

Anillos de Polinomios

Eugenio Miranda Palacios

5. Polinomios

5.1. Definiciones y primeras propiedades

Sea A un anillo conmutativo.

Definición 5.1. El conjunto de polinomios en la indeterminada X con coeficientes en A es el conjunto de todas las sumas formales finitas

$$f = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0$$

Este conjunto se representa por $A[X]$.

Obsérvese que X no es una variable. Es un elemento nuevo, indeterminado que no representa a ningún elemento de A (Al final de la edad media y en el renacimiento le llamaban “la cosa”, y los que manipulaban la cosa ,i.e. los algebristas, se llamaban “cosistas”).

En el conjunto de polinomios definimos una suma y un producto: Sean

$$\begin{aligned} f &= a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0 \\ g &= b_m X^m + b_{m-1} X^{m-1} + \cdots + b_1 X + b_0 \end{aligned}$$

dos polinomios. Supongamos que $m \leq n$, Tomamos $b_i = 0$ para todo $n \geq i > m$. Con este convenio definimos

$$\begin{aligned} f + g &= (a_n + b_n) X^n + \cdots + (a_1 + b_1) X + (a_0 + b_0). \\ fg &= a_n b_n X^{n+m} + (a_n b_{m-1} + a_{n-1} b_m) X^{n+m-1} + \\ &\quad \cdots + (a_1 b_0 + a_0 b_1) X + a_0 b_0 \end{aligned}$$

Teorema 5.2. El conjunto $A[X]$ con las dos operaciones definidas forma un anillo conmutativo que se llama anillo de polinomios en X con coeficientes en A .

Lema 5.3. La aplicación $\lambda : A \rightarrow A[X]$ definida por $\lambda(a) = a$ es un monomorfismo de anillos.

Normalmente se identifica cada elemento $a \in A$ con el polinomio $\lambda(a) \in A[X]$, con lo que A es un subanillo de $A[X]$.

Definición 5.4. Para un polinomio $f = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0 \neq 0$ el mayor índice n tal que $a_n \neq 0$ se llama *grado de f* y se representa por $gr(f)$. Si $f = 0$ definimos $gr(f) = -\infty$.

Cada uno de los sumandos $a_i X^i$ se llama *monomio o término (de grado i)* del polinomio f .

El término no nulo de mayor grado se llama *término líder*. El coeficiente $a_n \neq 0$ del término líder se llama *coeficiente líder* y el término de grado cero a_0 se llama *término constante*.

Un polinomio $f = X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0$ cuyo coeficiente líder vale 1 se llama *polinomio mónico*.

Un polinomio f se llama *constante* si $gr(f) \leq 0$, es decir, cuando $f \in \text{Im}(\lambda)$.

Teorema 5.5. Para cualquier anillo conmutativo A y cualesquiera polinomios $f, g \in A[X]$ se verifica

$$\begin{aligned} gr(f + g) &\leq \max(gr(f), gr(g)) \\ gr(fg) &\leq gr(f) + gr(g) \end{aligned}$$

Si A es un dominio de integridad se verifica

$$gr(fg) = gr(f) + gr(g)$$

Corolario 5.6. El anillo conmutativo A es un dominio de integridad si y sólo si $A[X]$ es un dominio de integridad.

En cualquier dominio de integridad es importante determinar el grupo de unidades y los elementos irreducible y primos, para poder estudiar sus propiedades de divisibilidad. En este sentido los primeros resultados son los siguientes:

Proposición 5.7. 1. Sea A un dominio de integridad. Los elementos invertibles de $A[X]$ son exactamente los invertibles de A

2. Todo polinomio $X - a \in A[X]$ es irreducible.

La propiedad mas importante de un anillo de polinomios es la siguiente:

Teorema 5.8 (Propiedad universal del anillo de polinomios). Sea A un anillo conmutativo, $\lambda : A \rightarrow A[X]$ la inclusión de A en el anillo de polinomios. Para todo anillo conmutativo B , todo homomorfismo de anillos $\sigma : A \rightarrow B$ y todo elemento $b \in B$ existe un único homomorfismo de anillos $ev_b : A[X] \rightarrow B$ tal que $(ev_b)\lambda = \sigma$ y $ev_b(X) = b$.

Esta propiedad se visualiza mejor en un diagrama: Dadas λ y σ existe un único ev_b que hace el siguiente diagrama conmutativo y aplica X en b :

$$\begin{array}{ccc} A & \xrightarrow{\lambda} & A[X] \\ & \searrow \sigma & \downarrow ev_b \\ & & B \end{array}$$

Demostración. Sea $f = \sum_{i=0}^n a_i X^i$. Definimos $ev_b(f) = \sum_{i=0}^n \sigma(a_i) b^i$, es decir, aplicamos σ a todos los coeficientes de f , sustituimos X por b y realizamos en B las operaciones indicadas. Es rutina comprobar que ev_b es un homomorfismo de anillos, que

$$ev_b(X) = b$$

y que $ev_b \cdot \lambda = \sigma$.

Sea ahora $\tau : A[X] \rightarrow B$ otro homomorfismo de anillos que verifique las mismas propiedades y sea $f = \sum_{i=0}^n a_i X^i \in A[X]$ arbitrario. Entonces

$$\tau(f) = \tau\left(\sum_{i=0}^n a_i X^i\right) = \sum_{i=0}^n \tau(a_i) \tau(X)^i = \sum_{i=0}^n \sigma(a_i) b^i = ev_b(f)$$

luego $\tau = ev_b$ es único. \square

El morfismo ev_b del teorema anterior se llama *morfismo de evaluación en b* . Se aplica sobre todo cuando σ es una inclusión, es decir que para todo $a \in A$, $\sigma(a) = a$. En este caso $ev_b(a_n X^n + \cdots + a_1 X + a_0) = a_n b^n + \cdots + a_1 b + a_0$ es el resultado de *evaluar f en b* y se representa por $ev_b(f) = f(b)$.

Definición 5.9. Un elemento $a \in A$ se llama *cero* o *raíz* de f si $f(a) = 0$.

Todo polinomio $f \in A[X]$ define una *aplicación polinómica* $\tilde{f} : A \rightarrow A$ mediante $\tilde{f}(a) = f(a)$. En general, distintos polinomios pueden definir la misma aplicación polinómica.

Ejemplo 5.10. Sea $A = \mathbb{Z}_2$ el anillo de las clases de restos módulo 2. Sean $f = 0$, $g = X^2 + X$, $h = X^3 + X$ polinomios de $\mathbb{Z}_2[X]$. Como polinomios son *distintos*, pero los tres definen la misma función polinómica $\mathbb{Z}_2 \rightarrow \mathbb{Z}_2$, a saber la función que aplica todo elemento (sólo hay dos) de \mathbb{Z}_2 en el cero de \mathbb{Z}_2 .

El proceso de construir el anillo de polinomios en una indeterminada puede aplicarse a cualquier anillo conmutativo, en particular a un mismo anillo de polinomios $A[X]$: Sea Y otra indeterminada. Definimos $A[X, Y] = A[X][Y]$, el anillo de polinomios en dos indeterminadas con coeficientes en A . Sus elementos son de la forma

$$f = \sum_{i,j} a_{ij} X^i Y^j$$

donde la suma es finita (En lugar de ello se suele decir que tomamos la suma sobre todos los pares i, j pero con $a_{ij} = 0$ para casi todo par (i, j) , es decir, para todos excepto un conjunto finito).

Mas generalmente, definimos inductivamente el anillo de polinomios en las indeterminadas X_1, \dots, X_n por la regla

$$A[X_1, \dots, X_n] = A[X_1, \dots, X_{n-1}][X_n]$$

. En otras palabras, consideramos a los elementos de $A[X_1, \dots, X_n]$ como polinomios en X_n con coeficientes en $A[X_1, \dots, X_{n-1}]$. Naturalmente existe un monomorfismo $\lambda : A \rightarrow A[X_1, \dots, X_n]$ y A se identifica con el subanillo $\text{Im}(\lambda)$ de $A[X_1, \dots, X_n]$.

Lema 5.11. *El anillo conmutativo A es un dominio de integridad si y sólo si lo es $A[X_1, \dots, X_n]$*

Demostración. Inducción sobre n . □

De la definición tenemos que todo elemento f de $A[X_1, \dots, X_n]$ se escribe de manera única como

$$f = \sum a_{i_1 \dots i_n} X_1^{i_1} \dots X_n^{i_n}$$

Aquí $a_{i_1 \dots i_n}$ esta determinado de manera única como el coeficiente en f del monomio $X_1^{i_1} \dots X_n^{i_n}$. Formalmente la suma anterior es infinita, pero de hecho sólo un número finito de coeficientes son distintos de cero. Ya que las indeterminadas conmutan entre sí con los elementos de A , el anillo $A[X_1, \dots, X_n]$ depende simétricamente de las X_i ; así que X_n no juega ningún papel especial. Podíamos haber escrito f como un polinomio en X_1 con coeficientes en $A[X_2, \dots, X_n]$ o escoger cualquier otra X_i .

Definición 5.12. Cada producto $M_i = X_1^{i_1} \dots X_n^{i_n}$ se llama *monomio primitivo*; el término correspondiente $a_{i_1 \dots i_n} X_1^{i_1} \dots X_n^{i_n}$ se llama *monomio* o *término monomial*; su *grado total* (o sencillamente *grado*) es $\sum i_j$, y el *grado en X_j* es i_j . El *grado de f* es el máximo de los grados de sus términos no nulos.

Por ejemplo $f = 2X_1^5 X_2^3 X_3 - X_1^2 X_3^3 + 7X_2^6$ es de grado 5 en X_1 , de grado 6 en X_2 y de grado 3 en X_3 ; el grado total de f es 9.

Definición 5.13. Un polinomio en el que todos los términos tienen el mismo grado total se llama *polinomio homogéneo* o también una *forma*. En una indeterminada las únicas formas son los monomios, pero ya para dos indeterminadas puede haber otros, por ejemplo las forma cuadráticas $aX^2 + bXY + cY^2$.

Un criterio práctico de homogeneidades es el siguiente

Lema 5.14. *El polinomio $f \in A[X_1, \dots, X_n]$ es homogéneo de grado k si y sólo si para otra indeterminada t se verifica que*

$$f(tX_1, \dots, tX_n) = t^k f(X_1, \dots, X_n).$$

A veces es conveniente ordenar los monomios. Incluso para propósito tan sencillo como escribir la expresión total de un polinomio es necesario un orden total de los monomios. Con frecuencia se usa el *orden lexicográfico* definido así: Entre monomios de distinto grado total, el de mayor grado precede al de menor grado. Entre monomios del mismo grado total, el monomio $X_1^{i_1} \dots X_n^{i_n}$ precede a $X_1^{j_1} \dots X_n^{j_n}$ si la primera diferencia no nula $i_1 - j_1, \dots, i_n - j_n$ es positiva.

Por ejemplo, $X_1^3 X_2 X_3^2$ precede a $X_1^3 X_3^3$ y es precedido por $X_1^3 X_2^2 X_3$. En cualquier polinomio, el primer término monomial (en el orden lexicográfico) entre los términos de grado máximo se llama el *término líder*

5.2. El algoritmo de la división con resto

Teorema 5.15 (Algoritmo general de división). *Sea A un anillo conmutativo y sean $f, g \in A[X]$ con el coeficiente líder de g invertible. Entonces existen únicos $q, r \in A[X]$ tales que $f = qg + r$ y $gr(r) < gr(g)$.*

Demostración. Inducción sobre $gr(f)$. Sean $f = a_n X^n + \dots + a_1 X + a_0$ y $g = b_m X^m + \dots + b_0$. Si $gr(f) < gr(g)$, tomamos $q = 0$ y $r = f$. Sea ahora $gr(f) = n \geq gr(g) = m$. Definimos

$$f_1 = f - (a_n b_m^{-1}) X^{n-m} g \quad (5.1)$$

Es inmediato que $gr(f_1) < gr(f)$ y por inducción existen $q_1, r \in A[X]$ tales que $f_1 = q_1 g + r$ con $gr(r) < gr(g)$. Despejando en 5.1 vemos que

$$f = (a_n b_m^{-1}) X^{n-m} g + f = ((a_n b_m^{-1}) X^{n-m} g + q_1) g + r$$

Definimos $q = (a_n b_m^{-1}) X^{n-m} g + q_1$ y tenemos demostrada la existencia de cociente y resto.

Para ver la unicidad, sea $f = qg + r = q_1 g + r$. Trasponiendo términos tenemos $(q - q_1)g = r - r_1$. Como el coeficiente líder de g es invertible se verifica

$$gr(g) > \max(gr(r), gr(r_1)) \geq gr(r - r_1) = gr((q - q_1)g) = gr(q - q_1) + gr(g)$$

lo que implica que $gr(q - q_1) = -\infty$ y $q - q_1 = 0$. Luego $q = q_1$ y por tanto $r = r_1$. \square

Corolario 5.16. *Sea K un cuerpo. Entonces $K[X]$ es un anillo euclídeo*

Demostración. En un cuerpo, todo elemento no nulo es invertible. Así que para todo polinomio no nulo g el coeficiente líder es invertible. Por el teorema anterior, para cualesquiera polinomios f, g con $g \neq 0$ existen únicos q, r tales que $f = qg + r$ con $gr(r) < gr(g)$. Esta es la segunda condición en la definición de anillo euclídeo.

Por otro lado todo cuerpo es un dominio de integridad, así que para dos polinomios no nulos f, g se verifica $gr(fg) = gr(f) + gr(g) \geq gr(f)$. esta es la primera condición de dicha definición

Por tanto $K[X]$ es euclídeo respecto a la función grado. \square

Corolario 5.17 (Teorema del resto). *Sea A un anillo conmutativo, a un elemento de A y $f \in A[X]$ un polinomio. Entonces existe un $q \in A[X]$ tal que*

$$f = (X - a)q + f(a)$$

y $(X - a)$ divide a f si y sólo si $f(a) = 0$.

Teorema 5.18. *Sea A un dominio de integridad y sea $f \in A[X]$. Sean $a_1, \dots, a_m \in A$ elementos distintos tales que $f(a_i) = 0$ para $i = 1, \dots, m$. Entonces $((X - a_1) \dots (X - a_m))$ divide a f .*

Demostración. Inducción sobre m . Para $m = 1$ esto es parte del teorema del resto. Sea $m > 1$. Por inducción $f = (X - a_1) \dots (X - a_{m-1})g$ con $g \in A[X]$. Evaluamos en a_m :

$$0 = f(a_m) = (a_m - a_1) \dots (a_m - a_{m-1})g(a_m)$$

Como los a_i distintos, $a_m - a_i \neq 0$ para $i = 1, \dots, m - 1$. Como A es un dominio de integridad, $g(a_m) = 0$. Por el teorema del grado $g = (X - a_m)g_1$. Sustituyendo en la expresión de f nos queda $f = (X - a_1) \dots (X - a_{m-1})(X - a_m)g_1$ y por tanto el producto $((X - a_1) \dots (X - a_m))$ divide a f . \square

Corolario 5.19. *Sea A un dominio de integridad y $f \in A[X]$, $f \neq 0$. El número de raíces de f en A es menor o igual al grado de f .*

Ejemplo 5.20. El teorema y corolarios anteriores son falsos para anillos conmutativos generales: Sea $f = X^2 - 1 \in \mathbb{Z}_8[X]$. En \mathbb{Z}_8 el polinomio f tiene cuatro raíces distintas: 1, 3, 5, 7. Además $(X - 1)(X - 3)$ no divide a f .

Corolario 5.21. *Sea A un dominio de integridad, a_1, \dots, a_{n+1} elementos distintos de A y $f, g \in A[X]$ tales que $gr(f), gr(g) \leq n$ y $f(a_i) = g(a_i)$ para $i = 1, \dots, n + 1$. Entonces $f = g$.*

Demostración. El polinomio $f - g$ tiene grado menor o igual a n y tiene $n + 1$ raíces distintas. Luego tiene que ser el polinomio cero. \square

Corolario 5.22. Sea A un dominio de integridad infinito y sean $f, g \in A[X]$ tales que para todo $a \in A$ se verifica $f(a) = g(a)$. Entonces $f = g$.

Este último corolario nos dice que si A es un dominio de integridad infinito, la correspondencia entre polinomios y funciones polinómicas es biyectiva.

El anterior corolario se generaliza a varias indeterminadas:

Teorema 5.23. Sea A un dominio de integridad y $f \in A[X_1, \dots, X_n]$ tal que para cualesquiera $a_1, \dots, a_n \in A$ se verifica $f(a_1, \dots, a_n) = 0$. Entonces $f = 0$.

Demostración. Inducción sobre n . □

Corolario 5.24 (Principio de irrelevancia de desigualdades algebraicas). Sea A un dominio de integridad y sean $f, g, h \in A[X]$, $h \neq 0$ tales que para $a_1, \dots, a_n \in A$, $h(a_1, \dots, a_n) \neq 0 \Rightarrow f(a_1, \dots, a_n) = g(a_1, \dots, a_n)$. Entonces $f = g$.

Demostración. El polinomio $(f - g)h$ se anula sobre todos los $a_1, \dots, a_n \in A$. Luego $(f - g)h = 0$. Como $A[X_1, \dots, X_n]$ es un dominio de integridad y $h \neq 0$, necesariamente $f - g = 0$. □

El principio de irrelevancia de desigualdades algebraicas se llama también *propiedad de densidad*, por su interpretación en geometría algebraica.

5.3. Factorización

Sea K un cuerpo. El anillo $K[X]$ es un dominio euclídeo y por tanto también es un dominio de factorización única. Vamos ahora a estudiar la factorización de polinomios en ese anillo. En primer lugar caracterizamos los elementos invertibles:

Lema 5.25. Las unidades de $K[X]$ son los polinomios constantes no nulos.

El primer teorema proporciona algunos polinomios irreducibles:

Teorema 5.26. Los polinomios de grado uno son irreducibles en $K[X]$.

Estos son los únicos irreducibles si y sólo si todo polinomio de $K[X]$ de grado positivo tiene una raíz en K .

Demostración. El primer resultado se deduce del teorema del grado.

Supongamos que todo polinomio irreducible es de grado uno. El anillo $K[X]$ es un dominio de factorización única, por tanto todo polinomio f no constante es divisible por un irreducible, así que existe un $b_1X - b_0$ con $b_1 \neq 0$ tal que $f = (b_1X - b_0)q$. Pero entonces $f(b_1/b_0) = 0$ y f tiene una raíz $b_1/b_0 \in K$.

A la inversa, si todo polinomio no constante tiene una raíz en K , sea f un polinomio irreducible y sea $a \in K$ tal que $f(a) = 0$. Por el teorema del resto $X - a$ divide a f . Como f es irreducible, debe ser asociado a $X - a$ y por tanto es de grado uno. □

Definición 5.27. Un cuerpo en que todo polinomio no constante tiene una raíz se llama *algebraicamente cerrado*.

El llamado *teorema fundamental del álgebra* dice que el cuerpo \mathbb{C} de los números complejos es algebraicamente cerrado. Este hecho fue conjeturado por D'Alembert y demostrado por primera vez por el gran Gauss en su tesis doctoral. Dicha demostración tenía una laguna, pero a lo largo de su vida Gauss proporcionó cinco demostraciones correctas distintas. Sin embargo todas esas demostraciones utilizan bastante maquinaria analítica (como es propio, porque la construcción de \mathbb{C} se basa en \mathbb{R} que es el objeto de estudio del análisis matemático). Desde un punto de vista puramente algebraico, el hecho de que \mathbb{C} sea algebraicamente cerrado es relativamente poco importante. Es mas importante demostrar que todo cuerpo K es un subcuerpo de otro cuerpo \bar{K} algebraicamente cerrado.

La factorización de polinomios con coeficientes en un cuerpo algebraicamente cerrado (como \mathbb{C}) es muy sencilla: Todo polinomio no constante es un producto de polinomios de grado uno.

Sobre los números reales es casi igual de fácil: Todo polinomio es no constante es un producto de polinomios irreducibles de grado uno y dos. Sobre el cuerpo \mathbb{Q} de los números racionales la situación es muy diferente: Existen polinomios irreducibles de todos los grados y para un polinomio $f \in \mathbb{Q}[X]$ dado puede ser difícil hallar sus factores. El resto de esta sección va encaminado a intentar factorizar polinomios en $\mathbb{Q}[X]$.

Vamos a establecer los teoremas en un contexto mas general: Sea A un dominio de factorización única y sea K su cuerpo de fracciones.

Definición 5.28. Para todo polinomio no nulo $f = a_n X^n + \dots + a_0 \in A[X]$ llamamos *contenido de f* a $c(f) = \text{m. c. d.}(a_n, \dots, a_0)$.

Un polinomio $f \in A[X]$ se llama *primitivo* si $c(f) = 1$.

Lema 5.29. *Todo polinomio $f \in A[X]$ se expresa como $f = c(f)f_1$ con f_1 primitivo.*

Teorema 5.30 (Lema de Gauss). *El producto de dos polinomios primitivos es primitivo.*

Demostración. Sean $f = a_n X^n + \dots + a_0$, $g = b_m X^m + \dots + b_0$ dos polinomios primitivos de $A[X]$. Sea $p \in A$ un primo de A arbitrario. Como f, g son primitivos, $\text{m. c. d.}(a_n, \dots, a_0) = 1 = \text{m. c. d.}(b_m, \dots, b_0)$ y en cada uno de ellos existe por lo menos un coeficiente no divisible por p . Sean a_i y b_j los primeros coeficientes no divisibles por p , de forma que para todo $k > i$, p divide a a_k y para todo $l > j$, p divide a b_l . En el polinomio producto fg consideramos el coeficiente del término de grado $i + j$:

$$c_{i+j} = (a_{i+j}b_0 + \dots + a_{i+1}b_{j-1}) + a_i b_j + (a_{i-1}b_{j+1} + \dots + a_0 b_{i+j})$$

Todos los términos del primer paréntesis (que puede ser vacío) son divisibles por p , como también lo son todos los términos del segundo paréntesis (que también puede ser vacío). Así que $c_{i+j} = q_1p + a_ib_j + q_2p$ con $q_1, q_2 \in A$. Si p dividiese a c_{i+j} , necesariamente $p \mid a_ib_j$ y como p es primo, dividiría a uno de los factores, lo cual es imposible. Luego p no divide a c_{i+j} .

Hemos demostrado que para todo primo $p \in A$ existe un coeficiente del producto $h = fg$ que no es divisible por p . Luego el máximo común divisor de los coeficientes de h es 1 y h es primitivo. \square

Corolario 5.31. *Para dos polinomios $f, g \in A[X]$, el contenido del producto es el producto de los contenidos, es decir $c(fg) = c(f)c(g)$.*

Teorema 5.32. *Sea $f \in A[X]$ primitivo. Entonces f es irreducible en $A[X]$ si y sólo si es irreducible en $K[X]$.*

Demostración. Supongamos que $f = gh$ es una factorización de f en $K[X]$. Multiplicando por un denominador común obtenemos $k = af = bg_1h_1$, donde $a, b \in A$ y los polinomios g_1, h_1 son primitivos. Por el lema de Gauss el producto g_1h_1 también es primitivo. Luego a y b son ambos contenidos del polinomio k , luego son asociados. Sea $b = ua$ con u invertible. Sustituyendo y simplificando nos queda $f = (uf_1)g_1$ donde $uf_1, g_1 \in A[X]$ son primitivos y $gr(uf_1) = gr(f)$, $gr(g_1) = gr(g)$. Luego f es factorizable en $A[X]$.

A la inversa, sea $f = gh$ una factorización en $A[X]$. Los polinomios f, g no son constantes y tienen sus coeficientes en K , luego esa misma es una factorización en $K[X]$.

Hemos visto que f es reducible en $A[X]$ si y sólo si es reducible en $K[X]$. El contrarrecíproco es el resultado buscado. \square

Corolario 5.33. *Los elementos irreducibles en $A[X]$ son de uno de los siguientes tipos:*

1. *Polinomios de grado cero que son irreducibles en A*
2. *Polinomios primitivos que son irreducibles en $K[X]$.*

Teorema 5.34. *Sea A un dominio de integridad. El anillo A es un dominio de factorización única si y sólo si $A[X]$ es un dominio de factorización única.*

Demostración. En primer lugar supongamos que $A[X]$ es un dominio de factorización única. Los elementos de A pertenecen a $A[X]$ y por tanto descomponen de manera única como producto de irreducibles en $A[X]$, necesariamente todos de grado cero. Por tanto todo $a \in A$ descompone de manera única como producto de irreducibles en A . Luego A es un dominio de factorización única.

A la inversa sea A un dominio de factorización única. Sea $f \in A[X]$ no cero. Descomponemos $f = c(f)f_1$ con f_1 primitivo. Descomponemos $c(f) = p_1 \dots p_t$ en producto de irreducibles en A y $f_1 = q_1 \dots q_s$ en producto de primitivos irreducibles en $K[X]$. Entonces $f = p_1 \dots p_t q_1 \dots q_s$ es una descomposición de f en producto de irreducibles en $A[X]$.

Sea ahora p un irreducible en $A[X]$ y sean $f, g \in A[X]$ tales que p divide al producto fg .

Si $gr(p) = 0$, entonces p es irreducible y primo en A y p divide al contenido $c(fg) = c(f)c(g)$. Luego p divide a $c(f)$ (en cuyo caso divide a f) o divide a $c(g)$ (en cuyo caso divide a g). Luego p es primo.

Si $gr(p) > 0$, entonces p es un polinomio primitivo irreducible y por tanto primo en $K[X]$. Luego p divide a f o a g en $K[X]$. Sea q un polinomio en $K[X]$ tal que $f = pq$. Extrayendo contenidos, vemos que q pertenece a $A[X]$ y por tanto p divide a f en $A[X]$. Luego p es primo en $A[X]$.

Hemos demostrado que todo polinomio de $A[X]$ descompone como producto de irreducibles y que todo irreducible es primo. Luego $A[X]$ es un dominio de factorización única. \square

Corolario 5.35. *Sea A un dominio de integridad. Entonces A es un dominio de factorización única si y sólo si $A[X_1, \dots, X_n]$ es un dominio de factorización única.*

Corolario 5.36. *Sea K un cuerpo. El anillo $K[X_1, \dots, X_n]$ es un dominio de factorización única.*

5.4. Criterios de irreducibilidad

En esta sección A es un dominio de factorización única y K es su cuerpo de fracciones, salvo mención expresa en contrario.

La factorización en el anillo de polinomios $A[X]$ presenta dos problemas prácticos relacionados entre sí:

1. Dado un polinomio $f \in A[X]$ determinar si es reducible o irreducible
2. Si f es reducible, factorizarlo en irreducibles.

Para el primer caso muchas veces basta tener criterios suficientes (es decir, que si un polinomio satisface el criterio, es irreducible. Si no lo satisface no podemos decir nada). Evidentemente, una solución general del segundo punto incluiría criterios necesarios y suficientes para que un polinomio dado sea irreducible.

Empezamos determinando los factores de grado uno:

Proposición 5.37. *Sean $f = a_n X^n + \dots + a_0, g = b_m X + b_0 \in A[X]$ con $a_n, b_m \neq 0$. Si g divide a f , necesariamente b_m divide a a_n y b_0 divide a a_0*

Demostración. Sea $h = c_k X^k + \cdots + c_0 \in A[X]$ tales que $f = gh$. Entonces el coeficiente líder del producto es $a_n = b_m c_k$ y el término independiente es $a_0 = b_0 c_0$ \square

Corolario 5.38 (Regla de Ruffini). Sea $f = a_n X^n + \cdots + a_0$ y sea $a/b \in K$ tal que m. c. d. $(a, b) = 1$ y $f(a/b) = 0$. Entonces a divide a a_n y b divide a a_0 .

La regla de Ruffini la describió ya Newton en su libro *Arithmetica Universalis* (publicado en 1707, cincuenta y ocho años antes del nacimiento de Ruffini), para determinar las raíces racionales y enteras de polinomios con coeficientes enteros. El corolario anterior permite usarla para hallar las raíces de polinomios con coeficientes en cualquier dominio de factorización única.

Ejemplo 5.39. Sea $f = X^4 + 4 \in \mathbb{Z}[X]$. Cualquier raíz racional suya debe ser de la forma a/b con $b \mid 1$ y $a \mid 4$. Luego las posibles raíces racionales de f son $1, -1, 2, -2, 4, -4$. Un cálculo rápido muestra que ninguno de estos números es raíz de f , luego el polinomio f no tiene raíces en \mathbb{Q} .

Ejemplo 5.40. Sea ahora $f = X^4 + 4 \in \mathbb{J}[X]$. Los divisores de 4 son ahora $1, 1 + i, 2, 2 + 2i, 4$ y sus asociados (todos los productos por las unidades $\pm 1, \pm i$). Un nuevo cálculo muestra que $f(1 + i) = f(1 - i) = f(-1 + i) = f(-1 - i) = 0$, luego f tiene cuatro raíces en \mathbb{J} y factoriza como

$$X^4 + 4 = (X - (i + i))(X - (i - i))(X - (-i + i))(X - (-i - i))$$

Un criterio de aplicación muy rápida es debido a un discípulo de Gauss:

Teorema 5.41 (Criterio de Eisenstein). Sea $f = a_n X^n + \cdots + a_0$ un polinomio primitivo y sea $p \in A$ un primo tal que $p \nmid a_n$, $p \mid a_i$ para $i = n - 1, \dots, 0$ y $p^2 \nmid a_0$. Entonces f es irreducible en $A[X]$.

Demostración. Supongamos que f es reducible, $f = gh$ con $g = b_m X^m + \cdots + b_0$ y $h = c_r X^r + \cdots + c_0$ con $m, r \geq 1$ y $n = m + r$. Como p no divide a $a_n = b_m c_r$, necesariamente $p \nmid b_m$ y $p \nmid c_r$. Como p divide a $a_0 = b_0 c_0$, p debe dividir a uno de los factores, sea $p \mid b_0$. Entonces p no divide a c_0 porque $p^2 \nmid a_0 = b_0 c_0$. Sea i tal que $p \nmid b_i$ pero $p \mid b_j$ para todo $j < i$. El coeficiente en f del término de grado i es $a_i = b_i c_0 + (b_{i-1} c_1 + \cdots + b_0 c_i)$. Todos los términos del paréntesis son divisibles por p y $p \nmid b_i c_0$, luego $p \nmid a_i$. Pero $i \leq m < n$, luego por la hipótesis $p \mid a_i$, contradicción. \square

Ejemplo 5.42. Sea $f = 2X^5 - 6X^3 + 9X^2 - 15 \in \mathbb{Z}[X]$. El polinomio f es primitivo porque m. c. d. $(2, -6, 9, -15) = 1$. El primo 3 divide a todos los coeficientes menos al líder, y $3^2 = 9$ no divide al término independiente, luego f es irreducible en $\mathbb{Z}[X]$.

Ejemplo 5.43. Sea $f = Y^3 + X^2Y^2 + XY + X \in K[X, Y]$ con K un cuerpo arbitrario. Como $K[X, Y] = A[Y]$ con $A = K[X]$ dominio euclídeo, aplicando el criterio de Eisenstein con el primo $X \in A[X]$ vemos que f es irreducible en $K[X, Y]$.

A veces el polinomio dado no satisface las condiciones del criterio de Eisenstein pero un transformado sencillo sí las satisface. Del siguiente lema podemos deducir entonces la irreducibilidad del polinomio original:

Lema 5.44. Sea A un dominio de integridad y sea $f \in A[X]$. Sea $a \in A$ arbitrario y sea $f_a(X) = f(X+a)$. Entonces f descompone como $f = gh$ con $\text{gr}(g), \text{gr}(h) > 0$ si y sólo si $f_a = g_a h_a$. En este caso $\text{gr}(g_a) = \text{gr}(g) > 0$ y $\text{gr}(h_a) = \text{gr}(h) > 0$.

Demostración. Cálculo trivial. \square

Corolario 5.45. Sea A un dominio de factorización única y sea $f \in A[X]$ primitivo. Sea $a \in A$ arbitrario tal que f_a sea primitivo. Entonces f es irreducible si y sólo si f_a es irreducible

Ejemplo 5.46. Sea $f = X^4 + 1 \in \mathbb{Z}[X]$. No podemos aplicar directamente el criterio de Eisenstein a f . Pero

$$f_1 = f(X+1) = (X+1)^4 + 1 = X^4 + 4X^3 + 6X^2 + 4X + 2$$

satisface las condiciones del criterio de Eisenstein con $p = 2$. Luego f_1 es irreducible en $\mathbb{Z}[X]$ y por tanto también lo es f .

Ejemplo 5.47. (Este ejemplo se remonta a Gauss). Sea $p \in \mathbb{Z}$ un primo. El polinomio

$$\Phi_p = \frac{X^p - 1}{X - 1} = X^{p-1} + X^{p-2} + \cdots + X + 1$$

se llama *p-ésimo polinomio ciclotómico*. Vamos a comprobar que Φ_p es irreducible en $\mathbb{Z}[X]$ (y por tanto en $\mathbb{Q}[X]$): Calculamos el desarrollo de $f = \Phi_p(X+1)$:

$$f = \frac{(X+1)^p - 1}{(X+1) - 1} = \frac{(\sum_{i=0}^p \binom{p}{i} X^i) - 1}{X} = \sum_{i=1}^p \binom{p}{i} X^{i-1}$$

Ahora p no divide al coeficiente líder $\binom{p}{p} = 1$, p divide a $\binom{p}{i}$ para $i = p-1, \dots, 1$ y p^2 no divide al término independiente $\binom{p}{1} = p$. Luego f es irreducible en $\mathbb{Z}[X]$ y por tanto también lo es Φ_p .

A veces se utiliza otra transformación del polinomio:

Definición 5.48. Sea $f = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0 \in A[X]$ un polinomio con $a_n, a_0 \neq 0$. Se llama *polinomio recíproco de f* al polinomio

$$f_{\text{rec}} = a_0 X^n + a_1 X^{n-1} + \cdots + a_{n-1} X + a_n = X^n f\left(\frac{1}{X}\right)$$

Lema 5.49. Sea $f \in A[X]$ primitivo. Entonces f es irreducible en $A[X]$ si y sólo si f_{rec} es irreducible.

Demostración. Los coeficientes de f_{rec} son los mismos que los de f , luego f_{rec} es primitivo. Sea ahora $f = gh$ con $m = gr(g)$, $r = gr(h)$ y $n = gr(f) = m + r$. Entonces

$$f_{rec} = X^n f\left(\frac{1}{X}\right) = X^m X^r g\left(\frac{1}{X}\right) h\left(\frac{1}{X}\right) = g_{rec} h_{rec}$$

□

Ejemplo 5.50. Sea $f = 6X^4 + 9X^3 - 3X^2 + 1 \in \mathbb{Z}[X]$. El primo $p = 3$ divide a todos los coeficientes menos al término independiente y $3^2 = 9$ no divide al coeficiente líder, luego f es irreducible.

Cuando se puede aplicar, el criterio de Eisenstein es una prueba muy rápida de irreducibilidad. Pero son muy pocos los polinomios a los que es aplicable. Existe otro criterio que se puede aplicar a mas polinomios y aunque falle, los resultados que se obtienen en su aplicación son útiles para intentar posteriormente la factorización del polinomio.

Todo homomorfismo de anillos $\sigma : A \rightarrow B$ define un homomorfismo $A[X] \rightarrow B[X]$ que también se denota por σ de la siguiente forma: Sea $f = a_n X^n + \cdots + a_0$. Entonces $\sigma(f) = \sigma(a_n)X^n + \cdots + \sigma(a_0)$.

Proposición 5.51. Sean A, B dos dominios de integridad con cuerpos de fracciones respectivos K y L . Sea $\sigma : A \rightarrow B$ un homomorfismo de anillos y sea $f \in A[X]$ un polinomio tal que $gr(\sigma(f)) = gr(f)$. Si $f = gh$, entonces $\sigma(f) = \sigma(g)\sigma(h)$ con $gr(\sigma(g)) = gr(g)$ y $gr(\sigma(h)) = gr(h)$.

Corolario 5.52 (Criterio de reducción). Si $\sigma(f)$ es irreducible en $L[X]$, entonces f es irreducible en $K[X]$.

Usualmente este criterio se aplica con $A = \mathbb{Z}$, $K = \mathbb{Q}$, $B = L = \mathbb{Z}_p$ y $\sigma : \mathbb{Z} \rightarrow \mathbb{Z}_p$ la proyección canónica que lleva cada entero n en su clase módulo p , o sea $\sigma(n) = \bar{n} = [n]_p$. En este caso se suele denotar $\sigma(f) = \tilde{f}$.

Ejemplo 5.53. Sea $p \in \mathbb{Z}$ un número primo. El polinomio $X^p - X - 1 \in \mathbb{Z}_p[X]$ es irreducible, luego $f = X^p - X - 1$ es irreducible en $\mathbb{Z}[X]$.

De la misma forma el polinomio $f = X^5 - 5X^4 - 6X - 1 \in \mathbb{Z}[X]$ es irreducible en $\mathbb{Z}[X]$ (porque módulo 5, $\sigma(f) = X^5 - X - 1 \in \mathbb{Z}_5[X]$).

El inverso del criterio de irreducibilidad es falso:

Ejemplo 5.54. El polinomio $f = X^3 - 3 \in \mathbb{Z}[X]$ es irreducible por el criterio de Eisenstein, pero módulo 2 $\sigma(f) = (X + 1)(X^2 + X + 1)$, luego puede ocurrir perfectamente que f sea irreducible y $\sigma(f)$ no lo sea.

La proposición 5.51 puede usarse combinando la información sobre los factores de f que se obtiene utilizando diversos primos:

Ejemplo 5.55. Sea $f = X^5 - 6X^4 + 5X^2 - X + 2$. Módulo 2 tenemos $\bar{f} = X^5 + X^2 + X = X(X^4 + X + 1)$ con ambos factores irreducibles. Si f es reducible, debe factorizar como producto de un polinomio de grado 1 por otro de grado 4.

Reduciendo módulo 3 queda $\bar{f} = X^5 - X^2 - X - 1 = (X^2 + 1)(X^3 - X - 1)$ con ambos factores irreducibles, así que si f fuese reducible debería factorizar como producto de un polinomio de grado 2 por otro de grado 3. Luego las factorizaciones módulo 2 y tres son incompatibles y f es irreducible en $\mathbb{Z}[X]$.

Ejemplo 5.56. Sea $f = X^4 - 22X^2 + 1 \in \mathbb{Z}[X]$. Reduciendo módulo 2 obtenemos $\bar{f} = X^4 + 1 = (X + 1)^4$, lo que no nos da información interesante. Módulo 3 es $\bar{f} = X^4 + 2X^2 + 1 = (X^2 + 1)^2$, luego si f factoriza en $\mathbb{Z}[X]$, debe hacerlo como producto de dos polinomios de grado 2, $f = gh$. Además los términos constantes de g y h deben ser divisores de 1 y congruentes con 1 módulo 3, luego ambos valen 1.

Supongamos que $f = (X^2 + aX + 1)(X^2 + bX + 1) = X^4 + (a + b)X^3 + (ab + 2)X^2 + (a + b)X + 1$. Comparando coeficientes debe ser $a + b = 0$ y $ab + 2 = -22$, así que $b = -a$ y $a^2 = 24$. Esta última ecuación no tiene solución con a entero, luego la factorización es imposible y f es irreducible.

5.5. Factorización en un número finito de pasos

Si el polinomio dado f es reducible, el problema es determinar los factores de f en un número finito de pasos (y en un tiempo razonable). En el libro *Arithmetica Universalis* citado antes, Newton describe cómo hallar los factores cuadráticos de un polinomio con coeficientes enteros. Esta es la traslación de dicho método a dominios de factorización única:

Sea A un dominio de factorización única con un número finito de unidades y sea $f = gh \in A[X]$ con

$$\begin{aligned} f &= a_n X^n + \cdots + a_0 \\ g &= b_2 X^2 + b_1 X + b_0 \end{aligned}$$

Entonces b_2 divide a a_n , b_0 divide a a_0 y $g(1) = b_2 + b_1 + b_0$ divide a $f(1) = a_n + \cdots + a_0$. Estas condiciones limitan a un número finito las posibilidades para b_2, b_0 y b_1 . Para cada una de las ternas (b_2, b_1, b_0) posibles construimos el polinomio g y probamos a dividir f por g . Así se determinan todos los factores cuadráticos de f .

Para limitar aún más el conjunto de posibles divisores se utiliza la condición de que $g(-1) = b_2 - b_1 + b_0$ divide a la suma alternada $(-1)^n f(-1) = a_n - a_{n-1} + \cdots \pm 1$. Además podemos utilizar la información que hayamos obtenido reduciendo diversos primos.

Ejemplo 5.57. Sea $f = X^4 + 4 \in \mathbb{Z}[X]$. Usando la regla de Ruffini vemos que f no tiene raíces enteras. Sea $g = b_2X^2 + b_1X + b_0$ un factor de f . Como f es mónico, podemos tomar $b_2 = 1$. Reduciendo módulo 2 tenemos $\bar{f} = X^4 = \bar{g}\bar{h}$, luego los términos constantes de g y del cociente h son pares. Como su producto es 4, necesariamente $b_0 = \pm 2$. Módulo 3 tenemos $\bar{f} = X^4 + 1 = (X^2 + X - 1)(X^2 - X - 1)$, luego $b_0 \equiv -1 \pmod{3}$, lo que nos deja $b_0 = 2$.

Ahora $1 + b_1 + 2 = b_1 + 3$ divide a $f(1) = 5$, lo que se verifica sólo para $b_1 = -8, -4, -2, 2$ y $1 - b_1 + 2 = -b_1 + 3$ divide a $f(-1) = 5$, lo que reduce las posibilidades a $b_1 = -2, 2$. Así que los únicos divisores de grados dos posibles son $g_1 = X^2 + 2X + 2$ y $g_2 = X^2 - 2X + 2$. Un cálculo fácil muestra que $f = g_1g_2$.

El anterior método de Newton fué extendido en 1793 por Friedrich von Schubert, quien mostró cómo hallar todos los factores de grado m en un número finito de pasos. Unos 90 años después Leopoldo Kronecker descubrió independientemente el método de Schubert. Desgraciadamente el método es muy ineficiente cuando $gr(f) \geq 5$ y es mejor utilizar métodos de reducción (descritos en [3] y [8]). El método de Kronecker se describe en la siguiente demostración:

Teorema 5.58 (Kronecker). *Sea A un dominio de factorización única con un número finito de unidades. Entonces es posible descomponer cualquier polinomio $f \in A[X]$ en factores irreducibles en un número finito de pasos.*

Demostración. Dado un polinomio $f \in A[X]$, el método consiste en determinar para cada $m < n/2$ un conjunto finito \mathcal{S} de polinomios entre los que están todos los divisores de f de grado menor o igual a m . Posteriormente se prueba a dividir f por cada uno de los polinomios del conjunto \mathcal{S} y así determinamos los divisores de grado menor o igual a m .

Si $f = gh$, para todo $a \in A$ se verifica que $f(a) = g(a)h(a)$, luego $g(a)$ divide a $f(a)$. Sean $a_0, \dots, a_m \in A$ elementos distintos. Para cada $i = 0, \dots, m$ sea D_i el conjunto de divisores de $f(a_i)$. Para cada sucesión $\mathbf{b} = (b_0, \dots, b_m) \in D_0 \times \dots \times D_m$ sea $g_{\mathbf{b}}$ el único polinomio de grado menor o igual a m que verifica $g_{\mathbf{b}}(a_i) = b_i$, $i = 0, \dots, m$ (El polinomio $g_{\mathbf{b}}$ es el *polinomio de interpolación*, que se obtiene por uno de los métodos de Newton o de Lagrange). El conjunto $\mathcal{S} = \{g_{\mathbf{b}} \mid \mathbf{b} \in D_0 \times \dots \times D_m\}$ es finito y contiene a todos los divisores de f de grado menor o igual a m . \square

En la práctica se achica bastante el conjunto \mathcal{S} utilizando la información que hayamos obtenido por reducción módulo diversos primos, igual que hicimos antes en el ejemplo 5.57

Ejemplo 5.59. Sea $f = X^6 - X^5 - X^4 + X^3 + X^2 - X - 1 \in \mathbb{Z}[X]$. Queremos encontrar los factores de grado menor o igual a tres, así que evaluamos f en cuatro ($=3+1$) puntos distintos. Elegimos los puntos $-2, -1, 0, 1$. Evaluamos: $f(-2) = 77$, $f(-1) = 1$, $f(0) = -1$, $f(1) = -1$. El conjunto $D_0 \times \dots \times D_3 = \{\pm 1, \pm 7, \pm 11, \pm 77\} \times \{\pm 1\} \times$

$\{\pm 1\} \times \{\pm 1\}$, así que en total hay que calcular $8 \cdot 2 \cdot 2 \cdot 2 = 64$ polinomios. Usando los interpoladores de Lagrange, estos polinomios son

$$\begin{aligned} f_b &= b_0 \frac{(X+1)X(X-1)}{(-2+1)(-2)(-2-1)} + b_1 \frac{(X+2)X(X-1)}{(-1+2)(-1)(-1-1)} \\ &+ b_2 \frac{(X+2)(X+1)(X-1)}{(0+2)(0+1)(0-1)} + b_3 \frac{(X+2)(X+1)X}{(1+2)(1+1)(1)} \\ &= b_0 \frac{X^3 - X}{-6} + b_1 \frac{X^3 + X^2 - 2X}{2} + b_2 \frac{X^3 + 2X^2 - X - 2}{-2} + b_3 \frac{X^3 + 3X^2 + 2X}{6} \end{aligned}$$

Calculamos los 64 polinomios. La mitad de ellos no tiene coeficientes enteros y los restantes se agrupan de dos en dos salvo el signo. Eligiendo uno de cada par de opuestos, nos quedan dieciseis polinomios:

b_0	b_1	b_2	b_3	g_b
1	1	1	1	1
7	1	1	1	$-X^3 + X + 1$
-11	1	1	1	$2X^3 - 2X + 1$
-77	1	1	1	$13X^3 - 13X + 1$
1	-1	1	1	$-X^3 - X^2 + 2X + 1$
7	-1	1	1	$-2X^3 - X^2 + 3X + 1$
-11	-1	1	1	$X^3 - X^2 + 1$
-77	-1	1	1	$12X^3 - X^2 - 11X + 1$
1	1	-1	1	$X^3 + 2X^2 - X - 1$
7	1	-1	1	$2X^2 - 1$
-11	1	-1	1	$3X^3 + 2X^2 - 3X - 1$
-77	1	-1	1	$14X^3 + 2X^2 - 14X - 1$
1	-1	-1	1	$X^2 + X - 1$
7	-1	-1	1	$-X^3 + X^2 + 2X - 1$
-11	-1	-1	1	$2X^3 + X^2 - X - 1$
-77	-1	-1	1	$13X^3 + X^2 - 12X - 1$

El polinomio f dado es mónico, así que buscamos factores mónicos. Repasando la lista anterior nos queda que sus posibles divisores mónicos de grado menor o igual que tres son

$$\begin{aligned}
&1 \\
&X^3 - X - 1 \\
&X^3 + X^2 - 2X + 1 \\
&X^3 - X^2 + 1 \\
&X^3 + 2X^2 - X - 1 \\
&X^2 + X - 1 \\
&X^3 - X^2 - 2X + 1
\end{aligned}$$

El 1 es trivial. Probando a dividir sucesivamente por cada uno de los otros obtenemos la factorización

$$f = (X^3 - X - 1)(X^3 - X^2 + 1)$$

y los dos factores son irreducibles (son de grado 3 y no tienen raíces enteras).

5.6. Polinomios simétricos

Sea A un anillo conmutativo y sean X_1, \dots, X_n indeterminadas. Sea S_n el grupo simétrico sobre $\{1, \dots, n\}$. Para toda permutación $\sigma \in S_n$ definimos

$$\sigma \cdot f(X_1, \dots, X_n) = f(X_{\sigma(1)}, \dots, X_{\sigma(n)})$$

Por ejemplo, sea $f = X_1^2 X_2 - X_3$ y sean $\rho = (1 \ 3)$, $\sigma = (1 \ 2 \ 3)$. Entonces $\rho \cdot f = X_3^2 X_2 - X_1$ y $\sigma \cdot f = X_2^2 X_3 - X_1$.

Definición 5.60. Un polinomio $f \in A[X_1, \dots, X_n]$ se llama *simétrico* si para toda permutación $\sigma \in S_n$ se verifica $\sigma \cdot f = f$.

Lema 5.61. El conjunto de polinomios simétricos es un subanillo de $A[X_1, \dots, X_n]$ que contiene al anillo A .

Sea Y otra indeterminada. Formamos el polinomio

$$\begin{aligned}
F(Y, X_1, \dots, X_n) &= (Y - X_1) \dots (Y - X_n) \\
&= Y^n - s_1 Y^{n-1} + \dots + (-1)^n s_n
\end{aligned}$$

con coeficientes en $A[X_1, \dots, X_n]$. Los polinomios coeficientes $s_1 = X_1 + \dots + X_n, \dots, s_n = X_1 \dots X_n$ son polinomios simétricos, y se llaman *polinomios simétricos elementales*. Obsérvese que el polinomio s_i es homogéneo de grado i .

Definición 5.62. Sea $a_e X_1^{e_1} \dots X_n^{e_n}$ un monomio no nulo. Se llama *peso* del monomio al entero $e_1 + 2e_2 + \dots + ne_n$.

Sea $g \in A[X_1, \dots, X_n]$. El *peso de* g es el mayor de los pesos de los monomios no nulos de g .

Teorema 5.63 (Teorema fundamental de los polinomios simétricos). *Sea A un dominio de integridad y sea $f \in A[X_1, \dots, X_n]$ un polinomio simétrico de grado d . Entonces existe un único polinomio $g \in A[X_1, \dots, X_n]$ de peso menor o igual a d tal que*

$$f(X_1, \dots, X_n) = g(s_1, \dots, s_n)$$

Demostración. Inducción sobre n y d .

Si $n = 1$, sólo hay una indeterminada, así que $s_1 = X_1$ y $g = f$ verifica las condiciones (el peso de f es igual al grado).

Sea ahora $n > 1$ y supongamos el teorema cierto para $n - 1$ indeterminadas. Si $d = 0$, el polinomio f es constante. Tomando $g = f$ se verifica el teorema (en este caso, el grado y el peso de f son ambos iguales a cero).

Finalmente sean $n > 1$, $d > 0$ y suponemos el teorema cierto para todo polinomio simétrico en n indeterminadas de grado menor que d . En el anterior polinomio F sustituimos $X_n = 0$. Obtenemos

$$\begin{aligned} F(Y, X_1, \dots, X_{n-1}, 0) &= (Y - X_1) \dots (Y - X_{n-1})Y \\ &= Y^n - (s_1)_0 Y^{n-1} + \dots + (-1)^{n-1} (s_{n-1})_0 Y \end{aligned}$$

donde $(s_i)_0$ se obtiene sustituyendo $X_n = 0$ en s_i .

Es inmediato que $(s_1)_0, \dots, (s_{n-1})_0$ son precisamente los polinomios simétricos elementales en X_1, \dots, X_{n-1} .

El polinomio $f(X_1, \dots, X_{n-1}, 0) \in A[X_1, \dots, X_{n-1}]$ es simétrico. Por la hipótesis de inducción sobre n , existe un polinomio $g_1 \in A[X_1, \dots, X_{n-1}]$ de peso menor o igual a d tal que $f(X_1, \dots, X_{n-1}, 0) = g_1((s_1)_0, \dots, (s_{n-1})_0)$. El polinomio

$$f_1(X_1, \dots, X_n) = f(X_1, \dots, X_n) - g_1(s_1, \dots, s_{n-1})$$

es simétrico y tiene grado menor o igual a d . Además $f_1(X_1, \dots, X_{n-1}, 0) = 0$, luego f_1 es divisible por X_n . Como es simétrico, también es divisible por X_1, \dots, X_{n-1} . Como estos factores son primos relativos, su producto divide a f_1 . Luego $f_1 = s_n f_2(X_1, \dots, X_n)$ con un polinomio $f_2 \in A[X_1, \dots, X_n]$ que es simétrico y de grado estrictamente menor que d . Por la inducción sobre d , existe un $g_2 \in A[X_1, \dots, X_n]$ de peso menor o igual a $d - n$ tal que

$$f_2(X_1, \dots, X_n) = g_2(s_1, \dots, s_n)$$

Sustituyendo obtenemos

$$f(X_1, \dots, X_n) = g_1(s_1, \dots, s_{n-1}) + s_n g_2(s_1, \dots, s_{n-1})$$

y cada término del miembro de la derecha tiene un peso menor o igual a d .

La unicidad se deduce del próximo teorema. \square

Teorema 5.64. Sea $g \in A[X_1, \dots, X_n]$. Entonces $g(s_1, \dots, s_n) = 0$ si y sólo si $g(X_1, \dots, X_n) = 0$.

Demostración. Inducción sobre n . Si $n = 1$ el resultado es trivial.

Sea ahora $n > 1$ y suponemos el resultado cierto para $n-1$ indeterminadas. Sea $g \in A[X_1, \dots, X_n]$ no nulo de grado mínimo tal que $g(s_1, \dots, s_n) = 0$. Escribimos g como un polinomio en X_n con coeficientes en X_1, \dots, X_{n-1} :

$$g = g_0 + \dots + g_d \cdot X_n^d$$

Sustituyendo X_i por s_i en el polinomio g tenemos

$$0 = g_0(s_1, \dots, s_{n-1}) + \dots + g_d(s_1, \dots, s_{n-1})s_n^d$$

Sustituyendo ahora $X_n = 0$ obtenemos

$$0 = g_0((s_1)_0, \dots, (s_{n-1})_0)$$

Pero los $(s_i)_0$ son los polinomios simétricos elementales en X_1, \dots, X_{n-1} . Por inducción $g_0(X_1, \dots, X_{n-1}) = 0$.

Ya que $g_0 = 0$ podemos escribir $g = f \cdot X_n$ con $f \in A[X_1, \dots, X_n]$ y por tanto $f(s_1, \dots, s_n)s_n = 0$, luego $f(s_1, \dots, s_n) = 0$ y f es de grado estrictamente menor que g , lo cual es imposible. \square

Ejemplo 5.65. Sea $f = (X_1 + X_2)(X_1 + X_3)(X_2 + X_3) \in \mathbb{Z}[X_1, X_2, X_3]$. Es fácil comprobar que f es un polinomio simétrico homogéneo de grado 3. Queremos encontrar un polinomio $g \in \mathbb{Z}[X_1, X_2, X_3]$ de peso menor o igual a 3 tal que $f = g(s_1, s_2, s_3)$. Para ello aplicamos la construcción de la demostración:

1. $f(X_1, 0, 0) = 0$, luego $g_1 = 0$.
2. $f(X_1, X_2, 0) = (X_1 + X_2)X_1X_2$. El resto del proceso de la demostración es trivial: $f(X_1, X_2, 0) = g((s_1)_0, (s_2)_0) = (s_1)_0(s_2)_0$.

3. La demostración construye ahora el polinomio

$$\begin{aligned}
f_1(X_1, X_2, X_3) &= f(X_1, X_2, X_3) - g(s_1, s_2) \\
&= (X_1 + X_2)(X_1 + X_3)(X_2 + X_3) \\
&\quad - (X_1 + X_2 + X_3)(X_1X_2 + X_1X_3 + X_2X_3) \\
&= (X_1^2X_2 + X_1^2X_3 + X_1X_2^2 + X_1X_3^2 + X_2^2X_3 + X_2X_3^2 + 2X_1X_2X_3) \\
&\quad - (X_1^2X_2 + X_1^2X_3 + X_1X_2^2 + X_1X_3^2 + X_2^2X_3 + X_2X_3^2 + 3X_1X_2X_3) \\
&= -X_1X_2X_3
\end{aligned}$$

$$\text{luego } f(X_1, X_2, X_3) = f_1(X_1, X_2, X_3) + X_1X_2X_3 = s_1s_2 - s_3.$$

Para escribir los polinomios simétricos se ha desarrollado una notación especial: Llamamos $\sum X_1^{i_1} \dots X_n^{i_n}$ a la suma de todos los monomios distintos que se obtienen al aplicar todas las permutaciones de S_n al monomio $X_1^{i_1} \dots X_n^{i_n}$. Por ejemplo si $n = 3$,

$$\begin{aligned}
\sum X_1^3 &= X_1^3 + X_2^3 + X_3^3 \\
\sum X_1^2X_2 &= X_1^2X_2 + X_1^2X_3 + X_2^2X_1 + X_2^2X_3 + X_3^2X_1 + X_3^2X_2
\end{aligned}$$

. Un polinomio simétrico general es una combinación lineal de términos de la forma $\sum X_1^{i_1} \dots X_n^{i_n}$ con coeficientes en A .

Ejemplo 5.66. Sea $f = \sum X_1^2X_2$ con $n = 3$. Calculamos $f(X_1, X_2, 0) = X_1^2X_2 + X_2^2X_1 = X_1X_2(X_1 + X_2) = (s_2)_0(s_1)_0$.

Ahora $f_1 = f - s_2s_1 = \sum X_1^2X_2 - (\sum X_1X_2)(\sum X_1) = -3X_1X_2X_3$. Luego $\sum X_1^2X_3 = s_1s_2 - 3s_3$.

Ejemplo 5.67. Seguimos tomando $n = 3$. Sea $f = \sum X_1^3$.

Entonces $f(X_1, 0, 0) = X_1^3 = (s_1)_{00}^3$.

El siguiente paso calcula $f(X_1, X_2, 0) - (s_1)_0^3 = -3(s_1)_0(s_2)_0$.

Luego $f(X_1, X_2, 0) = (s_1)_0^3 - 3(s_1)_0(s_2)_0$.

Finalmente calculamos $f_1(X_1, X_2, X_3) = f - (s_1^3 - 3s_1s_2) = 3X_1X_2X_3$,

así que $f = s_1^3 - 3s_1s_2 + 3s_3$.

Ejemplo 5.68. Sea $\Delta = \prod_{i < j} (X_i - X_j)$. El polinomio $d = \Delta^2$ es simétrico. Vamos a expresarlo en función de los polinomios simétricos elementales para $n = 3$.

1. En primer lugar $d(X_1, 0, 0) = ((X_1 - 0)((X_1 - 0)(0 - 0))^2 = 0$.

2. Ahora $d(X_1, X_2, 0) = ((X_1 - X_2)X_1X_2)^2$. Luego $d(X_1, X_2) = s_2^2 \cdot f_1$ con $f_1(X_1, X_2) = (X_1 - X_2)^2$.

$f_1(X_1, 0) = X_1^2$. Entonces

$$f_1(X_1, X_2) - (s_1)_0^2 = (X_1 - X_2)^2 - (X_1 + X_2)^2 = -4X_1X_2$$

y por tanto $f_1 = (s_1)_0^2 - 4(s_2)_0$.

3. Finalmente tenemos

$$\begin{aligned} g_1(X_1, X_2, X_3) &= d(X_1, X_2, X_3) - s_2^2(s_1^2 - 4s_2) \\ &= s_3 \cdot f_2 \end{aligned}$$

con $f_2 = 6 \sum X_1^2 X_2 - 4 \sum X_1^3 + 3X_1 X_2 X_3$. Por los dos ejemplos anteriores,

$$\begin{aligned} d &= s_1^2 s_2^2 - 4s_2^3 + s_3(6(s_1 s_2 - 3s_3) - 4(s_1^3 - 3s_1 s_2 + 3s_3) + 3s_3) \\ &= s_1^2 s_2^2 - 4s_2^3 - 4s_1^3 s_3 + 18s_1 s_2 s_3 - 27s_3^2 \end{aligned}$$

Existen otros tres métodos para expresar un polinomio simétrico en función de los simétricos elementales. Quizá el mas útil sea el *método de coeficientes indeterminados*: Descomponemos el polinomio simétrico dado en suma de polinomios simétricos homogéneos y expresamos cada uno de estos en función de los polinomios simétricos elementales. Para ello, expresamos cada uno de los polinomios homogéneos de grado d como suma con coeficientes indeterminados de todos los k monomios posibles en los s_i de peso d . Sustituimos las indeterminadas X_i por k conjuntos de valores concretos, lo que nos establece un sistema lineal de k ecuaciones en los coeficientes, sistema que resolvemos por los métodos de álgebra lineal.

Ejemplo 5.69. Sea $f = (X_1 + X_2 - X_3 - X_4)(X_1 - X_2 + X_3 - X_4)(X_1 - X_2 - X_3 + X_4)$. Es fácil comprobar que f es simétrico homogéneo de grado 3. La lista de todos los monomios posibles de peso 3 es la siguiente: $s_1^3, s_1 s_2, s_3$. Así que expresamos

$$f = a s_1^3 + b s_1 s_2 + c s_3$$

Ahora consideramos tres conjuntos de valores para los X_i de manera que nos quede un sistema determinado de tres ecuaciones lineales en a, b, c . Por ejemplo los valores

X_1	X_2	X_3	X_4	s_1	s_2	s_3
1	0	0	0	1	0	0
1	1	0	0	2	1	0
1	1	1	0	3	3	1

nos dan el sistema

$$\begin{aligned} f(1, 0, 0, 0) &= 1 = a \\ f(1, 1, 0, 0) &= 0 = 8a + 2b \\ f(1, 1, 1, 0) &= -1 = 27a + 9b + c \end{aligned}$$

que tiene la solución $a = 1, b = -4, c = 8$. Luego

$$f = s_1^3 - 4s_1 s_2 + 8s_3$$

Ejercicio 5.1. Sea $n \geq 5$. Expresar el polinomio $f = \sum x_1^2 x_2^2 x_3$ como un polinomio en los simétricos elementales.

Otro tipo de polinomios interesantes son los definidos a continuación:

Definición 5.70. Un polinomio $f \in A[X_1, \dots, X_n]$ se llama *alternado* si para toda permutación $\sigma \in s_n$ se verifica $\sigma \cdot f = \text{sgn}(\sigma)f$.

El polinomio alternado no nulo mas sencillo es el producto de todas las diferencias

$$\Delta = \prod_{i < j} (X_i - X_j)$$

Cada par ordenado de índices $i < j$ aparece exactamente una vez, así que en total hay $n(n-1)/2$ factores lineales y Δ es un polinomio homogéneo de grado $n(n-1)/2$. Cuando aplicamos una trasposición $(i j)$ a Δ , los factores se permutan entre sí, excepto el factor $X_i - X_j$ que se transforma en $X_j - X_i$, luego Δ cambia de signo.

Teorema 5.71. Sea A un dominio de integridad de característica distinta de 2. Todo polinomio f alternado de $A[X_1, \dots, X_n]$ es de la forma $f = \Delta g$, donde g es simétrico.

Demostración. Sustituyendo $X_2 = X_1$ obtenemos

$$f(X_1, X_1, \dots, X_n) = -f(X_1, X_1, \dots, X_n)$$

y como $\text{car}(A) \neq 2$, necesariamente $f(X_1, X_1, \dots, X_n) = 0$, luego $(X_1 - X_2)$ divide a f . De la misma forma $X_i - X_j$ divide a f para todo par $i < j$. Como estos polinomios son primos relativos, su producto divide a f así que existe un $g \in A[X_1, \dots, X_n]$ con $f = \Delta g$. Claramente $g = f/\Delta$ es un polinomio simétrico. \square

Corolario 5.72. Sea $f \in A[X_1, \dots, X_n]$ un polinomio alternado. Entonces $\text{gr}(f) \geq n(n-1)/2$.

5.7. La resultante

5.7.1. Introducción

El problema fundamental de la teoría de eliminación es el siguiente: Dados dos polinomios con coeficientes en un cuerpo F :

$$\begin{aligned} f &= a_n x^n + a_{n-1} x^{n-1} + \dots + a_0, & a_n &\neq 0 \\ g &= b_m x^m + b_{m-1} x^{m-1} + \dots + b_0, & b_m &\neq 0 \end{aligned} \quad (5.2)$$

determinar si tienen una raíz común en una extensión de F y en caso afirmativo hallarla. Para responder a esta cuestión, se busca una expresión que se anule sólo cuando f y g tienen una raíz común y que además sea calculable como función racional de los coeficientes de f y g . La más sencilla de tales expresiones es la resultante que vamos a definir y estudiar.

5.7.2. Definición

Sea K un cuerpo de descomposición para fg , así que en $K[X]$ tenemos:

$$\begin{aligned} f &= a_n(X - \alpha_1) \dots (X - \alpha_n) = a_n \prod_{i=1}^n (X - \alpha_i) \\ g &= b_m(X - \beta_1) \dots (X - \beta_m) = b_m \prod_{j=1}^m (X - \beta_j) \end{aligned} \quad (5.3)$$

La *resultante* de f y g viene definida por

$$R(f, g) = a_n^m b_m^n \prod_{i=1}^n \prod_{j=1}^m (\alpha_i - \beta_j) \quad (5.4)$$

5.7.3. Propiedades

1. $R(f, g) = 0 \Leftrightarrow \exists i, j$ tales que $\alpha_i = \beta_j$ (i.e., si f y g tienen una raíz en común)
2. $R(g, f) = (-1)^{nm} R(f, g)$
3. $R(f, g) = a_n^m \prod_{i=1}^n g(\alpha_i) = (-1)^{nm} b_m^n \prod_{j=1}^m f(\beta_j)$
4. $R(fg, h) = R(f, h)R(g, h)$, $R(f, gh) = R(f, g)R(f, h)$
5. Si $m = 0$ (i.e. si $g = b$ es un escalar), $R(f, b) = b^n$
6. $R(X^k, f) = a_0^k$; $R(f, X^k) = (-1)^{nk} a_0^k$
7. Si $g = fq + r$, $R(f, g) = a_n^{gr(g)-gr(r)} R(f, r)$
 Demostración: $R(f, g) = a_n^m \prod_{i=1}^n g(\alpha_i) = a_n^m \prod_{i=1}^n (f(\alpha_i)q(\alpha_i) + r(\alpha_i)) = a_n^n \prod_{i=1}^n r(\alpha_i) = a_n^{m-gr(r)} R(f, r)$
8. $R(X^k f, g) = b_0^k R(f, g)$; $R(f, X^k g) = (-1)^{nk} a_0^k R(f, g)$
9. $R(f, g)$ es un polinomio simétrico de grado m en las α_i
10. $R(f, g)$ es un polinomio simétrico de grado n en las β_j

11. $R(f, g)$ es un polinomio homogéneo de grado m en las a_i

Demostración: Por la propiedad 9, $R(f, g)$ es expresable como un polinomio en los polinomios simétricos elementales $\sigma_i = (-1)^i \frac{a_i}{a_0}$. Por el factor a_0^m todos los denominadores se simplifican.

12. $R(f, g)$ es un polinomio homogéneo de grado n en las b_j

13. El término $a_n^m b_0^n$ tiene coeficiente $+1$ en $R(f, g)$

Demostración: Dicho término sólo aparece al desarrollar $a_n^m b_0^n \prod_{i=1}^n (-\beta_i) = a_n^m b_0^n$.

5.8. El discriminante

El caso particular más importante de la resultante es cuando $g = f'$ (la derivada formal). En ese caso, $R(f, f') = 0 \Leftrightarrow f$ tiene raíces múltiples. Explícitamente, sean

$$\begin{aligned} f &= a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0 = a_n \prod_{i=1}^n (X - \alpha_i) \\ f' &= n a_n X^{n-1} + (n-1) a_{n-1} X^{n-2} + \dots + a_1 = a_n \sum_{j=1}^n \prod_{i \neq j} (X - \alpha_i) \\ f'(\alpha_j) &= a_n \prod_{i \neq j} (\alpha_j - \alpha_i) \\ R(f, f') &= a_n^{n-1} \prod_{j=1}^n f'(\alpha_j) = a_n^{2n-1} \prod_{j=1}^n \prod_{i \neq j} (\alpha_j - \alpha_i) \end{aligned} \quad (5.5)$$

5.8.1. Definición

Llamamos *discriminante de f* a $D(f) = a_n^{2n-2} \prod_{i < j} (\alpha_i - \alpha_j)^2$. Comparando con (5.5) obtenemos:

$$R(f, f') = (-1)^{\frac{n(n-1)}{2}} a_n D(f) \quad (5.6)$$

5.8.2. Propiedades

1. $f_1, f_2 \in F[X] \Rightarrow D(f_1 f_2) = D(f_1) D(f_2) R(f_1, f_2)^2$
2. $f_1, \dots, f_r \in F[X] \Rightarrow D(f_1 \dots f_r) = D(f_1) \dots D(f_r) R^2$ con $R \in F$

5.9. Métodos de cálculo

En esta sección nos planteamos encontrar una expresión explícita (o un método de cálculo) para $R(f, g)$ y $D(f)$ en función de los coeficientes de f y g . Para ello existen diversos métodos que pasamos a describir.

5.9.1. Cálculo directo

Las propiedades halladas para la resultante permiten calcular directamente el discriminante de polinomios particulares. Veamos algunos ejemplos:

1. Ejemplo: $f = X^n - 1 = \prod_{i=1}^n (X - \alpha_i)$, $f' = nX^{n-1}$

$$\begin{aligned} D(f) &= (-1)^{\frac{n(n-1)}{2}} R(f, f') = (-1)^{\frac{n(n-1)}{2}} \prod_{i=1}^n f'(\alpha_i) = (-1)^{\frac{n(n-1)}{2}} \prod_{i=1}^n n\alpha_i^{n-1} = \\ &= (-1)^{\frac{n(n-1)}{2}} n^n \left(\prod_{i=1}^n \alpha_i \right)^{n-1} = (-1)^{\frac{n(n-1)}{2} + n(n-1)} n^n = (-1)^{\frac{n(n-1)}{2}} n^n \end{aligned}$$

En particular si q es impar, $f = X^q - 1$, $D(f) = (-1)^{\frac{q-1}{2}} q^q$

2. Ejemplo: $f = X^{p-1} + X^{p-2} + \dots + X + 1$ p primo impar. Sea $g = X - 1$.
Entonces $fg = X^p - 1 \Rightarrow D(fg) = (-1)^{\frac{p-1}{2}} p^p$
 $g' = 1 \Rightarrow D(g) = R(g, g') = 1$ $R(f, g) = f(1) = p$.
Luego $D(fg) = D(f)D(g)R(f, g)^2 \Rightarrow D(f) = (-1)^{\frac{p-1}{2}} p^{p-2}$

3. Ejemplo:

$$\begin{aligned} f &= X^3 + aX + b = (X - \alpha_1)(X - \alpha_2)(X - \alpha_3) \\ f' &= 3X^2 + a \quad D(f) = -R(f, f') \\ R(f, f') &= \prod_{i=1}^3 (3\alpha_i^2 + a) = \prod_{i=1}^3 f'(\alpha_i) \end{aligned}$$

$$\text{Pero } f'(\alpha_i) = 3\alpha_i^2 + a = \frac{3\alpha_i^3 + a\alpha_i}{\alpha_i} = \frac{-2a\alpha_i - 3b}{\alpha_i}$$

$$\begin{aligned} \text{Llamamos } \beta_i &= 2a\alpha_i + 3b \Rightarrow \alpha_i = \frac{\beta_i - 3b}{2a} \text{ así que } \beta_i \text{ es raíz de } \left(\frac{X-3b}{2a}\right)^3 + a\frac{X-3b}{2a} + b \\ \Rightarrow \beta_1\beta_2\beta_3 &= 8a^3\left(\frac{27b^3}{8a^3} + \frac{3b}{2} - b\right) = 27b^3 + 4a^3b \end{aligned}$$

$$R(f, f') = \prod_{i=1}^3 f'(\alpha_i) = -\frac{\prod_{i=1}^3 (2a\alpha_i + 3b)}{\prod_{i=1}^3 \alpha_i} = -\frac{27b^3 + 4a^3b}{-b} = 27b^2 + 4a^3$$

y por tanto $D(f) = -(4a^3 + 27b^2)$

4. Ejemplo:

$$f = X^3 + aX^2 + b = (X - \alpha_1)(X - \alpha_2)(X - \alpha_3)$$

$$f' = 3X^2 + 2aX = X(3X + 2a) \quad f'(\alpha_i) = \alpha_i(3\alpha_i + 2a)$$

Sea $\beta_i = 3\alpha_i + 2a \Rightarrow \alpha_i = \frac{\beta_i - 2a}{3}$ y los β_i son raíces de $(\frac{X-2a}{3})^3 + a(\frac{X-2a}{3})^2 + b$
 $\Rightarrow \beta_1\beta_2\beta_3 = 3^3((\frac{2a}{3})^3 - a(\frac{2a}{3})^2 - b) = -(4a^3 + 27b)$

Luego $R(f, f') = \prod_{i=1}^3 f'(\alpha_i) = \prod_{i=1}^3 \alpha_i \prod_{i=1}^3 \beta_i = (-b)(-(4a^3 + 27b))$
y $D(f) = -R(f, f') = -b(4a^3 + 27b)$

5. Ejemplo:

$$f = X^5 + aX + b = \prod_{i=1}^5 (X - \alpha_i) \quad f' = 5X^4 + a$$

$$f'(\alpha_i) = 5\alpha_i^4 + a = \frac{5\alpha_i^5 + a\alpha_i}{\alpha_i} = \frac{-4a\alpha_i - 5b}{\alpha_i}$$

Llamamos $\beta_i = 4a\alpha_i + 5b \Rightarrow \alpha_i = \frac{\beta_i - 5b}{4a}$ así que β_i es raíz de $(\frac{X-5b}{4a})^5 + a(\frac{X-5b}{4a}) + b$
 $\Rightarrow \prod_{i=1}^5 \beta_i = (4a)^5((\frac{5b}{4a})^5 + \frac{5b}{4} - b) = (5b)^5 + 4^4 a^5 b$

$$D(f) = R(f, f') = \prod_{i=1}^5 f'(\alpha_i) = -\frac{\prod_{i=1}^5 \beta_i}{\prod_{i=1}^5 \alpha_i} = 5^5 b^4 + 4^4 a^5$$

6. Ejemplo:

$$f = X^5 + aX^4 + b = \prod_{i=1}^5 (X - \alpha_i)$$

$$f' = 5X^4 + 4aX^3 = X^3(5X + 4a) \quad f'(\alpha_i) = \alpha_i^3(5\alpha_i + 4a)$$

Sea $\beta_i = 5\alpha_i + 4a \Rightarrow \alpha_i = \frac{\beta_i - 4a}{5}$ y los β_i son raíces de $(\frac{X-4a}{5})^5 + a(\frac{X-4a}{5})^4 + b$
 $\Rightarrow \prod_{i=1}^5 \beta_i = 5^5((\frac{4a}{5})^5 - a(\frac{4a}{5})^4 - b) = -(4^4 a^5 + 5^5 b)$

Luego $D(f) = R(f, f') = \prod_{i=1}^5 f'(\alpha_i) = \prod_{i=1}^5 \alpha_i^3 \prod_{i=1}^5 \beta_i = b^3(4^4 a^5 + 5^5 b)$

7. Ejemplo:

$$f = X^n + aX + b = \prod_{i=1}^n (X - \alpha_i) \quad f' = nX^{n-1} + a$$

$$f'(\alpha_i) = n\alpha_i^{n-1} + a = \frac{n\alpha_i^n + a\alpha_i}{\alpha_i} = \frac{-(n-1)a\alpha_i - nb}{\alpha_i}$$

Llamamos $\beta_i = (n-1)a\alpha_i + nb \Rightarrow \alpha_i = \frac{\beta_i - nb}{(n-1)a}$ así que β_i es raíz de $(\frac{X-nb}{(n-1)a})^n + a(\frac{X-nb}{(n-1)a}) + b \Rightarrow \prod_{i=1}^n \beta_i = (-1)^n((n-1)a)^n(((n-1)a)^n - \frac{nb}{n-1} + b) = (nb)^n + (-1)^n(n-1)^{n-1}a^n(-b)$

$$R(f, f') = \prod_{i=1}^n f'(\alpha_i) = (-1)^n \frac{\prod_{i=1}^n \beta_i}{\prod_{i=1}^n \alpha_i} = n^n b^{n-1} + (-1)^{n-1}(n-1)^{n-1}a^n$$

$$D(f) = (-1)^{\frac{n(n-1)}{2}} R(f, f') = (-1)^{\frac{n(n-1)}{2}} (n^n b^{n-1} + (-1)^{n-1}(n-1)^{n-1}a^n)$$

8. Ejemplo:

$$f = X^n + aX^{n-1} + b = \prod_{i=1}^n (X - \alpha_i) \quad f' = nX^{n-1} + (n-1)aX^{n-2} = X^{n-2}(nX + (n-1)a) \quad f'(\alpha_i) = \alpha_i^{n-2}(n\alpha_i + (n-1)a)$$

Sea $\beta_i = n\alpha_i + (n-1)a \Rightarrow \alpha_i = \frac{\beta_i - (n-1)a}{n}$ y los β_i son raíces de $(\frac{X-(n-1)a}{n})^n + a(\frac{X-(n-1)a}{n})^{n-1} + b \Rightarrow \prod_{i=1}^n \beta_i = (-1)^n(n^n((-\frac{(n-1)a}{n})^n + a(-\frac{(n-1)a}{n})^{n-1} - b)) = -(n-1)^{n-1}a^n + (-1)^n n^n b$

Luego $R(f, f') = \prod_{i=1}^n f'(\alpha_i) = \prod_{i=1}^n \alpha_i^{n-2} \prod_{i=1}^n \beta_i = (-b)^{n-2}(-(n-1)^{n-1}a^n + (-n)^n b)$
y $D(f) = (-1)^{\frac{n(n-1)}{2}} R(f, f') = (-1)^{\frac{(n-1)(n+2)}{2}} b^{n-2}((n-1)^{n-1}a^n + (-1)^{n-1}n^n b)$

5.9.2. Método modular

A partir de la propiedad 7 de la resultante puede desarrollarse un método muy económico para el cálculo de la resultante de algunos pares especiales de polinomios: En primer lugar, sean

$$f = a_n X^n + \dots + a_0 \quad g = X - b$$

Dividiendo f entre g obtenemos:

$$f = gf_1 + f(b)$$

Por las propiedades de la resultante obtenemos:

$$R(f, g) = (-1)^n R(g, f) = (-1)^n R(g, f(b)) = (-1)^n f(b) \quad (5.7)$$

Sean ahora

$$f = a_n X^n + \dots + a_0 \quad g = b_m X^m + \dots + b_0$$

y sean p, q_i, r, s_j tales que

$$pg \prod_{i=1}^k (X - q_i) \equiv r \prod_{j=1}^l (X - s_j) \pmod{f} \quad (5.8)$$

Entonces

$$R(f, p)R(f, g)R(f, \prod_{i=1}^k (X - q_i)) = a_n^{m+k-l} R(f, r)R(f, \prod_{j=1}^l (X - s_j)) \quad (5.9)$$

Pero por (5.7)

$$\begin{aligned} R(f, p) &= p^n & R(f, r) &= r^n \\ R(f, \prod_{i=1}^k (X - q_i)) &= \prod_{i=1}^k R(f, X - q_i) = \prod_{i=1}^k (-1)^n f(q_i) \\ R(f, \prod_{j=1}^l (X - s_j)) &= \prod_{j=1}^l R(f, X - s_j) = \prod_{j=1}^l (-1)^n f(s_j) \end{aligned}$$

Despejando en (5.9),

$$R(f, g) = (-1)^{n(k+l)} a_n^{m+k-l} \frac{r^n \prod_{j=1}^l f(s_j)}{p^n \prod_{i=1}^k f(q_i)}$$

Ejemplo: Sean

$$f = X^5 - X^2 + 15 \quad g = f' = 5X^4 - 2X$$

Tomamos $k = 1, l = 2, p_1 = 1, p_0 = 0$ y calculamos:

$$Xg = 5X^5 - 2X^2 = 5f + 3X^2 - 75 \equiv 3(X - 5)(X + 5) \pmod{f}$$

$$\begin{aligned} D(f) = R(f, f') &= (-1)^{5(1+2)} \frac{3^5 f(5) f(-5)}{f(0)} = \\ &= -3^5 \frac{(5^5 - 5^2 + 15)((-5)^5 - (-5)^2 + 15)}{15} = -3^5 \frac{10^2 - 5^{10}}{15} = 3^4 5(5^8 - 4) \end{aligned}$$

5.9.3. Por el algoritmo de Euclides

Dividiendo g por f obtenemos $g = fq + r$ con $\text{gr}(r) < \text{gr}(f)$. Por las propiedades 7 y 2,

$$R(f, g) = a_n^{m-\text{gr}(r)} R(f, r) = (-1)^{m\text{gr}(r)} a_n^{m-\text{gr}(r)} R(r, f)$$

Por inducción sobre el grado llegamos a $\text{gr}(r) = 0$ y aplicamos la propiedad 5.

1. Ejemplo:

$$\begin{aligned} f &= aX + b \\ f' &= a \\ R(f, f') &= a \\ D(f) &= 1 \end{aligned}$$

2. Ejemplo:

$$\begin{aligned} f &= aX^2 + bX + c \\ f' &= 2aX + b \\ f &= \left(\frac{1}{2}X + \frac{b}{4a}\right)f' + \left(c - \frac{b^2}{4a}\right) \\ R(f, f') &= R(f', f) = (2a)^2 R(f', c - \frac{b^2}{4a}) = (2a)^2 \left(c - \frac{b^2}{4a}\right) = a(4ac - b^2) \\ D(f) &= (-1)^{\frac{2-1}{2}} \frac{1}{a} R(f, f') = b^2 - 4ac \end{aligned}$$

3. Ejemplo:

$$\begin{aligned} f &= X^3 + aX + b \\ f' &= 3X^2 + a \\ f &= \frac{1}{2}Xf' + r \quad r = \frac{2a}{3}X + b \\ f' &= \left(\frac{9}{2a}X - \frac{27b}{4a^2}\right)r + r_1 \quad r_1 = \frac{27b^2 + 4a^3}{4a^2} \\ R(f, f') &= R(f', f) = 3^2 R(f', r) = 3^2 R(r, f') = 3^2 \left(\frac{2a}{3}\right)^2 R(r, r_1) = 4a^2 \frac{27b^2 + 4a^3}{4a^2} \\ D(f) &= -R(f, f') = -(4a^3 + 27b^2) \end{aligned}$$

5.9.4. Determinante de Euler-Sylvester-Cayley

Multiplicando f sucesivamente por $1, X, \dots, X^{m-1}$ y g por $1, X, X^{n-1}$ e igualando a cero nos queda el siguiente sistema de $(n + m)$ ecuaciones en las $(n + m)$ incógnitas $1, X, X^2, \dots, X^{n+m-1}$:

$$\begin{array}{cccccccccccccc} X^{m-1}f & = & a_n X^{n+m-1} & + & a_{n-1} X^{n+m-2} & + & \dots & + & a_0 X^{m-1} & & & & & = & 0 \\ X^{m-2}f & = & & & a_n X^{n+m-2} & + & \dots & + & a_1 X^{m-1} & + & a_0 X^{m-2} & & & = & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1f & = & & & & & & & a_n X^n & + & \dots & + & a_0 & = & 0 \\ X^{n-1}g & = & b_m X^{n+m-1} & + & b_{m-1} X^{n+m-2} & + & \dots & + & b_0 X^{n-1} & & & & & = & 0 \\ X^{n-2}g & = & & & b_m X^{n+m-2} & + & \dots & + & b_1 X^{n-1} & + & b_0 X^{n-2} & & & = & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1g & = & & & & & & & b_m X^m & + & \dots & + & b_0 & = & 0 \end{array}$$

Por el teorema de Rouché, este sistema tendrá solución si y sólo si el determinante de los coeficientes es cero. Este determinante se llama *resultante de*

Euler-Sylvester-Cayley:

$$C(f, g) = \begin{vmatrix} a_n & a_{n-1} & \dots & a_0 & 0 & \dots & 0 \\ 0 & a_n & a_{n-1} & \dots & a_0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_n & a_{n-1} & \dots & a_0 \\ b_m & b_{m-1} & \dots & b_0 & 0 & \dots & 0 \\ 0 & b_m & b_{m-1} & \dots & b_0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & b_m & b_{m-1} & \dots & b_0 \end{vmatrix}$$

Vamos a ver que $C(f, g) = R(f, g)$: Como $C(f, g) = C(a_n, \dots, a_0, b_m, \dots, b_0)$ y los a_i y b_j son polinomios simétricos en α_i y β_j respectivamente, obtenemos que $C(f, g)$ es un polinomio simétrico en α_i y β_j . Por otra parte, si f y g tienen una raíz común, el anterior sistema lineal tiene solución, luego $\forall i, j (\alpha_i - \beta_j) \mid C(f, g) \Rightarrow R(f, g) \mid C(f, g)$. Contando grados vemos que el cociente tiene grado cero (i.e. es una constante). Luego $C(f, g) = \lambda R(f, g)$. Pero el término $a_n^m b_m^n$ aparece con coeficiente +1 en $C(f, g)$ y en $R(f, g) \Rightarrow \lambda = 1$.

1. Ejemplo: Tomando $f = aX^2 + bX + c, g = f' = 2aX + b$ tenemos:

$$R(f, g) = \begin{vmatrix} a & b & c \\ 2a & b & 0 \\ 0 & 2a & b \end{vmatrix} = ab^2 + 4a^2c - 2ab^2 = a(4ac - b^2)$$

2. Ejemplo: $f = X^3 + aX + b, g = 3X^2 + a$

$$R(f, g) = \begin{vmatrix} 1 & 0 & a & b & 0 \\ 0 & 1 & 0 & a & b \\ 3 & 0 & a & 0 & 0 \\ 0 & 3 & 0 & a & 0 \\ 0 & 0 & 3 & 0 & a \end{vmatrix} = 4a^3 + 27b^2$$

5.9.5. Determinante de Bezout

La resultante de Cayley proporciona una expresión sencilla y elegante para $R(f, g)$. Sin embargo, el orden del determinante es $(m + n)$, muy alto para los cálculos prácticos. Vamos a desarrollar otro método basado en la misma idea pero donde el determinante que va a aparecer es de orden $\max(m, n)$. En primer lugar consideramos $m = n$, o sea que f y g son del mismo grado. Definimos los elementos:

$$c_{ij} = \begin{cases} a_j b_i - a_i b_j & \text{si } 0 \leq i, j \leq n \\ 0 & \text{en otro caso} \end{cases}$$

Obsérvese que $c_{ij} = -c_{ji}$ y que $c_{ii} = 0$.

Si todos los c_{ij} son cero, existe un λ tal que $g = \lambda f$. En lo que sigue excluimos este caso. Consideremos ahora los polinomios:

$$h_i = b_i f - a_i g \quad i = 0, 1, \dots, n \quad (5.10)$$

Si $c_{ij} \neq 0$, del sistema:

$$\begin{aligned} h_i &= b_i f - a_i g \\ h_j &= b_j f - a_j g \end{aligned}$$

obtenemos:

$$\begin{aligned} f &= \frac{a_j}{c_{ij}} h_i - \frac{a_i}{c_{ij}} h_j \\ g &= \frac{b_j}{c_{ij}} h_i - \frac{b_i}{c_{ij}} h_j \end{aligned}$$

luego h_i, h_j tienen un cero en común si y sólo si f y g tienen un cero en común, y las raíces comunes de f y g son precisamente las raíces comunes a todos los h_i .

Formemos ahora los polinomios:

$$\begin{aligned} g_0 &= h_n &= b_n f - a_n g &= \sum_i^{n-1} d_{0i} X^i \\ g_1 &= x g_0 + h_{n-1} &= (b_n X + b_{n-1}) f - (a_n X + a_{n-1}) g &= \sum_i^{n-1} d_{1i} X^i \\ g_2 &= x g_1 + h_{n-2} &= (b_n X^2 + b_{n-1} X + b_{n-2}) f - (a_n X^2 + a_{n-1} X + a_{n-2}) g &= \sum_i^{n-1} d_{2i} X^i \\ &\vdots &\vdots &\vdots \\ g_{n-1} &= x g_{n-2} + h_1 &= (b_n X^{n-1} + \dots + b_1) f - (a_n X^{n-1} + \dots + a_1) g &= \sum_i^{n-1} d_{n-1i} X^i \end{aligned} \quad (5.11)$$

Los g_i tienen un cero en común \Leftrightarrow los h_i tienen un cero en común $\Leftrightarrow f$ y g tienen un cero en común. Veamos la forma general de los coeficientes d_{ki} . Por construcción,

$d_{0i} = c_{ni}$, $d_{ki} = d_{k-1,i-1} + c_{n-k,i}$. Demostraremos por inducción sobre k que

$$d_{ki} = \sum_{j=0}^k c_{n-j,i-k+j} \quad (5.12)$$

Para $k = 0$ es trivial. Supongámoslo cierto para $k - 1$. Entonces

$$d_{ki} = d_{k-1,i-1} + c_{n-k,i} = \sum_{j=0}^{k-1} c_{n-j,i+j-k} + c_{n-k,i}$$

Las raíces comunes de f y g dan lugar a soluciones no triviales del sistema:

$$\begin{aligned} d_{0n-1}X^{n-1} + \dots + d_{01}X + d_{00}1 &= 0 \\ d_{1n-1}X^{n-1} + \dots + d_{11}X + d_{10}1 &= 0 \\ &\vdots \\ d_{n-1,n-1}X^{n-1} + \dots + d_{n-1,1}X + d_{n-1,0}1 &= 0 \end{aligned}$$

Llamamos *resultante de Bezout de f y g* al determinante de este sistema:

$$B(f, g) = \begin{vmatrix} d_{0n-1} & \dots & d_{00} \\ \vdots & \vdots & \vdots \\ d_{n-1,n-1} & \dots & d_{n-1,0} \end{vmatrix}$$

Como cada c_{ij} es homogéneo de grado 1 en a_i y en b_j , d_{ki} también es homogéneo de grado 1 en ambos, y $B(f, g)$ es un polinomio homogéneo en las a_i y en las b_j de grado $2n$. Igual que para la resultante de Cayley, $B(f, g)$ es cero cuando f y g tienen una raíz en común, luego $B(f, g) = \lambda R(f, g)$ y contando grados, $\lambda \in F$.

Para determinar λ observamos el término $a_n^n b_0^n$. En la resultante de Cayley este término sólo aparece en el desarrollo de la diagonal principal y por tanto tiene coeficiente +1. En $B(f, g)$ aparece en el producto de todos los $c_{0n} = a_n b_0 - a_0 b_n$ de la diagonal secundaria, luego tiene coeficiente $\text{sgn}(\sigma)$ siendo $\sigma = (1 \ n)(2 \ n-1) \dots$ luego $\text{sgn}(\sigma) = (-1)^{\frac{n(n-1)}{2}}$ y por tanto $\lambda = (-1)^{\frac{n(n-1)}{2}}$, $B(f, g) = (-1)^{\frac{n(n-1)}{2}} R(f, g)$.

En caso de que $\text{gr}(g) = m \leq \text{gr}(f) = n$, tomamos $g_1 = X^{n-m}g$, formamos la resultante de Bezout de f y g_1 y utilizamos la propiedad 8 de $R(f, g)$:

$$B(f, X^{n-m}g) = (-1)^{\frac{n(n-1)}{2}} R(f, X^{n-m}g) = (-1)^{\frac{n(n-1)}{2} + n(n-m)} a_0^{n-m} R(f, g)$$

Para calcular el discriminante de un polinomio, $g = f'$, $m = n - 1$ y nos queda:

$$B(f, Xf') = (-1)^{\frac{n(n-1)}{2} + n} a_0 R(f, f') = (-1)^n a_0 a_n D(f)$$

así que

$$D(f) = \frac{(-1)^n}{a_n a_0} B(f, Xf')$$

además, en este caso los c_{ij} tienen una forma sencilla: Sean

$$\begin{aligned} f &= a_n X^n + a_{n-1} X^{n-1} + \dots + a_0 \\ g &= Xf' = n a_n X^n + \dots + a_1 X = b_n X^n + \dots + b_1 \end{aligned}$$

luego $b_i = i a_i$, $c_{ij} = (j - i) a_i a_j$, y

$$d_{ki} = \sum_{j=0}^k c_{n-j, i+j-k} = \sum_{j=0}^k (i + 2j - k - n) a_{n-j} a_{i+j-k} = - \sum_{j=0}^k (n - i + k - 2j) a_{n-j} a_{i+j-k}$$

y nos queda la expresión:

$$D(f) = \frac{1}{a_n a_0} \begin{vmatrix} -d_{0,n-1} & \dots & -d_{0,0} \\ \vdots & \ddots & \vdots \\ -d_{n-1,n-1} & \dots & -d_{n-1,0} \end{vmatrix}$$

1. Ejemplo: $f = aX^2 + bX + c$

$$D(f) = \frac{1}{ac} \begin{vmatrix} ab & 2ac \\ 2ac & bc \end{vmatrix} = b^2 - 4ac$$

2. Ejemplo: $f = X^3 + aX + b$

$$D(f) = \frac{1}{b} \begin{vmatrix} 0 & 2a & 3b \\ 2a & 3b & 0 \\ 3b & 0 & ab \end{vmatrix} = -(4a^3 + 27b^2)$$

3. Ejemplo: $f = a_3X^3 + a_2X^2 + a_1X + a_0$

$$D(f) = \frac{1}{a_3 a_0} \begin{vmatrix} a_3 a_2 & 2a_3 a_1 & 3a_3 a_0 \\ 2a_3 a_1 & 3a_3 a_0 + a_2 a_1 & 2a_2 a_0 \\ 3a_3 a_0 & 2a_2 a_0 & a_1 a_0 \end{vmatrix}$$

4. Ejemplo: $f = a_4X^4 + a_3X^3 + a_2X^2 + a_1X + a_0$

$$D(f) = \frac{1}{a_4 a_0} \begin{vmatrix} a_4 a_3 & 2a_4 a_2 & 3a_4 a_1 & 4a_4 a_0 \\ 2a_4 a_2 & 3a_4 a_1 + a_3 a_2 & 4a_4 a_0 + 2a_3 a_1 & 3a_3 a_0 \\ 3a_4 a_1 & 4a_4 a_0 + 2a_3 a_1 & 3a_3 a_0 + a_2 a_1 & 2a_2 a_0 \\ 4a_4 a_0 & 3a_3 a_0 & 2a_2 a_0 & a_1 a_0 \end{vmatrix}$$

5. Ejemplo: $f = a_5X^5 + a_4X^4 + a_3X^3 + a_2X^2 + a_1X + a_0$

$$D(f) = \frac{1}{a_5 a_0} \begin{vmatrix} a_5 a_4 & 2a_5 a_3 & 3a_5 a_2 & 4a_5 a_1 & 5a_5 a_0 \\ 2a_5 a_3 & 3a_5 a_2 + a_4 a_3 & 4a_5 a_1 + 2a_4 a_2 & 5a_5 a_0 + 3a_4 a_1 & 4a_4 a_0 \\ 3a_5 a_2 & 4a_5 a_1 + 2a_4 a_2 & 5a_5 a_0 + 3a_4 a_2 + a_3 a_2 & 4a_4 a_0 + 2a_3 a_1 & 3a_3 a_0 \\ 4a_5 a_1 & 5a_5 a_0 + 3a_4 a_1 & 4a_4 a_0 + 2a_3 a_1 & 3a_3 a_0 + a_2 a_1 & 2a_2 a_0 \\ 5a_5 a_0 & 4a_4 a_0 & 3a_3 a_0 & 2a_2 a_0 & a_1 a_0 \end{vmatrix}$$

Referencias

- [1] J. A. Beachy and W. D. Blair, *Abstract Algebra*, Waveland Press 1996
- [2] Z. I. Borevich and I. R. Shafarevich, *Number Theory*, Academic Press 1966
- [3] L. Childs, *A Concrete Introduction to Higher Algebra*, Springer 1979
- [4] H. Cohn, *A Classical Invitation to Algebraic Numbers and Class Fields*, Springer 1978
- [5] P. M. Cohn, *Classic Algebra*, Wiley and sons 2000
- [6] D. S. Dummit and R. M. Foote, *Abstract Algebra* 2nd ed., Prentice-Hall 1999
- [7] J. A. Gallian, *Contemporary Abstract Algebra* 5th edition, Houghton Mifflin 2002
- [8] D. E. Knuth, *The Art of Computer Programming, vol. 2: Seminumerical Algorithms* 3rd edition, Addison-Wesley 1998
- [9] S. Lang, *Algebra* 3rd edition, Addison-Wesley 1993
- [10] I. Niven, H. S. Zuckerman and H. L. Montgomery, *An Introduction to the Theory Of Numbers*, Wiley 1991