Facultad de Matemáticas
Departamento de Álgebra

Capítulo 1

Introducción

1.1. Repaso de polinomios

Operaremos siempre sobre el cuerpo complejo $\mathbb{C}$. Consideraremos siempre los anillos de polinomios $A = \mathbb{C}[x_1, \ldots, x_n]$, donde $n > 1$ es un entero, y $P = \mathbb{C}[y_0, y_1, \ldots, y_n]$.

Notas 1.1.1.– Recordemos que un polinomio se llama *homogéneo* (o una *forma*, por usar el lenguaje clásico) si todos sus monomios tienen el mismo grado. En general, un polinomio es una suma finita de formas.

1.1.1.1. Sea P_m el conjunto de las formas de grado m de P, junto con el cero. Evidentemente, P_m es un $\mathbb{C}$-espacio vectorial, y el conjunto de monomios de grado m en $(y_0, y_1 \ldots, y_n)$ es una base de él. El número de estos monomios es igual al de combinaciones con repetición de los elementos $(y_0, y_1 \ldots, y_n)$ tomados de m en m, con lo que

$$\dim(P_m) = \binom{n+m}{m}.$$

1.1.1.2. Para cada $m \in \mathbb{Z}_0$ hay una función lineal

$$\begin{array}{rcl} P_m & \longrightarrow & A \\ F(y_0, y_1 \ldots, y_n) & \longrightarrow & F^a = F(1, x_1, \ldots, x_n) \end{array}$$

llamada deshomogeneización. Por ejemplo, la deshomogeneización de

$$F(y_0, y_1, y_2) = y_0^2 + y_0 y_1 + y_0 y_2 + y_1^2 + y_1 y_2 + y_2^2$$

es

$$F^a = 1 + x_1 + x_2 + x_1^2 + x_1\,x_2 + x_2^2\,.$$

Por otra parte, hay una función

$$\begin{array}{rcl} A & \longrightarrow & P \\ f(x_1,\ldots,x_n) & \longrightarrow & f^h \end{array}$$

definida de la forma siguiente:

1. $0^h = 0$.
2. Si $0 \neq f(x_1,\ldots,x_n) = \sum_{i=0}^{m} f_i(x_1,\ldots,x_n)$ es la descomposición de f en suma de formas, con $f_m \neq 0$ (es decir, el grado de f es m), entonces

$$f^h = \sum_{i=0}^{m} y_0^{m-i} f_i(y_1,\ldots,y_n)$$

A esta función se la llama *homogeneización*. Nótese que, si f es ya homogéneo, su homogeneización es él mismo. Obsérvese, asimismo, que f^h nunca es divisible por y_0. Por ejemplo, si

$$f(x_1,x_2) = 1 + 2\,x_1 + 3\,x_2 + 4\,x_1^2 + 5\,x_2^2 + 6\,x_1^3\,,$$

la homogeneización de $f(x_1,x_2)$ es

$$f^h = y_0^3 + 2\,y_1\,y_0^2 + 3\,y_2\,y_0^2 + 4\,y_1^2\,y_0 + 5\,y_2^2\,y_0 + 6\,y_1^3\,.$$

1.1.1.3. La homogeneización no es un homomorfismo de anillos porque no es compatible con la suma (aunque sí con el producto). Por ejemplo, si $f(x_1,x_2)$ es el polinomio de arriba y $g(x_1,x_2) = 1 - 6\,x_1^3$, entonces

$$\begin{array}{rcl} g^h & = & y_0^3 - 6\,y_1^3 \\ f^h + g^h & = & 2\,y_0^3 + 2\,y_1\,y_0^2 + 3\,y_2\,y_0^2 + 4\,y_1^2\,y_0 + 5\,y_2^2\,y_0 \\ & = & y_0(2\,y_0^2 + 2\,y_1\,y_0 + 3\,y_2\,y_0 + 4\,y_1^2 + 5\,y_2^2) \\ f + g & = & 2 + 2\,x_1 + 3\,x_2 + 4\,x_1^2 + 5\,x_2^2 \\ (f+g)^h & = & 2\,y_0^2 + 2\,y_1\,y_0 + 3\,y_2\,y_0 + 4\,y_1^2 + 5\,y_2^2 \end{array}$$

es decir, $f^h + g^h = y_0\,(f+g)^h$.

1.1.1.4. Para ver que la homogenización es compatible con el producto, razonamos así. Tomemos dos polinomios no nulos $f, g \in \mathbb{C}[x_1, \ldots, x_n]$, y sean M_1, M_2 sus respectivos grados (grado de la forma de mayor grado); el grado de fg es $M_1 + M_2$. Para hallar fg se multiplica cada término de f por cada uno de g y luego se reducen términos semejantes en la suma resultante. La homogeneización de fg consiste, pues, en tomar cada término de grado m que aparece en fg y multiplicarlo por $y_0^{M_1+M_2-m}$. Para homogeneizar f y g por separado se multiplica cada término de grado $m_1 \leq M_1$ que aparezca en f por $y_0^{M_1-m_1}$ y cada término de grado $m_2 \leq M_2$ que aparezca en g por $y_0^{M_2-m_2}$. Luego multiplicamos todos estos términos homogeneizados y reducimos términos semejantes. Los términos anteriores de f y g darán lugar a uno de fg de grado $m_1 + m_2$ que, para homogeneizarlo, hay que mutiplicarlo por $y_0^{M_1+M_2-(m_1+m_2)}$. Esto prueba nuestro aserto.

1.1.1.5. Si $f \in \mathbb{C}[x_1, \ldots, x_n]$, está claro que $(f^h)^a = f$; vamos a probar que, si $F \in \mathbb{C}[y_0, y_1, \ldots, y_n]$ es homogéneo no divisible por y_0, entonces $(F^a)^h = F$. En efecto, para que fueran distintos, tendría que bajar de grado F al pasar a F^a. Pero esto no ocurre porque no es divisible por y_0 y no hay cancelaciones posibles, ya que F es homogéneo: dos monomios en $(y_1, \ldots, y_n)$ con los mismos exponentes tienen el mismo exponente de y_0.

Notas 1.1.2.– Vamos a dar unas propiedades cruciales de los polinomios.

1.1.2.1. Dado un polinomio $0 \neq F \in \mathbb{C}[y_0, y_1, \ldots, y_n]$ de grado m se verifica que F es homogéneo si y sólo si, si t es una indeterminada, entonces

$$F(ty_0, ty_1, \ldots, ty_n) = t^m F(y_0, y_1, \ldots, y_n)\,.$$

En otras palabras, F es homogéneo si y sólo si, para todo $t \in \mathbb{C}$, es una

$$F(ty_0, ty_1, \ldots, ty_n) = t^m F(y_0, y_1, \ldots, y_n)\,.$$

La demostración de este hecho es intuitiva.

1.1.2.2. (Teorema de Euler) Sea $0 \neq F \in \mathbb{C}[y_0, y_1, \ldots, y_n]$ un polinomio homogéneo de grado m; se verifica que

$$y_0 \frac{\partial F}{\partial y_0} + y_1 \frac{\partial F}{\partial y_1} + \cdots + y_n \frac{\partial F}{\partial y_n} = mF$$

En efecto, basta demostrarlo para un monomio. Supongamos que

$$F = y_0^{i_0} y_1^{i_1} \cdots y_n^{i_n}\,, \quad \sum_{j=0}^{n} i_j = m\,;$$

tenemos que

$$y_0\frac{\partial F}{\partial y_0} + y_1\frac{\partial F}{\partial y_1} + \cdots + y_n\frac{\partial F}{\partial y_n} =$$
$$i_0 y_0^{i_0} y_1^{i_1} \cdots y_n^{i_n} + i_1 y_0^{i_0} y_1^{i_1} \cdots y_n^{i_n} + \cdots + i_n y_0^{i_0} y_1^{i_1} \cdots y_n^{i_n} =$$
$$mF$$

1.1.2.3. Los anillos de polinomios son factoriales, es decir, todo polinomio de grado positivo se descompone en producto finito de irreducibles. Como la forma de mayor grado de un producto de polinomios es el producto de las formas de mayor grado de los factores, se deduce que un polinomio homogéneo factoriza en producto de polinomios homogéneos. Un resultado muy sencillo, pero esencial, es que una forma en dos variables se descompone en producto de formas lineales. En efecto, sea $F(y_0, y_1)$ una forma de grado m; se puede escribir siempre $F(y_0, y_1) = y_0^r F_1(y_0, y_1)$, donde $r \geq 0$, F_1 es de grado $s = m - r$ y F_1 no es divisible por y_0, luego tiene un término en y_1^s con un coeficiente $a \in \mathbb{C}$ distinto de cero. Así podemos escribir

$$F_1(y_0, y_1) = y_0^s \left(\frac{1}{y_0}\right)^s F_1(y_0, y_1) = y_0^s F_1\left(1, \frac{y_1}{y_0}\right).$$

Como este último es un polinomio en la única variable y_1/y_0 con coeficientes en $\mathbb{C}$, se puede descomponer en un producto de factores lineales, luego tenemos

$$F_1(y_0, y_1) = y_0^s \left(\frac{y_1}{y_0} - \alpha_1\right) \cdots \left(\frac{y_1}{y_0} - \alpha_s\right) = (y_1 - \alpha_1 y_0) \cdots (y_1 - \alpha_s y_0)$$

donde $\alpha_1, \ldots, \alpha_s \in \mathbb{C}$. Por tanto,

$$F(y_0, y_1) = y_0^r (y_1 - \alpha_1 y_0) \cdots (y_1 - \alpha_s y_0),$$

lo que prueba nuestro aserto.

Antes de continuar, hemos de hacer unas observaciones, que van a ser un punto sencillo, pero básico, en todo lo que seguirá.

Notas 1.1.3.– Para todo polinomio $0 \neq F \in \mathbb{C}[y_0, y_1, \ldots, y_n]$, existe un vector $\mathbf{u} \in \mathbb{C}^{n+1}$ tal que $F(\mathbf{u}) \neq 0$. Éste es el llamado *principio de identidad de polinomios*: un polinomio sobre un cuerpo infinito es cero si y sólo si es cero como función.

1.1.3.1. Se puede proyectivizar la situación. Consideremos una forma $0 \neq F(y_0, y_1 \ldots, y_n)$, de grado $m > 0$ y el espacio proyectivo $\mathbb{P}^n_{\mathbb{C}}$ de dimensión n sobre $\mathbb{C}$. Sea $(a_0 : a_1 : \ldots : a_n) \in \mathbb{P}^n_{\mathbb{C}}$; por 1.1.2.1, para todo $t \in \mathbb{C}\setminus\{0\}$, $F(ta_0, ta_1 \ldots, ta_n) = t^m F(a_0, a_1 \ldots, a_n)$. Por tanto, el hecho de que F se anule o no en un punto, no depende de la elección de un vector particular que represente al punto. Por ello tiene sentido hablar del conjunto $V = V(F)$ de los puntos de $\mathbb{P}^n_{\mathbb{C}}$ que anulan a F. Este conjunto se llama una *hipersuperficie proyectiva*. El principio de identidad nos dice que $\mathbb{P}^n_{\mathbb{C}}$ no es una hipersuperficie proyectiva.

1.1.3.2. Mediante un cambio de sistema de referencia en $\mathbb{P}^n_{\mathbb{C}}$, si es preciso, podemos arreglarnos para que V no pase por el punto $(0 : 0 : \ldots : 0 : 1)$. Un cambio de sistema de referencia, al ser lineal en las variables, lleva F sobre una forma G del mismo grado. El hecho de que $(0 : 0 : \ldots : 0 : 1) \notin V$ se traduce en que $G(y_0, y_1 \ldots, y_n)$ tiene un término en y_n^m con coeficiente un elemento de $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$. Este hecho es muy importante: por un cambio lineal homogéneo en las variables, podemos hacer que la trasformada de F tenga un término en y_n^m con coeficiente no nulo.

1.1.3.3. Utilizando la inmersión usual del espacio afín en el proyectivo, se pueden introducir las topologías clásicas, que nos ayudarán en nuestro razonamientos. Sabemos que $\mathbb{P}^n_{\mathbb{C}}$ es el conjunto de las rectas vectoriales de $\mathbb{C}^{n+1}$. Dado un vector $\mathbf{u} \in \mathbb{C}^{n+1}$, un polidisco de centro $\mathbf{u}$ y radio $\varepsilon \in \mathbb{R}_+$ es el conjunto

$$K(\mathbf{u}, \varepsilon) = \{\mathbf{z} \in \mathbb{C}^{n+1} \mid |z_i - u_i| < \varepsilon\,,\ \forall i = 0, 1, \ldots, n\}\,.$$

Los polidiscos determinan la topología usual sobre $\mathbb{C}^{n+1}$. Ahora bien, si $\mathbf{u} \neq \mathbf{0}$ y $K(\mathbf{u}, \varepsilon)$ es un polidisco con $\varepsilon < |\mathbf{u}|$ (lo que implica que $\mathbf{0} \notin K(\mathbf{u}, \varepsilon)$), el conjunto de rectas vectoriales determinadas por los vectores de $K(\mathbf{u}, \varepsilon)$ es, por definición un polidisco proyectivo del punto $\langle \mathbf{u} \rangle$ (la recta vectorial que contiene a $\mathbf{u}$). Los polidiscos proyectivos son la base de la topología clásica sobre $\mathbb{P}^n_{\mathbb{C}}$.

La inmersión clásica del espacio afín en el proyectivo funciona así: se elige un hiperplano vectorial H de $\mathbb{C}^{n+1}$, cuya proyectivización es un hiperplano proyectivo, como hiperplano del infinito. Las rectas vectoriales contenidas en H son los puntos del infinito; las no contenidas son los puntos afines. Se reconstruye la inmersión clásica $\mathbb{C}^n \hookrightarrow \mathbb{P}^n_{\mathbb{C}}$ tomando un hiperplano afín cualquiera $H' \parallel H$, $H' \neq H$; hay una biyección entre las rectas proyectivas no contenidas en H y los puntos de $H' = \mathbb{C}^n$. Ésta es la inmersión clásica de $\mathbb{C}^n$ en $\mathbb{P}^n_{\mathbb{C}}$. Los polidiscos proyectivos que no cortan a H en vectores distintos de $\mathbf{0}$ cortan a H' en abiertos de una base, que determina la misma topología que la de los polidiscos en $\mathbb{C}^n$.

Con estas nociones de topología queda ya claro que una hipersuperficie en $\mathbb{P}^n_{\mathbb{C}}$ es cerrada (su complementario es abierto: toda función continua, forma en

este caso, que no se anula en un punto tiene un polidisco centrado en el punto en el que no se anula).

1.2. Eliminación

Sea A un dominio factorial y sean

$$\begin{aligned} f(X) &= a_0X^n + a_1X^{n-1} + \cdots + a_{n-1}X + a_n\,,\ a_0 \neq 0 \\ g(X) &= b_0X^m + b_1X^{m-1} + \cdots + b_{m-1}X + b_m\,,\ b_0 \neq 0 \end{aligned}$$

dos polinomios en $A[X]$ de grado positivo.

Proposición 1.2.1.– *Las condiciones siguientes son equivalentes:*

1. *$f(X)$ y $g(X)$ tienen en $A[X]$ un máximo común divisor de grado positivo, esto es, tienen un divisor común no constante.*
2. *Existen dos polinomios $h_1(X)$, $h_2(X) \in A[X]$ no nulos que verifican:*
   a) $\operatorname{gr}(h_1) \leq m-1$, $\operatorname{gr}(h_2) \leq n-1$,
   b) $h_1(X)f(X) = h_2(X)g(X)$

Demostración:

1) $\Longrightarrow$ 2)

Sea $h(X)$ un factor común no constante de $f(X)$ y $g(X)$; entonces se puede escribir:

$$\begin{aligned} f(X) &= h_2(X)h(X) \\ g(X) &= h_1(X)h(X)\,, \end{aligned}$$

donde $\operatorname{gr}(h_2) \leq n-1$ y $\operatorname{gr}(h_1) \leq m-1$. De esta manera es:

$$h_1(X)f(X) = h_1(X)h_2(X)h(X) = h_2(X)g(X)\,.$$

2) $\Longrightarrow$ 1)

De $h_1(X)f(X) = h_2(X)g(X)$ se deduce que no todo factor irreducible de grado positivo de $f(X)$ puede dividor a $h_2(X)$ ya que $\operatorname{gr}(h_2) < \operatorname{gr}(f)$. Por tanto, alguno de ellos divide a $g(X)$, lo que prueba la propòsición. □

Facultad de Matemáticas
Departamento de Álgebra

Designemos por K al cuerpo de fracciones de A y sean S_1, S_2 las matrices siguientes:

1. S_1 es la matriz $(m+n)\times m$ dada por

$$S_1 = \begin{pmatrix} a_0 & & & \\ a_1 & a_0 & & \\ \vdots & \vdots & \ddots & \\ a_{m-1} & a_{m-2} & \cdots & a_0 \\ a_m & a_{m-1} & \cdots & a_1 \\ \vdots & \vdots & & \vdots \\ a_n & a_{n-1} & \cdots & a_{n-m+1} \\ & a_n & \cdots & a_{n-m+2} \\ & & \ddots & \vdots \\ & & & a_n \end{pmatrix}$$

   donde las posiciones no ocupadas contienen ceros.

2. S_2 es la matriz $(m+n)\times n$ dada por

$$S_2 = \begin{pmatrix} b_0 & & & \\ b_1 & b_0 & & \\ \vdots & \vdots & \ddots & \\ b_{n-1} & b_{n-2} & \cdots & b_0 \\ b_n & b_{n-1} & \cdots & b_1 \\ \vdots & \vdots & & \vdots \\ b_m & b_{m-1} & \cdots & b_{m-n+1} \\ & b_m & \cdots & b_{m-n+2} \\ & & \ddots & \vdots \\ & & & b_m \end{pmatrix}$$

   donde las posiciones no ocupadas contienen ceros.

3. $S = (S_1|S_2)$ es la matriz que se obtiene poniendo juntas, de izquierda a derecha, las matrices S_1 y S_2, por este orden.

Por ejemplo, si $n = 4$ y $m = 3$ se obtiene:

$$\begin{pmatrix} a_0 & 0 & 0 & b_0 & 0 & 0 & 0 \\ a_1 & a_0 & 0 & b_1 & b_0 & 0 & 0 \\ a_2 & a_1 & a_0 & b_2 & b_1 & b_0 & 0 \\ a_3 & a_2 & a_1 & b_3 & b_2 & b_1 & b_0 \\ a_4 & a_3 & a_2 & 0 & b_3 & b_2 & b_1 \\ 0 & a_4 & a_3 & 0 & 0 & b_3 & b_2 \\ 0 & 0 & a_4 & 0 & 0 & 0 & b_3 \end{pmatrix}.$$

Ejemplo 1.2.2.– En Maple se calcula la traspuesta de esta matriz con la orden `linalg[sylvester]`, porque la matriz se llama *matriz de Sylvester*.

Proposición 1.2.3.– *Las condiciones siguientes son equivalentes:*

1. $f(X)$ *y* $g(X)$ *tienen un divisor común no constante.*
2. $\det(S) = 0$

Demostración: Por la proposición 1.2.1 se verifica que $f(X)$ y $g(X)$ tienen un factor común no constante si y sólo si existen dos polinomios no nulos de $A[X]$,

$$\begin{aligned} h_1(X) &= c_0X^{m-1} + c_1X^{m-2} + \cdots + c_{m-1} \\ h_2(X) &= d_0X^{n-1} + d_1X^{n-2} + \cdots + d_{n-1} \end{aligned}$$

tales que

$$\begin{aligned} &(a_0X^n + a_1X^{n-1} + \cdots + a_n)(c_0X^{m-1} + c_1X^{m-2} + \cdots + c_{m-1}) \\ = \\ &(b_0X^m + b_1X^{m-1} + \cdots + b_m)(d_0X^{n-1} + d_1X^{n-2} + \cdots + d_{n-1}) \end{aligned}$$

o, lo que es lo mismo, si el sistema

$$\left\{\begin{array}{rcl} a_0c_0 & = & b_0d_0 \\ a_1c_0 + a_0c_1 & = & b_1d_0 + b_0d_1 \\ & \vdots & \\ a_nc_{m-2} + a_{n-1}c_{m-1} & = & b_md_{n-2} + b_{m-1}d_{n-1} \\ a_nc_{m-1} & = & b_md_{n-1} \end{array}\right.$$

tiene solución en A distinta de la $c_i = 0$, $d_j = 0$, para todo $i = 0, 1, \ldots, m-1$, $j = 0, 1, \ldots, n-1$. Ese sistema tiene solución no trivial en A si y sólo si la tiene en K, por ser homogéneo, y este hecho es equivalente a que el determinante de la matriz de los coeficientes es cero. Pero, salvo signo, ese determinante es igual al $\det(S)$, de donde la conclusión. □

Definición 1.2.4.– *El elemento $R = \det(S) \in A$ será designado con el nombre de resultante de $f(X)$ y $g(X)$. A la resultante de un polinomio $f(X)$ y su derivada $f'(X)$ se le llamará discriminante de $f(X)$.*

Nota 1.2.5.– De la observación de la matriz de Sylvester se deduce que R es un polinomo homogéneo de grado m en los coeficientes a_r, $r = 0, 1, \ldots, n$, y a la vez homogéneo de grado n en los coeficientes b_s, $s = 0, 1, \ldots, n$.

Ejemplo 1.2.6.– En Maple se pueden hallar resultantes y discriminantes sin pasar por la matriz de Sylvester, con las órdenes sistema `resultant` y `discrim`. Hay que tener en cuenta que Maple no da como discriminante exactamente la resultante entre un polinomio y su derivada, sino algo distinto que explicamos. Sea, como de costumbre,

$$f(X) = a_0X^n + a_1X^{n-1} + \cdots + a_{n-1}X + a_n\,,\ a_0 \neq 0$$

y R la resultante entre $f(X)$ y df/dX. Lo que Maple da como discriminante es:

$$(-1)^{n(n-1)/2}\frac{R}{a_0}\,.$$

Corolario 1.2.7.– *Supongamos que k es un cuerpo y sea $f(X) \in k[X]$ un polinomio de grado positivo que tiene todas sus raíces en k. Si D es el discriminante de $f(X)$ se verifica que $f(X)$ tiene todas sus raíces simples (resp. tiene alguna raíz múltiple) si y sólo si $D \neq 0$ (resp. $D = 0$).*

Demostración: Supongamos que $f(X)$ tiene una raíz múltiple $a \in K$. Entonces es $(X-a)|f(X)$ por el teorema de Ruffini. Pero, al ser múltiple la raíz, también anula a la derivada, luego $(X-a)|f'(X)$. Así, por la proposición 1.2.3, es $D = 0$.

Recíprocamente, supongamos que $D = 0$; entonces $f(X)$ y $f'(X)$ tienen un factor común no constante en $k[X]$ por la proposición 1.2.3. Como $f(X)$ factoriza en $k[X]$ en producto de factores lineales, deben tener $f(X)$ y $f'(X)$ un factor común lineal, es decir, una raíz común. Así, esa raíz es múltiple para $f(X)$. □

Volviendo a la situación del inicio de esta sección tenemos

Proposición 1.2.8.– *Sea I el ideal de $A[X]$ engendrado por $f(X)$ y $g(X)$. Se verifica que la resultante R pertenece a I.*

Demostración: Si $R = 0$ no hay nada que probar; supongamos, pues, $R \neq 0$ y consideremos las expresiones

$$\begin{aligned}
X^{m-1}f(X) &= a_0X^{n+m-1} + a_1X^{n+m-2} + \cdots + a_nX^{m-1}\\
X^{m-2}f(X) &= a_0X^{n+m-2} + \cdots + a_{n-1}X^{m-1} + a_nX^{m-2}\\
\vdots &\quad \vdots\\
f(X) &= a_0X^n + \cdots + a_n\\
X^{n-1}g(X) &= b_0X^{n+m-1} + b_1X^{n+m-2} + \cdots + b_mX^{n-1}\\
X^{n-2}g(X) &= b_0X^{n+m-2} + \cdots + b_{m-1}X^{n-1} + b_mX^{n-2}\\
\vdots &\quad \vdots\\
g(X) &= b_0X^m + \cdots + b_m
\end{aligned}$$

como un sistema de ecuaciones, con coeficientes en el cuerpo de fracciones de $A[X]$, en las incógnitas $\{X^{n+m-1}, X^{n+m-2}, \ldots, X, 1\}$ que están en los miembros de la derecha (los términos independientes son los miembros de la izquierda). La matriz de los coeficientes del sistema es la traspuesta de S, luego su determinante es la resultante $R \neq 0$. La regla de Cramer de resolución de sistemas de ecuaciones lineales nos dice entonces que $1 = T/R$, donde T es el determinante de la matriz que se obtiene al sustituir la última columna de la traspuesta de S por la de los términos independientes anteriores. Así $R = T$, y calculando T por el desarrollo por los adjuntos de la última columna tenemos una expresión del tipo $R = \alpha(X)f(X) + \beta(X)g(X)$, con $\alpha(X), \beta(X) \in A[X]$. Esto prueba la proposición. □

Vamos a dedicar el resto se la sección a demostrar que se puede expresar la resultante de dos polinomios con coeficientes en un cuerpo de una forma bastante simple, en función de las raíces de estos polinomios.

Nota 1.2.9.– Sea k un cuerpo y $f(X) \in k[X]$,

$$f(X) = a_0X^n + a_1X^{n-1} + \cdots + a_{n-1}X + a_n$$

un polinomio de grado positivo que tiene todas sus raíces en k. Sean estas raíces $\{x_1, \ldots, x_n\}$ y, para cada entero i, $1 \leq i \leq n$, designemos por $\sigma_i(f)$ a la suma de todos los productos distintos de i factores, con índices distintos,

que se puedan formar con los elementos $\{x_1, \ldots, x_n\}$. Por ejemplo, si $n = 4$, tenemos:

$$\begin{aligned}
\sigma_1(f) &= x_1 + x_2 + x_3 + x_4 \\
\sigma_2(f) &= x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4 \\
\sigma_3(f) &= x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4 \\
\sigma_4(f) &= x_1x_2x_3x_4
\end{aligned}$$

Nótese que, en cada $\sigma_i(f)$ el conjunto de los índices de las x_i es justamente el de las combinaciones ordinarias de $\{1, \ldots, n\}$ tomados de i en i, es decir,

$$\sigma_i(f) = \sum_{1 \le r_1 < \ldots < r_i \le n} x_{r_1} \cdots x_{r_i}\,.$$

Puesto que

$$a_0X^n + a_1X^{n-1} + \cdots + a_{n-1}X + a_n = a_0(X - x_1)(X - x_2)\cdots(X - x_n)\,,$$

se deduce que

$$\begin{aligned}
a_n &= (-1)^n a_0 \sigma_n(f) \\
a_{n-1} &= (-1)^{n-1} a_0 \sigma_{n-1}(f) \\
\vdots &\quad \vdots \\
a_1 &= -a_0 \sigma_1(f)\,,
\end{aligned}$$

y a estas igualdades se las llamará las *relaciones de Cardano* para el polinomio f.

Nótese que cada $\sigma_i(f)$ verifica la propiedad de que, permutando de cualquier forma los índices de las x_j, no varía $\sigma_i(f)$. Un polinomio, en un número cualquiera de argumentos, que verifique esta condición, se llama una *función simétrica*. Así, los coeficientes de $f(X)$ son funciones simétricas de las raíces. Nótese que la suma, la diferencia y el producto de funciones simétricas es simétrica.

Para razonar con resultantes conviene escribir todo en función de variables. Esto permite luego particularizar para un cuerpo cualquiera. Cuando operemos con una colección cualquiera de variables $\mathbf{z} = (z_1, \ldots, z_p)$, si $\mathbf{i} = (i_1, \ldots, i_p) \in \mathbb{Z}_0^p$, designaremos por $\mathbf{z}^{\mathbf{i}}$ al monomio $z_1^{i_1} \cdots z_p^{i_p}$, y por $\mathrm{gr}\,(\mathbf{z}^{\mathbf{i}})$ a su grado, que es

$$\mathrm{gr}\,(\mathbf{z}^{\mathbf{i}}) = \sum_{l=1}^{p} i_l\,.$$

El orden lexicográfico de monomios siempre se referirá a

$$z_1 >_{\text{lex}} z_2 >_{\text{lex}} \cdots >_{\text{lex}} z_p\,.$$

Para elementos $\alpha = (\alpha_1, \ldots, \alpha_p) \in \mathbb{Z}_0^p$ consideraremos el mismo grado

$$\text{gr}\,(\alpha) = \sum_{l=1}^{p} \alpha_l$$

y el mismo orden lexicográfico, es decir,

$$\alpha = (\alpha_1, \ldots, \alpha_p) >_{\text{lex}} (\beta_1, \ldots, \beta_p) = \beta$$

si y sólo si, comenzando por la izquierda, la primera componente de α que difiere de la correspondiente de β es mayor.

A partir de ahora razonaremos bajo los presupuestos que describimos. Tomamos un cuerpo base k y variables $a_0, b_0, \mathbf{x} = (x_1, \ldots, x_n)$, e $\mathbf{y} = (y_1, \ldots, y_m)$. Consideramos el dominio factorial $A = k[a_0, b_0, \mathbf{x}, \mathbf{y}]$ y los polinomios siguientes en $A[X]$, donde X es una nueva variable:

$$\begin{aligned} f(X) &= a_0(X - x_1)\cdots(X - x_n)\,, \\ g(X) &= b_0(X - y_1)\cdots(X - y_m)\,. \end{aligned}$$

Por otra parte, consideramos los elementos

$$\begin{aligned} \sigma_i(\mathbf{x}) &= \sum_{1 \le r_1 < \ldots < r_i \le n} x_{r_1} \cdots x_{r_i} \\ \sigma_j(\mathbf{y}) &= \sum_{1 \le s_1 < \ldots < s_j \le m} y_{s_1} \cdots y_{s_j} \end{aligned}$$

y ponemos

$$\begin{aligned} a_i &= (-1)^i a_0 \sigma_i(\mathbf{x}) \\ b_j &= (-1)^j b_0 \sigma_j(\mathbf{y}) \end{aligned}$$

donde $i = 1, \ldots, n$, $j = 1, \ldots, m$. A las $\sigma_i(\mathbf{x})$ se les llama las *funciones simétricas elementales* en los argumentos $\mathbf{x}$.

Teorema 1.2.10.– *Los elementos*

$$\{\sigma_i(\mathbf{x}), \sigma_j(\mathbf{y}) \mid i = 1, \ldots, n,\ j = 1, \ldots, m\}$$

son algebraicamente independientes sobre k, es decir, son variables sobre $k[a_0, b_0]$.

La demostración de este teorema, requiere un resultado previo, del que se deduce fácilmente.

Lema 1.2.11.– *Sea A un dominio, $\mathbf{z} = (z_1, \ldots, z_p)$ variables sobre A, y sea, para cada $i = 1, \ldots, p$, $\sigma_i(\mathbf{z})$ la correspondiente función simétrica elemental. Entonces, los elementos $\{\sigma_1(\mathbf{z}), \ldots, \sigma_p(\mathbf{z})\}$ son algebraicamente independientes sobre A.*

Demostración: Nótese que $\sigma_i(\mathbf{z})$ es homogéneo de grado i, y su término mayor para el orden lexicográfico es $z_1 z_2 \cdots z_i$.

Consideremos la función $\varphi : \mathbb{Z}_0^p \to \mathbb{Z}_0^p$ definida por

$$\varphi(\mathbf{i}) = \varphi(i_1, \ldots, i_p) = \left(\sum_{l=1}^{p} i_l, \sum_{l=2}^{p} i_l, \ldots, i_p \right) .$$

Por ejemplo, si $p = 4$,

$$\varphi(i_1, i_2, i_3, i_4) = (i_1 + i_2 + i_3 + i_4, i_2 + i_3 + i_4, i_3 + i_4, i_4) .$$

Es trivial ver que esta función es inyectiva porque, fijado un elemento $\mathbf{j} \in \varphi(\mathbb{Z}_0^p)$, el conjunto de los elementos $\mathbf{i} \in \mathbb{Z}_0^p$ que van sobre él está determinado por el sistema lineal

$$\begin{array}{ccccccccc} i_1 & + & i_2 & + & \cdots & + & i_p & = & j_1 \\ & & i_2 & + & \cdots & + & i_p & = & j_2 \\ & & & & & \ddots & & & \\ & & & & & & i_p & = & j_p \end{array}$$

que tiene solución única. Nótese que ese mismo sistema lineal nos indica que

$$\varphi(\mathbb{Z}_0^p) = \{(i_1, \ldots, i_p) \mid i_1 \geq i_2 \geq \cdots \geq i_p\} .$$

Hecha esta observación preliminar, vamos con la demostración del lema en sí. Introduzcamos una nueva serie de variables, $\mathbf{u} = (u_1, \ldots, u_p)$; se trata de probar que, si $0 \neq h(\mathbf{u}) \in A[\mathbf{u}]$, entonces $h(\sigma(\mathbf{z})) \neq 0$, donde

$$\sigma(\mathbf{z}) = (\sigma_1(\mathbf{z}), \ldots, \sigma_p(\mathbf{z})) .$$

Designemos por $\omega : k[\mathbf{u}] \to k[\mathbf{z}]$ al homomorfismo de sustitución definido por:

$$\omega(u_i) = \sigma_i(\mathbf{z}) , \quad i = 1, \ldots, p .$$

Para completar nuestra demostración introducimos en $A[\mathbf{u}]$ lo que llamamos el *grado simétrico*, $\mathrm{gr}_s(-)$ en la forma siguiente: si $\mathbf{u}^{\mathbf{i}}$ es un monomio, ponemos $\mathrm{gr}_s(\mathbf{u}^{\mathbf{i}}) = \mathrm{gr}(\varphi(\mathbf{i}))$ (esta última vez el grado ordinario).

Descomponemos $h(\mathbf{u})$ en suma de formas respecto del grado simétrico, y consideramos la forma $h_m(\mathbf{u})$ del mayor grado simétrico m. Para cada monomio $\mathbf{u}^{\mathbf{i}}$ de grado simétrico m que aparece en $h_m(\mathbf{u})$, el término mayor para el orden lexicográfico en $\sigma(\mathbf{z})^{\mathbf{i}}$ es $\mathbf{u}^{\varphi(\mathbf{i})}$. Como φ es inyectiva, el mayor de ellos no puede cancelar. Esto prueba nuestro lema. □

Sin mucho esfuerzo adicional se puede dar el teorema de las funciones simétricas. Seguimos con $A[\mathbf{u}]$ y consideremos el grupo simétrico S_p actuando sobre los números $\{1,\dots,p\}$. Es decir, este grupo es el de las biyecciones de ese conjunto de números. Se puede hacer actuar S_p sobre los monomios de $A[\mathbf{u}]$ poniendo, para $\tau \in S_p$,

$$\tau(\mathbf{u}^{\mathbf{i}}) = \tau(u_1^{i_1}\cdots u_p^{i_p}) = u_1^{\tau(i_1)}\cdots u_p^{\tau(i_p)} = \mathbf{u}^{\tau(\mathbf{i})}\,.$$

De manera análoga se puede hacer actuar τ sobre todo $A[\mathbf{u}]$ poniendo, para $\tau \in S_p$

$$\tau\left(\sum_{\text{finita}} a_{\mathbf{i}}\mathbf{u}^{\mathbf{i}}\right) = \sum_{\text{finita}} a_{\mathbf{i}}\mathbf{u}^{\tau(\mathbf{i})}\,.$$

Un polinomio $f(\mathbf{u}) \in A[\mathbf{u}]$ se llama *simétrico* si

$$\tau(f(\mathbf{u})) = f(\mathbf{u})\,, \quad \forall \tau \in S_p\,.$$

El resultado importante es el siguiente

Teorema 1.2.12.– *Todo polinomio simétrico $f \in A[\mathbf{z}]$ se puede escribir (de manera única por el lema 1.2.11) como un polinomio en las funciones simétricas elementales, es decir*

$$f = g(\sigma(\mathbf{z}))\,, \quad g(\mathbf{u}) \in A[\mathbf{u}]\,.$$

Si f es homogéneo de grado p, entonces g es homogéneo respecto del grado simétrico, y su grado simétrico es p.

Demostración: Consideremos sobre $A[\mathbf{z}]$ el orden monomial `tdeg` que, como se recuerda, se define así:

$$\mathbf{z}^{\mathbf{i}} < \mathbf{z}^{\mathbf{j}} \iff \begin{cases} \sum_{l=1}^{p} i_l < \sum_{l=1}^{p} j_l & \text{ó} \\ \sum_{l=1}^{p} i_l = \sum_{l=1}^{p} j_l & \text{y} \quad \mathbf{z}^{\mathbf{i}} <_{\text{lex}} \mathbf{z}^{\mathbf{j}} \end{cases}$$

donde, para el orden lexicográfico, se cuenta $z_1 > \cdots > z_p$.

La demostración se hace por recurrencia sobre el término líder $t(f)$ para `tdeg`. Si $t(f) = c \in \mathbb{C}$, el resultado es claro. Supongamos, pues, que el término líder $t(f) = c\mathbf{z}^{\mathbf{a}}$ con $c \in \mathbb{C} \setminus \{0\}$ y $\mathbf{a} \neq 0$. Como, con este término, tienen que parecer todos los de la forma $c\mathbf{z}^{\mathbf{b}}$, donde $\mathbf{b}$ es cualquier permutación de $\mathbf{a}$, debe ser necesariamente $a_1 \geq \cdots \geq a_p$, con lo que $\mathbf{a} \in \varphi(\mathbb{Z}_0)$. Sea $\mathbf{d} \in \mathbb{Z}_0^p$ tal que $\varphi(\mathbf{d}) = \mathbf{a}$; el polinomio $f - c\sigma(\mathbf{z})^{\mathbf{d}}$ es simétrico y su término líder es menor que el anterior. Esto termina nuestra demostración.

□

Nota 1.2.13.– La demostración del teorema 1.2.10 es ahora trivial. Basta tomar $A_1 = k[a_0, b_0]$. Entonces, $\mathbf{x} = (x_1, \ldots, x_n)$ son algebraicamente independientes sobre A_1, luego lo son $(\sigma_1(\mathbf{x}), \ldots, \sigma_n(\mathbf{x}))$. Ahora bien,

$$A_1[\sigma_1(\mathbf{x}), \ldots, \sigma_n(\mathbf{x})] \subset A_1[x_1, \ldots, x_n] = A$$

y $(y_1, \ldots, y_m)$ son algebraicamente independientes sobre A. Por tanto, $(\sigma_1(\mathbf{y}), \ldots, \sigma_m(\mathbf{x}))$ son algebraicamente independientes sobre A, lo que prueba el teorema.

Seguimos en la situación y con los presupuestos de la página 12.

Notas 1.2.14.–

1.2.14.1. Puesto que la resultante R de $f(X)$ y $g(X)$ es un polinomio en los coeficientes de ambos, las relaciones de Cardano (nota 1.2.9) nos dicen que $R \in k[a_0, b_0, \mathbf{x}, \mathbf{y}]$, simétrico en $\mathbf{x}$ y en $\mathbf{y}$ por separado.

1.2.14.2. La división polinómica en varias variables permite escribir

$$\begin{aligned} R(a_0, b_0, \mathbf{x}, \mathbf{y}) &= Q_{i,j}(a_0, b_0, \mathbf{x}, \mathbf{y})(x_i - y_j) \\ &\quad + r_{i,j}(a_0, b_0, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n, \mathbf{y}) \end{aligned}$$

Puesto que $R = 0$ al hacer $x_i = y_j$, y puesto que $r_{i,j}$ no depende de x_i, debe ser idénticamente nulo, con lo que $x_i - y_j$ divide a R. Como todos los polinomios $x_i - y_j$ son primos entre sí, si tienen algún índice distinto, R debe ser divisible por el producto

$$R' = \prod_{i=1}^{n} \prod_{j=1}^{m} (x_i - y_j)\,.$$

1.2.14.3. Tenemos que, considerando R' como un polinomio en $\mathbf{x}$ con coeficientes en $A = k[\mathbf{y}]$, es simétrico, y su término mayor para `tdeg` es

$$x_1^m \cdots x_n^m\,.$$

Así, $R' = R_1(\sigma(\mathbf{x}))$, donde aparece necesariamente el término $\sigma_n(\mathbf{x})^m$. Por las relaciones de Cardano, $R' = R_1(a_1/a_0, \ldots, a_n/a_0)$ es un polinomio de grado m. Así, $a_0^m R'$ es un polinomio homogéneo de grado m en $(a_0, a_1 \ldots, a_n)$.

1.2.14.4. Operando de la misma forma con $a_0^m R'$ vemos que $a_0^m b_0^n R'$ es un polinomio homogéneo de grado $n+m$ en $(a_0, a_1 \ldots, a_n), (b_0, b_1 \ldots, b_m)$ y que $R'|R$ (la resultante), que es también homogéneo de grado $m+n$.

1.2.14.5. En resumen, el desohomogeneizado R' de $a_0^m b_0^n R'$ respecto de $a_0^m b_0^m$ debe dividir al deshomogeneizado $R/(a_0^m b_0^n)$ de la resultante. El término mayor para el orden lexicográfico de $a_0^m b_0^n R'$ respecto de $(a_0, a_1, \ldots a_m, b_0, b_1, \ldots, b_n)$ es $a_0^m b_m^n$, que es el mismo que el de la resultante. Esto prueba la primera parte del teorema que enunciamos a continuación.

Teorema 1.2.15.– *Se verifica que*

$$R = a_0^m b_0^n \prod_{i=1}^{n} \prod_{j=1}^{m} (x_i - y_j) = a_0^m \prod_{i=1}^{n} g(x_i) = (-1)^{mn} b_0^n \prod_{j=1}^{m} f(y_j)\,.$$

Demostración: Restan por demostrar las igualdades segunda y tercera. Puesto que

$$g(X) = b_0 \prod_{j=1}^{m} (X - y_j)\,,$$

es

$$\prod_{i=1}^{n} g(x_i) = b_0^n \prod_{i=1}^{n} \prod_{j=1}^{m} (x_i - y_j)\,,$$

y así

$$R = a_0^m \prod_{i=1}^{n} g(x_i)\,.$$

Puesto que

$$f(X) = a_0 \prod_{i=1}^{n} (X - x_i)\,,$$

es

$$\prod_{j=1}^{m} f(y_j) = a_0^m \prod_{j=1}^{m} \prod_{j=1}^{m} (y_j - x_i) = (-1)^{mn} a_0^m \prod_{j=1}^{m} \prod_{j=1}^{m} (x_i - y_j)\,,$$

y así

$$R = (-1)^{mn} b_0^n \prod_{j=1}^{m} f(y_j)\,.$$

□

Teorema 1.2.16.– *El discriminante de $f(X)$ es*

$$\Delta(f) = (-1)^{n(n-1)/2} a_0^{2n-1} \prod_{1 \le i < j \le n} (x_i - x_j)^2 .$$

Demostración: Tenemos que, como

$$f(X) = a_0(X - x_1) \cdots (X - x_n)$$

es

$$f'(X) = \frac{df}{dX} = a_0 \sum_{i=1}^{n} (X - x_1) \cdots (X - x_{i-1})(X - x_{i+1}) \cdots (X - x_n) ,$$

que es un polinomio en X de grado $m - 1$. Por el teorema 1.2.15,

$$\begin{aligned}
\Delta(f) &= a_0^{n-1} \prod_{i=1}^{n} f'(x_i) \\
&= a_0^{2n-1} \prod_{i=1}^{n} (x_i - x_1) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_n) \\
&= a_0^{2n-1} \prod_{i=1}^{n} (-1)^{i-1} (x_1 - x_i) \cdots (x_{i-1} - x_i)(x_i - x_{i+1}) \cdots (x_i - x_n) \\
&= a_0^{2n-1} (-1)^{0+1+\cdots+(n-1)} \prod_{1 \le i < j \le n} (x_i - x_j)^2 \\
&= a_0^{2n-1} (-1)^{n(n-1)/2} \prod_{1 \le i < j \le n} (x_i - x_j)^2
\end{aligned}$$

□

Nota 1.2.17.– Nótese que las fórmulas obtenidas para resultantes y discriminantes considerando el caso de incógnitas valen para cualquier cuerpo, sustituyendo las incógnitas por los elementos del cuerpo. Es el *principio de especialización*. Por otro lado, de ambas fórmulas reultan patentes los dos hechos anunciados anteriormente:

1. Que dos polinomios tienen una raíz común si y sólo si su resultante es cero.
2. Que un polinomio tiene una raíz múltiple si y sólo si su discriminante es cero.