdots, h_k, 1 - yg \rangle = \mathbb{R}(u_1, \cdots, u_m)[x_1, \cdots, x_n, y]$. Luego, como la base reducida de Gröbner de un ideal es única, su base debe ser $\{1\}$.

- (3⇒ 2): Si la base de Gröbner de $\langle h_1, h_2, \cdots, h_k, 1 - yh \rangle$ es $\{1\}$, deben existir polinomios $p_1, \cdots, p_{k+1} \in \mathbb{R}(u_1, \cdots, u_m)[x_1, \cdots, x_n, y]$ tales que

$$1 = p_1 h_1 + \cdots + p_k h_k + p_{k+1}(1 - yg)$$

Imitando el procedimiento utilizado en la demostración de la versión fuerte del *Nullstellensatz* (teorema 7.20), se sustituye y por $\frac{1}{g}$ en la igualdad anterior, obteniendo

$$1 = p'_1 h_1 + \cdots + p'_k h_k$$

donde

$$p'_i = p_i \left(x_1, \cdots, x_n, \frac{1}{g} \right),$$

y los h_i's no cambian pues no dependen de y. Como todos los denominadores de los p'_i son alguna potencia de g, se puede encontrar una potencia r tal que $g^r p_i \in \mathbb{R}(u_1, \cdots, u_m)[x_1, \cdots, x_n]$ $\forall\ i = 1, \cdots, k$. Entonces, multiplicando la igualdad anterior por g^r se obtiene

$$g^r = (g^r p'_1) h_1 + \cdots + (g^r p'_k) h_k \in \widetilde{H},$$

por lo tanto, $g \in \sqrt{\widetilde{H}}$. ■

El teorema anterior constituye entonces el resultado culmen de este trabajo, permitiendo utilizar el cálculo de bases de Gröbner para determinar algorítmicamente si un resultado geométrico es deducible genéricamente a partir de otros. Este proceso será entendido como *demostrar de forma automatizada* un teorema de geometría, y se puede resumir de la siguiente manera.

Problema 5. (Demostrar un teorema de geometría euclideana) Demostrar que un teorema de geometría euclideana τ es verdadero (siempre y cuando no se degeneren las hipótesis).

Solución. Si tanto las hipótesis como la conclusión se pueden escribir en forma de ecuaciones polinomiales (como los enunciados de la proposición 8.1), entonces el siguiente proceso permite verificar la validez del teorema en $\mathbb{C}$ (lo que implicaría que también es cierto en $\mathbb{R}$, pero no se puede utilizar para determinar su falsedad en $\mathbb{R}$).

1. Elegir un sistema de coordenadas cartesiano y traducir las hipótesis y conclusión de τ en ecuaciones polinomiales. Para ello utilizar las variables $u_1, \cdots, u_m$ para los parámetros libres y $x_1, \cdots, x_n$ para aquellos que quedan determinados por las construcciones anteriores.

2. Utilizando el algoritmo de Buchberger (en cualquiera de sus implementaciones, por ejemplo el algoritmo 6.2) calcular una base de Gröbner G para el ideal

$$\langle h_1, \cdots, h_k, 1 - yg \rangle \in \mathbb{R}(u_1, \cdots, u_m)[x_1, \cdots, x_n]$$

   donde h_i son las hipótesis de τ y g su conclusión, y es una variable más.

3. Dependiendo de la versión del algoritmo de Buchberger utilizada, si este no reduce la base de Gröbner, utilizar el algoritmo 5.1 para reducir la base de Gröbner G en G'.

4. Si $G' = \{1\}$ el teorema τ es verdadero en $\mathbb{C}$, y por tanto en $\mathbb{R}$, por lo que el teorema de geometría es cierto en el espacio cartesiano $\mathbb{R}^d$ utilizado. Si $G' \neq \{1\}$, el teorema no es verdadero en su generalización para $\mathbb{C}$, aunque no puede concluirse nada sobre su veracidad en el espacio $\mathbb{R}^d$

8.3. Ejemplo: El hexagrama de Pascal

En esta sección se da una demostración automatizada del teorema del hexagrama de Pascal, ejemplificando así la utilidad del método de demostración automatizada con bases de Gröbner. A criterio personal del autor, el teorema de Pascal es el candidato perfecto para poner a prueba el método desarrollado, pues desde su enunciado mismo menciona objetos eminentemente algebraicos como lo son las cónicas.

El teorema en cuestión fue descubierto por el filósofo y matemático Blaise Pascal alrededor de 1640, teniendo sólo 16 años de edad. Como señala Coxeter (1967), la prueba original de Pascal es desconocida, pues se perdió el manuscrito. Sin embargo, es sabido que G. W. Leibniz leyó e incluso

halagó la prueba del joven Pascal (Coxeter, 1967). Dados los resultados y métodos disponibles en la época de Pascal, Coxeter (1967) sugiere que la prueba original pudo haber estado limitada únicamente al caso en el que la cónica fuera una circunferencia, probablemente utilizando alguna forma del teorema de Menelao. En todo caso, Pascal estaba consciente de que su teorema se aplicaba a todas las cónicas, según su ensayo sobre cónicas "*Essay pour les coniques*", el cuál ha sobrevivido hasta la fecha (Coxeter, 1967).

Actualmente, como menciona Borodzik y Zoladek (2002), se suele presentar una de dos demostraciones del hexagrama de Pascal: una utilizando tétradas armónicas para el caso del círculo y geometría proyectiva para la generalización a cónicas; u otra que depende del teorema de Bacharach-Cayley, por medio del cual se cuenta el número de intersecciones de curvas cúbicas en el plano proyectivo. La demostración automatizada que se proveerá a continuación, aunque no permita ver los pasos de una demostración formal, permitirá verificar la validez del teorema en los casos genéricos. Se utilizará también la metodología de bases de Gröbner para probar uno de los casos degenerados del teorema de Pascal, siendo este el caso del teorema de Pappus.

8.3.1. El Teorema El siguiente es el enunciado del teorema que aparece en Coxeter (1967), limitada al caso de la circunferencia como seguramente lo habría enunciado originalmente B. Pascal.

> **Teorema** (Hexagrama de Pascal)**.** Si los seis puntos de un hexágono yacen sobre la misma circunferencia y las tres parejas de lados paralelos se intersectan, entonces los tres puntos de intersección son colineales.

Es relevante mencionar que a pesar de que este enunciado exija que los pares de lados opuestos se intersecten y que el hexágono tenga seis vértices, el teorema sigue siendo válido en dichos casos degenerados (véase Coxeter, 1967). Sin embargo en dichos casos es necesario dar una nueva interpretación de las hipótesis y conclusión del teorema, por lo que no se considerarán estos casos en este trabajo. De igual forma, al sustituir la palabra circunferencia por cónica, el teorema sigue siendo válido incluso para cónicas degeneradas como lo son un par de rectas, este caso constituye el teorema de Pappus que se demostrará más adelante.

Se enuncia entonces a continuación el teorema de una manera más adecuada para su posterior análisis

> **Teorema** (Hexagrama de Pascal-Versión 2)**.** Sean A, B, C, A', B' y C' seis puntos cualesquiera sobre una cónica Γ. Si los pares de rectas AB' con $A'B$, AC' con $A'C$, y BC' con $B'C$ se intersectan en los puntos X, Y y Z, respectivamente. Entonces, X, Y y Z son colineales.

8.3.2. Demostración La implementación de la demostración automatizada se hará con el sistema algebraico computacional SAGE (Stein *et al.*, 2015) y se puede encontrar en el Apéndice 11.1. SAGE es un software matemático gratis y de código abierto, construido sobre otros numerosos paquetes de código abierto como: NumPy, SciPy, matplotlib, Sympy, Maxima, GAP, FLINT y R, combinados con un lenguaje común basado en Python.

Obsérvese que la cónica Γ se puede representar como la variedad $\mathcal{V}(\gamma) \subseteq \mathbb{R}^2$, donde γ es una curva cuadrática (Weisstein, s.f.), por lo que tiene la forma:

$$\Gamma : \gamma(a, b, c, d, f, g; x, y) = ax^2 + by^2 + cxy + dx + fy + g = 0, \tag{8.1}$$

donde $a, b, c, d, f, g \in \mathbb{R}$ son parámetros que caracterizan a Γ. Nótese sin embargo que estos 6 parámetros no son únicos por cada cónica, pues sus múltiplos caracterizan a la misma variedad. Por la construcción que se hará de las hipótesis del teorema, será útil que los parámetros de Γ queden determinados por la elección de algunos de los puntos de la cónica, por lo que el teorema se demostrará identificando dos casos (más adelante se verá que basta considerar únicamente dos casos para demostrar los casos genéricos del teorema): (1) cuando $a \neq 0$ y (2) cuando $a = 0$.

Caso 1: $a \neq 0$**.** Dividiendo la ecuación 8.1 y renombrando las variables adecuadamente, se caracteriza la cónica Γ mediante

$$\Gamma : \gamma(1, b, c, d, f, g; x, y) = x^2 + by^2 + cxy + dx + fy + g = 0 \tag{8.2}$$

la cual tiene exactamente cinco parámetros que la determinan de forma biunívoca. Dado que cinco puntos determinan una cónica (Weisstein, s.f.), los cinco parámetros quedarán determinados luego de elegir los cinco primeros puntos del hexágono. Se comienza entonces por elegir un sistema de coordenadas cartesianas, por simplicidad se elige el sistema con origen en A y el punto $B = (0, 1)$, por lo que se definen los primeros cinco puntos:

$$A = (0, 0) \qquad B = (0, 1) \qquad C = (u_1, u_2) \qquad A' = (u_3, u_4) \qquad B' = (u_5, u_6)$$

Luego, los parámetros b, c, d, f y g deben ser tales que los cinco puntos anteriores satisfagan la ecuación 8.2. Para A, eso se traduce en

$$\gamma(1, b, c, d, f, g; 0, 0) = g = 0 \Rightarrow g = 0.$$

Luego, sabiendo que $g = 0$, para B se obtiene:

$$\gamma(1, b, c, d, f, 0; 0, 1) = b + f = 0 \Rightarrow f = -b.$$

Por lo que la ecuación 8.2 se reduce a la ecuación más simple

$$\Gamma : \gamma(1, b, c, d, -b, 0; x, y) = x^2 + by^2 + cxy + dx - by = 0. \tag{8.3}$$

Entonces b, c y d deben cumplir con la ecuación 8.3 para los puntos C, A' y B', con lo que se

plantean las primeras hipótesis del teorema

$$h_1 = u_1^2 + bu_2^2 + cu_1u_2 + du_1 - bu_2 = 0$$
$$h_2 = u_3^2 + bu_4^2 + cu_3u_4 + du_3 - bu_4 = 0$$
$$h_3 = u_5^2 + bu_6^2 + cu_5u_6 + du_5 - bu_6 = 0$$

Para parametrizar el sexto punto, C', de la cónica, nótese que su primera coordenada se puede elegir arbitrariamente, pero la segunda debe ser tal que C' cumpla también con la ecuación 8.3. Los puntos X, Y y Z quedan totalmente determinados luego de construir los puntos del hexágono, por lo que todas sus coordenadas son dependientes, así se parametrizan el resto de los puntos de la construcción como

$$C' = (u_7, x_1) \qquad X = (x_2, x_3) \qquad Y = (x_4, x_5) \qquad Z = (x_6, x_7)$$

Entonces se agrega la hipótesis que $C' \in \Gamma$

$$h_4 = u_7^2 + bx_1 + cu_7x_1 + du_1 - bx_1 = 0,$$

y las hipótesis que X, Y y Z son los puntos de intersección de los lados del hexágono, utilizando el inciso III) de la proposición 8.1

$$X = AB' \cap A'B$$
$$A, B', X \text{ son colineales} : h_5 = u_5(u_6 - x_3) - (u_5 - x_2)u_6 = 0$$
$$A', B, X \text{ son colineales} : h_6 = u_3(x_3 - 1) - (u_4 - 1)x_2 = 0$$

$$Y = AC' \cap A'C$$
$$A, C', Y \text{ son colineales} : h_7 = u_7(x_1 - x_5) - (u_7 - x_4)x_1 = 0$$
$$A', C, Y \text{ son colineales} : h_8 = (u_1 - u_3)(u_2 - x_5) - (u_1 - x_4)(u_2 - u_4)$$

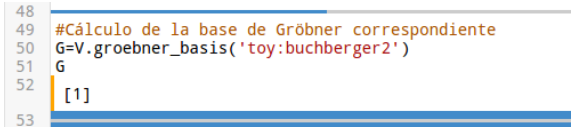
$$Z = BC' \cap B'C$$
$$B, C', Z \text{ son colineales} : h_9 = u_7(x_1 - x_7) - (u_7 - x_6)(x_1 - 1) = 0$$
$$B', C, Z \text{ son colineales} : h_{10} = (u_1 - u_5)(u_2 - x_7) - (u_1 - x_6)(u_2 - u_6) = 0$$

Finalmente, la conclusión X, Y y Z colineales se traduce como

$$g = (x_3 - x_5)(x_4 - x_6) - (x_2 - x_4)(x_5 - x_7).$$

De esta forma, se obtienen diez hipótesis $h_1, h_2, \cdots, h_{10}$, y una conclusión g para el teorema, traducidos a polinomios con parámetros $u_1, \cdots, u_7$ y variables dependientes $b, c, d, x_1, \cdots, x_7$. Pasando al paso 2) (según la solución al problema 5), se debe encontrar la base de Gröbner G para

Figura 8.3: Resultado Implementación en SAGE Teorema de Pascal Caso $a \neq 0$



el ideal:

$$\langle h_1, h_2, \cdots, h_{10}, 1 - yg \rangle \subseteq \mathbb{R}(u_1, \cdots, u_7)[b, c, d, x_1, \cdots, x_7, y].$$

Utilizando el algoritmo de Buchberger. Para ello se utilizan las líneas 45-47 del código en el Apéndice 11.1, utilizando la función `V.groebner_basis('toy:buchberger2')` de SAGE, donde el argumento `'toy:buchberger2'` le indica que utilice la implementación del algoritmo de Buchberger utilizando los tres criterios de Buchberger como en el algoritmo 6.2. Es importante notar que el ideal `V` al que se le obtiene su base de Gröbner, es un ideal del anillo

$$R = \mathbb{Q}(u_1, \cdots, u_7)[b, c, d, x_1, \cdots, x_7, y]$$

con campo $\mathbb{Q}$ en lugar de $\mathbb{R}$. Sin embargo, es claro que todos los polinomios $h_1, \cdots, h_1, g \in R$, y dado que el algoritmo de Buchberger únicamente realiza operaciones elementales con los coeficientes de los polinomios (suma, resta, multiplicación y división), la base de Gröbner de `V` es la misma sin importar si se calcula en $\mathbb{R}$ o $\mathbb{Q}$.

Finalmente, ejecutando el código del Apéndice 11.1, se obtiene el resultado esperado (Figura 8.3):

$$G = \{1\}$$

por lo que, según el teorema 8.7, g se deduce genéricamente de $h_1, \cdots, h_{10}$ y queda demostrado el teorema de Pascal para casos no degenerados cuando la cónica correspondiente tiene coeficiente $a \neq 0$.

Caso 2: $a = 0$**.** Se considera $b \neq 0$, pues si b también fuera cero, la cónica, como se verá después, se degenera en un par de rectas, reduciéndose así al teorema de Pappus que se demuestra después. Dividiendo la ecuación 8.1 por b, haciendo $a = 0$ y renombrando las variables adecuadamente, se caracteriza la cónica Γ mediante

$$\Gamma : \gamma(0, 1, c, d, f, g; x, y) = y^2 + cxy + dx + fy + g = 0 \tag{8.4}$$

la cual tiene exactamente cuatro parámetros que la determinan de forma biunívoca. Estos parámetros quedarán determinados luego de elegir los cuatro primeros puntos del hexágono. Se elige entonces el mismo sistema de coordenadas que en el Caso 1, por lo que se definen los primeros cuatro puntos:

$$A = (0, 0) \qquad B = (0, 1) \qquad C = (u_1, u_2) \qquad A' = (u_3, u_4).$$

Para A, la ecuación 8.4 se traduce en

$$\gamma(0,1,c,d,f,g;0,0) = g = 0 \Rightarrow g = 0.$$

Luego, sabiendo que $g = 0$, para B se obtiene:

$$\gamma(0,1,c,d,f,0;0,1) = 1 + f = 0 \Rightarrow f = -1.$$

Nótese de esta última igualdad que si $b = 0$ en lugar de 1, f debería ser cero también, por lo que la cónica se reduciría a la ecuación

$$\gamma(0,0,c,d,0;x,y) = cxy + dx = 0 \Leftrightarrow x = 0 \text{ y } cy + d = 0,$$

que es una pareja de rectas para cualesquiera valores de c y d, caso que se demostrará más adelante.

Regresando al caso $a = 0, b \neq 0$, la ecuación 8.4 se convierte en la ecuación más simple

$$\Gamma : \gamma(0,1,c,d,-1,0;x,y) = y^2 + cxy + dx - y = 0. \tag{8.5}$$

Entonces c y d deben cumplir con la ecuación 8.5 para los puntos C y A', con lo que se plantean las primeras hipótesis del teorema

$$h_1 = u_2^2 + cu_1u_2 + du_1 - u_2 = 0,$$
$$h_2 = u_4^2 + cu_3u_4 + du_3 - u_4 = 0.$$

Nótese ahora que aunque la primera coordenada de los puntos B' y C' puede ser arbitraria, la segunda queda determinada por la cónica Γ, mientras que para los puntos X, Y y Z, las dos coordenadas quedan totalmente determinadas por los puntos del hexágono. Así, se escriben el resto de puntos como

$$B' = (u_5, x_1) \qquad C' = (u_6, x_2) \qquad X = (x_3, x_4) \qquad Y = (x_5, x_6) \qquad Z = (x_7, x_8).$$

Se agregan la hipótesis $B' \in \Gamma$ y $C' \in \Gamma$

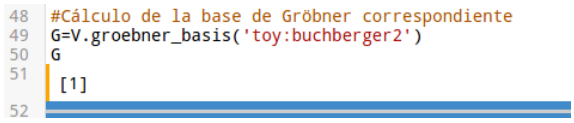
$$h_3 = x_1^2 + cu_5x_1 + du_5 - x_1 = 0,$$
$$h_4 = x_2^2 + cu_6x_2 + du_6 - x_2 = 0,$$

y las hipótesis que X, Y y Z son los puntos de intersección de los lados del hexágono:

$$X = AB' \cap A'B$$
$$A, B', X \text{ son colineales} : h_5 = u_5(x_1 - x_4) - (u_5 - x_3)x_1 = 0$$
$$A', B, X \text{ son colineales} : h_6 = u_3(x_4 - 1) - (u_4 - 1)x_3 = 0$$

$$Y = AC' \cap A'C$$
$$A, C', Y \text{ son colineales} : h_7 = u_6(x_2 - x_6) - (u_6 - x_5)x_2 = 0$$
$$A', C, Y \text{ son colineales} : h_8 = (u_1 - u_3)(u_2 - x_6) - (u_1 - x_5)(u_2 - u_4) = 0$$

Figura 8.4: Resultado Implementación en SAGE Teorema de Pascal Caso $a \neq 0$



$$Z = BC' \cap B'C$$

$$B, C', Z \text{ son colineales} : h_9 = u_6(x_2 - x_8) - (u_6 - x_7)(x_2 - 1) = 0$$

$$B', C, Z \text{ son colineales} : h_{10} = (u_1 - u_5)(u_2 - x_8) - (u_1 - x_7)(u_2 - x_1) = 0$$

Finalmente, la conclusión X, Y y Z colineales se traduce como

$$g = (x_4 - x_6)(x_5 - x_7) - (x_3 - x_5)(x_6 - x_8)$$

De esta forma, se obtienen diez hipótesis y una conclusión para el teorema, traducidos a polinomios con parámetros $u_1, \cdots, u_6$ y variables dependientes $c, d, x_1, \cdots, x_8$. Ahora bien, se debe encontrar la base de Gröbner G para el ideal:

$$\langle h_1, h_2, \cdots, h_{10}, 1 - yg \rangle \subseteq \mathbb{R}(u_1, \cdots, u_6)[c, d, x_1, \cdots, x_8, y].$$

Utilizando el algoritmo de Buchberger. Para ello se utilizan las líneas 44-46 del código en el Apéndice 11.2, utilizando la función `V.groebner_basis('toy:buchberger2')` de SAGE, tal y como se hizo en el Caso 1.

Finalmente, ejecutando el código del Apéndice 11.2, se obtiene el resultado esperado (Figura 8.4):

$$G = \{1\}$$

por lo que, según el teorema 8.7, g se deduce genéricamente de $h_1, \cdots, h_{10}$ y queda demostrado el teorema de Pascal para casos no degenerados cuando la cónica correspondiente tiene coeficientes $a = 0$ y $b \neq 0$.

8.3.3. Caso Degenerado: Teorema de Pappus Aunque ya se demostró el caso genérico del teorema de Pascal, resulta interesante considerar el caso cuando la cónica se degenera en un par de rectas (lo que sucede cuando $a = b = 0$ en la ecuación 8.1). Este caso lo constituye el teorema de Pappus, el cual fue demostrado por primera vez hacia el año 300 A. de C. por Pappus de Alejandría, y cuyo rol en los fundamentos de la geometría proyectiva fue reconocido hasta 16 siglos después (Coxeter, 1967). El enunciado del teorema es como sigue

> **Teorema** (Pappus)**.** Sean A, B y C tres puntos sobre una recta, y A', B' y C', puntos sobre otra. Si las rectas AB', AC' y BC' se intersectan con las rectas $A'B$, $A'C$ y $B'C$ en los puntos X, Y y Z, respectivamente. Entonces los puntos X, Y y Z son colineales.

Se procede entonces a dar la demostración del teorema siguiendo los pasos de la solución al problema 5. Por simplicidad, se elige el sistema coordenado con origen en A y $B = (0,1)$. Considerando que C y C' únicamente tendrán una coordenada libre, pues deben caer en las rectas AB y $A'B'$, respectivamente, y que X, Y y Z están completamente determinados, se obtiene:

$$A = (0,0), \qquad B = (0,1), \qquad C = (0,u_1)$$
$$A' = (u_2,u_3), \qquad B' = (u_4,u_5), \qquad C' = (u_6,x_1)$$
$$X = (x_2,x_3), \qquad Y = (x_4,x_5), \qquad Z = (x_6,x_7)$$

Dado que A, B y C son colineales por construcción, las hipótesis restantes del teorema se traducen como:

$$A', B' \text{ y } C' \text{colineales: } h_1 = (u_3-u_5)(u_4-u_6) - (u_2-u_4)(u_5-x_1) = 0$$
$$X = A'B \cap AB'$$
$$A', B \text{ y } X \text{colineales: } h_2 = u_2(x_3-1) - (u_3-1)x_2 = 0$$
$$A, B' \text{ y } X \text{colineales: } h_3 = u_4(u_5-x_3) - (u_4-x_2)u_5 = 0$$
$$Y = A'C \cap AC'$$
$$A', C \text{ y } Y \text{colineales: } h_4 = (u_1-u_3)x_4 - (u_1-x_5)u_2 = 0$$
$$A, C' \text{ y } Y \text{colineales: } h_5 = u_6(x_1-x_5) - (u_6-x_4)x_1 = 0$$
$$Z = B'C \cap BC'$$
$$B', C \text{ y } Z \text{colineales: } h_6 = (u_1-u_5)x_6 - (u_1-x_7)u_4 = 0$$
$$B, C' \text{ y } Z \text{colineales: } h_7 = u_6(x_1-x_7) - (u_6-x_6)(x_1-1) = 0$$

Y la conclusión se traduce en

$$X, Y \text{ y } Z \text{ colineales: } g = (x_3-x_5)(x_4-x_6) - (x_2-x_4)(x_5-x_7) = 0$$

Finalmente, se encuentra la base de Gröbner del ideal

$$\langle h_1, \cdots, h_7, 1-gy \rangle \subseteq \mathbb{R}(u_1, \cdots, u_6)[x_1, \cdots, x_7, y]$$

ejecutando las líneas 43-45 del código del Apéndice 11.3 (Figura 8.5), con lo que se obtiene

$$G = \{1\}$$

Por lo que se ha demostrado el teorema de Pappus genéricamente.

Figura 8.5: Resultado Implementación en SAGE Teorema de Pappus

```
46
47  #Calcular base de Gröbner
48  G=V.groebner_basis('toy:buchberger2')
49  G
50
51
    [1]
```

9 CONCLUSIONES

Aquí concluye la introducción a la teoría de bases de Gröbner que se propuso hacer en este trabajo. Las aplicaciones de esta teoría son innumerables y actualmente se sigue investigando su aplicación a toda una gama de distintos problemas. Note que la base de todo el desarrollo mostrado, fue extrapolar el algoritmo de la división de polinomios sobre una variable a los anillos de polinomios en varias variables. Ante el problema de la no unicidad del residuo del algoritmo de reducción en varias variables, Buchberger construyó las bases de su teoría que se muestran en este trabajo. El desarrollo iniciado por Buchberger es impresionante, pues no sólo define las bases de Gröbner como los objetos que permiten solucionar este problema, sino que demuestra su existencia para todo ideal de polinomios en varias variables y proporciona un algoritmo para calcularlas. Se espera que la manera en la que fueron expuestas estas ideas en este escrito, permitan al lector apreciar lo admirable del trabajo de Buchberger, a la vez que la simpleza y belleza de sus ideas.

Con el objetivo de mostrar esta poderosa herramienta a los interesados en el tema, aun si no han llevado antes un curso completo de álgebra abstracta, se presentó en el Capítulo 2 una recopilación de los resultados sobre anillos, campos e ideales que fueron necesarios durante el desarrollo de este trabajo. En el Capítulo 3 se trabajó con el anillo de polinomios en una variable, demostrando algunas de las propiedades más importantes y conocidas del mismo. El desarrollo del Capítulo 3 se hizo de tal forma que la construcción de los anillos de polinomios en varias variables del Capítulo 4, fuera inmediata. Aunque muchas de las propiedades de $\mathbb{K}[x]$ se extrapolan a $\mathbb{K}[x_1, \cdots, x_n]$, se mostró por qué el algoritmo de la división de $\mathbb{K}[x]$ es inútil en $\mathbb{K}[x_1, \cdots, x_n]$. Fue necesario entonces estudiar los órdenes de monomios y sus propiedades, los cuáles permiten construir un algoritmo de reducción que extiende las ideas del de la división de $\mathbb{K}[x]$ a $\mathbb{K}[x_1, \cdots, x_n]$. Además, en este desarrollo, se demostraron dos resultados muy importantes en el álgebra que merecen mencionarse: el lema de Dickson y el teorema de las bases de Hilbert, los cuáles son imprescindibles para el funcionamiento del algoritmo de Buchberger, asegurando la finitud de las bases de Gröbner.

En el Capítulo 5 se introdujo las bases de Gröbner, mostrando que en efecto solucionan el problema de la unicidad del residuo del algoritmo de reducción, como se quería. Luego, en el Capítulo 6, se discutió el algoritmo de Buchberger para calcularlas. Estudiando la complejidad de este algoritmo, se vio que no puede ser menor a $2^{2^{O(d)}}$, d el mayor grado de los generadores del ideal, sin embargo se introdujeron algunas mejoras para evitar redundancias en el algoritmo, y se argumentó que la mayoría de problemas que surgen en las aplicaciones se pueden resolver en tiempos mucho menores a este. Aún así, la complejidad del algoritmo de Buchberger y otros algoritmos

para calcular bases de Gröbner, es un tema que se investiga ampliamente en la actualidad. Al final del Capítulo 6, se muestra también las primeras aplicaciones de las bases de Gröbner: el determinar si un polinomio cualquiera pertenece o no a un ideal dado I; y el caracterizar la forma del anillo cociente $\mathbb{K}[x_1, \cdots, x_n]/I$.

En los últimos Capítulos, 7 y 8, se muestran otras aplicaciones de las bases de Gröbner. Las aplicaciones del Capítulo 7 son quizás las más útiles, pues se muestra cómo resolver sistemas de ecuaciones utilizando este método. A su vez, en el mismo Capítulo, se muestra uno de los resultados más importantes de la geometría algebraica: el *Nullstellensatz*, o teorema de los ceros de Hilbert. Si bien es cierto que la demostración de Hilbert no utilizaba bases de Gröbner, esta metodología permite demostrar de forma sencilla este resultado, cuya comprensión antes estaba destinada únicamente a los estudiantes de posgrado.

La aplicación mostrada en el Capítulo 8, es quizás una de las más interesantes de la teoría. El encontrar métodos para demostrar teoremas de forma automatizada, es actualmente un tema muy investigado, y buena parte de los investigadores se dedican a los demostradores algebraicos, como los que utilizan bases de Gröbner. Aquí, es necesario exponer las dificultades y limitaciones que tienen las pruebas de los teoremas del hexagrama de Pascal y de Pappus, presentadas. Aunque está claro que se dio una demostración de los mismos en sus casos *genéricos*, únicamente se mencionó que estos son los casos en los que la *forma* de la configuración construida por el teorema no contenga degeneraciones (*i.e.* puntos distintos que colapsan en uno, rectas que se intersectan que se vuelven paralelas, etc.), sin embargo no fue posible dar un listado concreto de cuáles casos no fueron considerados por el algoritmo mostrado. Estos casos corresponden con las condiciones sobre los parámetros u_i que no pueden hacerse cero, durante el cálculo de la base de Gröbner; en otras palabras, son todos aquellos polinomios sobre los u_i que aparecen en algún denominador durante la ejecución del algoritmo de Buchberger. Si bien enlistar estos polinomios no es tarea fácil, lo más complicado es traducir de regreso estos polinomios a las configuraciones geométricas que los generan. A pesar de esta limitación, se pudo hallar una prueba poco usual para el teorema del hexagrama de Pascal, que toma en consideración desde un principio una cónica cualquiera, en lugar del camino clásico de probarlo para el caso de la circunferencia y luego extrapolar el resultado. A su vez se identificó la dificultad de determinar cuáles son los casos *degenerados* que no se demostraron con este acercamiento, por lo que queda como una pregunta abierta para futuras investigaciones.

10 BIBLIOGRAFÍA

Adams, W. W., y Loustaunau, P. (1994). *An introduction to Gröbner bases.* Providence, R.I: American Mathematical Society.

Arora, S. (2009). *Computational complexity a modern approach.* Cambridge New York: Cambridge University Press.

Bardet, M. (2005, November 2002). On the complexity of a Gröbner basis algorithm. En F. Chyzak (Ed.), *Algorithms seminar 2002–2004* (p. 85-92).

Becker, T. (1993). *Gröbner bases : a computational approach to commutative algebra.* New York: Springer-Verlag.

Borodzik, M., y Zoladek, H. (2002). The Pascal theorem and some its generalizations. *Topological Methods in Nonlinear Analysis*, *19*, 77-90.

Buchberger, B. (2006). Bruno buchberger's phd thesis 1965: An algorithm for finding the basis elements of the residue class ring of a zero dimensional polynomial ideal. *Journal of Symbolic Computation*, *41*(3-4), 475-511.

Buchberger, B. (2013, 18 de octubre). *Re: What is the advantage of Wu Wenjun's method over Gröbner basis method?* [Comentario de forum web]. http://www.researchgate.net/post/What_is_the_advantage_of_Wu_Wenjuns_method_over_Groebner_basis_method.

Cox, D. A., Little, J., y O'Shea, D. (2007). *Ideals, varieties, and algorithms* (3.ª ed.). Springer.

Coxeter, H. S. M. (1967). *Geometry revisited.* Washington, D.C: Mathematical Association of America.

Dasgupta, S., Papadimitriou, C. H., y Vazirani, U. V. (2008). *Algorithms.* Boston: McGraw-Hill Higher Education.

Dummit, D. (2004). *Abstract algebra.* Hoboken, NJ: John Wiley and Sons, Inc.

Ferro, A., y Gallo, G. (1988). Automated theorem proving in elementary geometry. *Le Matematiche*, *XLIII*(I), 195-224.

Fine, B. (1997). *The fundamental theorem of algebra.* New York: Springer.

Fröberg, R. (1997). *An introduction to Gröbner bases.* Chichester New York: Wiley.

Garcia, A., y Lequain, Y. (2003). *Elementos de álgebra.* Instituto de Matemáqtica Pura e Aplicada.

Herstein, I. N. (1975). *Topics in algebra.* New York: Wiley.

Kunz, E. (1985). *Introduction to commutative algebra and algebraic geometry.* Boston: Birkhäuser.

Mayr, E. W. (1997). Some complexity results for polynomial ideals. *Journal of Complexity*, *13*(3), 303-325.

Stein, W., y cols. (2015). Sage Mathematics Software (Versión 6.7) [Manual de software informático]. (`http://www.sagemath.org`)

Stillwell, J. (2010). *Mathematics and its history.* New York: Springer.

Stokes, T., y Bulmer, M. (2001). A complex change of variables for geometrical reasoning. En J. Gebert (Ed.), *Automated deduction in geometry : third international workshop, ADG 2000, Zurich, Switzerland, September 25-27, 2000 : revised papers* (p. 143-153). Berlin New York: Springer.

Stoutemyer, D. R. (2012). *Subtotal ordering – a pedagogically advantageous algorithm for computing total degree reverse lexicographic order.*

Weisstein, E. W. (s.f.). *Conic section. [en MathWorld—A Wolfram Web Resource].* Disponible en http://mathworld.wolfram.com/ConicSection.html

Wu, J. (1997). Mechanical geometry theorem proving based on groebner bases. *J. of Comput. Sci. & Technol.*, *12*(1), 10-16.

11 APÉNDICES

11.1. Implementación Hexagrama: Caso 1

```
#Funciones para traducir proposiciones geométricas a polinomios
#Colinealidad
colineales1(a1,b1,a2,b2,a3,b3)=(b2-b1)*(a3-a2)-(b3-b2)*(a2-a1)
colineales=lambda M,N,O: colineales1(M[0],M[1],N[0],N[1],O[0],O[1])
conica1(a,b,c,d,f,x,y)=a*x^2+b*y^2+c*x*y+d*x+f*y
conica=lambda a,b,c,d,e,P: conica1(a,b,c,d,e,P[0],P[1])

#Primero se encuentra la cónica que pasa por cinco puntos cualesquiera
#Anillos a utilizar
A.<u1,u2,u3,u4,u5,u6,u7>=PolynomialRing(QQ,7)
F=FractionField(A)
R.<b,c,d,x1,x2,x3,x4,x5,x6,x7,y>=PolynomialRing(F,11)


#Puntos
A=(0,0); B=(0,1); C=(u1,u2)
A1=(u3,u4); B1=(u5,u6); C1=(u7,x1)
X=(x2,x3); Y=(x4,x5); Z=(x6,x7)

#Hipótesis
#C, A1, B1 y C1 caen sobre la cónica
h1=conica(1,b,c,d,-b,C)
h2=conica(1,b,c,d,-b,A1)
h3=conica(1,b,c,d,-b,B1)
h4=conica(1,b,c,d,-b,C1)

#X = A1B ∩ AB1
h5=colineales(A,B1,X)
h6=colineales(A1,B,X)

#Y = A1C ∩ AC1
h7=colineales(A,C1,Y)
h8=colineales(A1,C,Y)

#Z = B1C ∩ BC1
h9=colineales(B,C1,Z)
```

```
h10=colineales(B1,C,Z)

#Conclusión
g=colineales(X,Y,Z)

#Ideal Radical
V=R.ideal([h1,h2,h3,h4,h5,h6,h7,h8,h9,h10,1-y*g])

#Cálculo de la base de Gröbner correspondiente
G=V.groebner_basis('toy:buchberger2')
G
```

11.2. Implementación Hexagrama: Caso 2

```
#Funciones para traducir proposiciones geométricas a polinomios
#Colinealidad
colineales1(a1,b1,a2,b2,a3,b3)=(b2-b1)*(a3-a2)-(b3-b2)*(a2-a1)
colineales=lambda M,N,O: colineales1(M[0],M[1],N[0],N[1],O[0],O[1])
conica1(a,b,c,d,e,x,y)=a*x^2+b*y^2+c*x*y+d*x+e*y
conica=lambda a,b,c,d,e,P: conica1(a,b,c,d,e,P[0],P[1])

#Primero se encuentra la cónica que pasa por cuatro puntos cualesquiera (al hacer
    a=0)
#Anillos a utilizar
A.<u1,u2,u3,u4,u5,u6>=PolynomialRing(QQ,6)
F=FractionField(A)
R.<c,d,x1,x2,x3,x4,x5,x6,x7,x8,y>=PolynomialRing(F,11)

#Puntos
A=(0,0); B=(0,1); C=(u1,u2)
A1=(u3,u4); B1=(u5,x1); C1=(u6,x2)
X=(x3,x4); Y=(x5,x6); Z=(x7,x8)

#Hipóteses
#C, A1, B1 y C1 caen sobre la cónica
h1=conica(0,1,c,d,-1,C)
h2=conica(0,1,c,d,-1,A1)
h3=conica(0,1,c,d,-1,B1)
h4=conica(0,1,c,d,-1,C1)

#X = A1B ∩ AB1
h5=colineales(A,B1,X)
h6=colineales(A1,B,X)

```

```
#Y = A1C ∩ AC
h7=colineales(A,C1,Y)
h8=colineales(A1,C,Y)

#Z = B1C ∩ BC1
h9=colineales(B,C1,Z)
h10=colineales(B1,C,Z)

#Conclusión
g=colineales(X,Y,Z)

#Ideal Radical
V=R.ideal([h1,h2,h3,h4,h5,h6,h7,h8,h9,h10,1-g*y])

#Cálculo de la base de Gröbner correspondiente
G=V.groebner_basis('toy:buchberger2')
G
```

11.3. Implementación Teorema de Pappus

```
#Funciones para traducir proposiciones geométricas a polinomios
#Colinealidad
colineales1(a1,b1,a2,b2,a3,b3)=(b2-b1)*(a3-a2)-(b3-b2)*(a2-a1)
colineales=lambda M,N,O: colineales1(M[0],M[1],N[0],N[1],O[0],O[1])

#Anillo en el que se trabajará
#Campo de fracciones de los parámetros libres
A.<u1,u2,u3,u4,u5,u6>=PolynomialRing(QQ,6)
F=FractionField(A)
#Anillo
R.<x1,x2,x3,x4,x5,x6,x7,y>=PolynomialRing(F,8)

#Definición de Puntos
A=(0,0); B=(0,1); C=(0,u1)
A1=(u2,u3); B1=(u4,u5); C1=(u6,x1)
X=(x2,x3); Y=(x4,x5); Z=(x6,x7)

#Hipótesis del teorema
#A,B y C son colineales por construcción

#A1, B1 y C1 son colineales
h1=colineales(A1,B1,C1)

#X = A1B ∩ AB1
```

```
25 h2=colineales(A1,B,X)
26 h3=colineales(A,B1,X)
27
28 #Y = A1C ∩ AC1
29 h4=colineales(A1,C,Y)
30 h5=colineales(A,C1,Y)
31
32 #Z = B1C ∩ BC1
33 h6=colineales(B1,C,Z)
34 h7=colineales(B,C1,Z)
35
36 #Conclusión
37 #X, Y y Z colineales
38 g=colineales(X,Y,Z)
39
40 #Ideal a utilizar
41 V=R.ideal([h1,h2,h3,h4,h5,h6,h7,1-g*y])
42
43 #Calcular base de Gröbner
44 G=V.groebner_basis('toy:buchberger2')
45 G
```