Universidad de Valladolid

Facultad de Ciencias

Trabajo de Fin de Máster

Máster en Matemáticas

Polinomios sobre anillos de división

Realizado por: Miguel Herreros Gaona

Tutelado por: Jose Ramón Brox López

Curso 2024-2025

Índice general

Introducción

La teoría sobre los polinomios con coeficientes en un cuerpo está ampliamente estudiada y es muy diversa, tanto en cuanto a las áreas en las que se basa, como en las áreas que se basan en ella como herramienta para probar determinados resultados. Sin embargo, ¿qué sucede cuándo los coeficientes no son conmutativos? ¿Cómo definimos qué es un polinomio entonces? ¿Teoremas tan importantes como el teorema del factor o el teorema fundamental del álgebra se mantienen? ¿Cómo son las raíces?

En este Trabajo de Fin de Máster se busca hacer una recopilación y explicación de los principales resultados de la teoría de polinomios sobre anillos de división y ahondaremos en los polinomios definidos sobre los cuaternios, el anillo de división más fácilmente reconocible.

Empezaremos el capítulo 1 haciendo un repaso de teoría de grupos, introduciendo qué son los anillos de división y demostrando las principales propiedades de los mismos, en particular, demostraremos el pequeño teorema de Wedderburn. En el capítulo 2, definiremos los dos tipos polinomios, a izquierda y generales, que se pueden definir sobre los mismos. Para los polinomios a izquierda veremos la generalización del teorema del factor y los dos teoremas de Gordon-Motzkin, que generalizan el teorema fundamental del álgebra a los polinomios a izquierda. Por otro lado, para los polinomios generales, daremos su construcción y los dos principales resultados, también de Gordon y Motzkin, acerca de la existencia de polinomios de cierto grado, n, con un determinado número de soluciones, h, cuya cota está determinada por el grado. Por último, en el capítulo 3 realizaremos un estudio más extenso sobre los polinomios a izquierda (o, equivalentemente, a derecha) en los cuaternios, pues el teorema de Niven nos dice que un anillo de división centralmente finito es algebraicamente cerrado para este tipo de polinomios si y solo si es el álgebra de cuaternios sobre algún cuerpo real cerrado. A continuación expondremos el método de Janovská-Opfer para calcular las raíces de estos polinomios y cerraremos el capítulo clasificando las raíces de un cierto subconjunto de los polinomios generales sobre los cuaternios.

Capítulo 1

Preliminares.

1.1. Teoría de grupos.

Vamos a empezar el trabajo haciendo una pequeña introducción a la teoría de grupos finitos con los principales resultados técnicos que usaremos más adelante.

Definición 1.1.1. Sea G grupo finito y $S \leq G$. Llamaremos *índice de S en G* al número de clases laterales a izquierda (resp. a derecha) de S en G y lo denotaremos por $[G : S]$.

Teorema 1.1.2 (Lagrange)**.** *Sea G un grupo finito y $S \leq G$. Entonces*

$$[G : S] = \frac{|G|}{|S|}.$$

Demostración. Notemos que cada elemento de G pertenece a única clase lateral con respecto a S, como hay exactamente $[G : S]$ clases laterales y todas ellas tienen el mismo número de elementos, que además es igual a $|S|$, tenemos que:

$$|G| = [G : S] \cdot |S|. \qquad \square$$

Definición 1.1.3. Sea G un grupo y S un conjunto. Una acción de grupo de G en S es un homomorfismo $f : G \to Perm(S)$ de G al grupo de permutaciones de S. Este homomorfismo nos permite asociar a cada elemento $g \in G$, una permutación $f_g \in Perm(S)$, luego dado $s \in S$, tenemos que $f_g(s) \in S$. Definimos $g \cdot s := f_g(s)$. Con esta notación abreviada, como f es homomorfismo de grupos, tenemos las siguientes propiedades:

- Para todo $g, h \in G$ y $s \in S$, tenemos que $g(hs) = f_g(f_h(s)) = (f_g \circ f_h)(s) = f_{gh}(s) = (gh)s$.
- Dado $e \in G$ el elemento neutro, tenemos que $\forall s \in S$, $es = f_e(s) = id(s) = s$.

Con esta notación multiplicativa, a veces no explicitaremos el homomorfismo f y simplemente diremos que G *actúa sobre* S.

Definición 1.1.4. Sea G un grupo finito que actúa sobre un conjunto S. Dado $s \in S$, llamaremos *órbita de s* bajo la acción de G al subconjunto:

$$orb_G(s) := \{gs \,|\, g \in G\} \subset S.$$

De forma similar, llamaremos *estabilizador de s* bajo la acción de G al subgrupo:

$$stab_G(s) := \{g \in G \,|\, gs = s\} \leq G.$$

Si no da lugar a confusión, denotaremos las órbitas y los estabilizadores sin los subíndices.

Teorema 1.1.5 (Teorema de órbita-estabilizador)**.** *Sea G un grupo finito que actúa sobre un conjunto finito S. Entonces tenemos que, para todo $s \in S$:*

$$|orb(s)| = [G : stab(s)] = \frac{|G|}{|stab(s)|}.$$

Demostración. Fijemos $s \in S$. Sea $\phi : G \to orb(s)$, $\phi(g) = gs$. Notemos que ϕ es claramente sobreyectiva. Sean $g, h \in G$ tales que $\phi(g) = \phi(h)$, es decir, $gs = hs \iff s = g^{-1}hs \iff g^{-1}h \in stab(s)$. Luego tenemos que $g \equiv h \pmod{stab(s)}$, por lo tanto, $\tilde{\phi} : G/stab(s) \to orb(s)$, $\tilde{\phi}([g]) \mapsto gs$ es una biyección que está bien definida por las consideraciones anteriores. Por lo tanto:

$$|orb(s)| = |G/stab(s)| = [G : stab(s)].$$

Aplicando el teorema de Lagrange 1.1.2 obtenemos la última igualdad. □

Definición 1.1.6. Llamaremos *conjugación* a la siguiente acción de G en G.

$$\begin{aligned} c : G \times G &\to G \\ (g, h) &\mapsto c_g(h) = ghg^{-1} \end{aligned}$$

Definición 1.1.7. Sea G un grupo y $S \leq G$. Definimos el *centralizador de S en G* como $C_G(S) := \{g \in G \,|\, gs = sg,\ \forall s \in S\}$.

A $Z(G) := C_G(G) = \{g \in G \,|\, gg' = g'g,\ \forall g' \in G\}$ lo llamaremos *centro de G*.

Observación. A los centralizadores de la forma $C_G(\{x\})$ los denotaremos por $C_G(x)$ y notemos que estos se corresponden con los estabilizadores, $stab(x)$, bajo la acción de conjugación.

Al igual que con las órbitas y estabilizadores, si no da lugar a confusión, omitiremos los subíndices de los centralizadores.

Proposición 1.1.8 (Ecuación de clase)**.** *Sea G un grupo finito. Sea m el número de clases de conjugación de G con más de un elemento y x_i, $\forall i = 1, \ldots, m$ un representante de cada clase. Tenemos que:*

$$|G| = |Z(G)| + \sum_{i=1}^{m} [G : C(x_i)].$$

Demostración. Notemos que las clases de conjugación son las órbitas bajo la acción de conjugación y que si, dado $y \in G$, $|orb(y)| = 1$, entonces $y \in Z(G)$. Por lo tanto, usando el teorema 1.1.5 tenemos que:

$$|G| = |Z(G)| + \sum_{i=1}^{m} orb(x_i) = |Z(G)| + \sum_{i=1}^{m} [G : C(x_i)]. \qquad \square$$

Proposición 1.1.9. *Sea $(G, \cdot)$ un grupo y $S \subset G$ un subconjunto finito de G tal que $(S, \cdot)$ es un semigrupo. Entonces S es un subgrupo de G.*

Demostración. Notemos que, al ser G un grupo, S hereda que $\forall a, b, c \in S \subset G$ se cumple que $ab = ac \implies b = c$

Consideremos ahora, dado $a \in S$ la aplicación $\mu_a : S \to S$ tal que $\mu_a(s) = as$. Tenemos que μ_a es una biyección al ser inyectiva por 1.1 y S finito. Esto nos lleva a que $\exists b \in S \subset G$ tal que $\mu_a(b) = ab = a$ y, como $b \in G$, tenemos que b es precisamente el elemento neutro de G, luego $e = b \in S$. Por otro lado, esto nos lleva a que a que $\exists c \in S \subset G$ tal que $\mu_a(c) = ac = e$, que por unicidad del inverso en G tenemos que $c = a^{-1} \in S$. Por tanto S es un semigrupo con elemento neutro e inversos para todos sus elementos, es decir, un grupo. $\square$

1.2. Anillos de división.

El objeto de estudio de este trabajo son los polinomios definidos sobre un anillo de división, pero antes de ello, debemos definir qué es un anillo de división y qué propiedades tiene. Uno de los libros más influyentes en la teoría de anillos de división es *Skew fields: theory of general division rings*, de P.M. Cohn (ver [Coh95]). Sin embargo, para este trabajo, hemos decidido usar como bibliografía básica para los anillos de división *A first course in noncommutative rings*, de T.Y. Lam (ver [Lam01]), pues se trata de un libro más didáctico que recopila los principales resultados sobre anillos de división (sección §13) y sobre un tipo de polinomios que se pueden definir sobre los mismos (sección §16), todos ellos con las demostraciones actualizadas y simplificadas.

Definición 1.2.1 (Anillo de división)**.** Sea $(D, +, \cdot)$ un anillo unitario. Diremos que D es un *anillo de división* si se cumple que para todo $x \in D$ no nulo existe un único elemento x^{-1} al que llamaremos *elemento inverso* tal que $x \cdot x^{-1} = x^{-1} \cdot x = 1$, donde 1 es la unidad de D.

Definición 1.2.2. Sea D un anillo de división y $S \subset D$. Llamaremos *centralizador de S en D* al conjunto

$$C_D(S) := \{x \in D \,|\, xs = sx, \forall s \in S\}.$$

Proposición 1.2.3. *Dado $S \subset D$, el centralizador de S, $C_D(S)$, es un subanillo de división.*

Demostración. Sean $x, y \in C_D(S)$ y $s \in S$. Notemos que

$$(x+y)s = xs + ys = sx + sy = s(x+y) \implies x+y \in C_D(S)$$

$$(xy)s = xsy = s(xy) \implies (xy) \in C_D(S)$$

$$xs = sx \implies x^{-1}xsx^{-1} = x^{-1}sxx^{-1} \implies sx^{-1} = x^{-1}s \implies x^{-1} \in C_D(S)$$

Por otro lado, $0, 1 \in C_D(D) \subset C_D(S)$. Luego $C_D(S)$ es un subanillo de división de D. □

Proposición 1.2.4. *Sea $Z(D) := C_D(D) = \{x \in D \,|\, xy = yx,\ \forall y \in D\}$. Entonces $Z(D)$ es un subanillo de división de D, que además es un cuerpo, al que llamaremos centro de D.*

Nota: En el caso de que $S = \{s\}$, hablaremos del centralizador de s en D y lo denotaremos por $C_D(s)$

Demostración. Sabemos por 1.2.3 que $Z(D)$ es un subanillo de división. Como además, $Z(D)$ es conmutativo por definición, tenemos que $Z(D)$ es un cuerpo. □

Nota: Notemos que D es un espacio vectorial sobre $Z(D)$ con el producto heredado de D. De hecho, es una $Z(D)$-álgebra. Ambos hechos se usarán como herramienta en múltiples demostraciones a lo largo del trabajo

---

Definición 1.2.5. Diremos que D es *centralmente finito* si es un $Z(D)$-espacio vectorial de dimensión finita y *centralmente infinito* en caso contrario.

Definición 1.2.6. Sea $\Xi_n = \{\xi \in \mathbb{C} \,|\, \xi^n = 1, \xi^m \neq 1\ \forall m < n\}$ el conjunto de las n-raíces primitivas de la unidad. Denotamos por $\Phi_n(x) = \prod_{\xi \in \Xi_n}(x - \xi) \in \mathbb{Z}[x]$ el n-ésimo polinomio ciclotómico. Un resultado conocido de la teoría de cuerpos (ver teorema 4.2.6 en [Wei09]) es que todos los polinomios ciclotómicos son irreducibles sobre los racionales.

Proposición 1.2.7. *Se cumple que:*

1. *$\forall n$ el polinomio $x^n - 1 = \prod_{d|n} \Phi_d(x)$.*
2. *Si $d \,|\, n$ y $d < n$, entonces $\Phi_n(x) \,|\, \frac{x^n-1}{x^d-1}$.*

Demostración. 1. Notemos que las raíces de $x^n - 1$ son precisamente las n-raíces de la unidad y, además, forman un grupo multiplicativo. Por el teorema 1.1.2, tenemos que los órdenes de dichas raíces serán divisores de n, luego agrupando por órdenes, tenemos que

$$x^n - 1 = \prod_{\xi \in \mathbb{C}, \xi^n = 1} (x - \xi) = \prod_{d \,|\, n} \prod_{\xi \in \Xi_d} (x - \xi) = \prod_{d \,|\, n} \Phi_d(x).$$

2. Notemos que $x^d - 1 = \prod_{t|d} \Phi_t(x)$. Además, tenemos que $t \mid n$ y como $n > d$, n no divide a d, luego $\Phi_n(x)$ no es un factor de $x^d - 1$. Por lo tanto $x^n - 1 = \Phi_n(x) \cdot (x^d - 1) \cdot \prod_{k \mid n, k \nmid d, k \neq n} \Phi_k(x)$. □

Teorema 1.2.8 (Pequeño teorema de Wedderburn)**.** *Todo anillo de división finito es un cuerpo.*

Demostración. Supongamos que D es finito y veamos que es conmutativo. Por 1.2.4 sabemos que $Z(D)$ es un subanillo de D que es un cuerpo, en particular, tenemos que $Z(D)$ también es finito, luego $|Z(D)| = q$, con q potencia de primo.

Si $n := dim_{Z(D)}(D) = 1$, tendremos que $D = Z(D)$, luego D es conmutativo y por tanto un cuerpo. Supongamos que $n > 1$ y procedamos por reducción al absurdo.

Consideremos el grupo multiplicativo $D^* = D \setminus \{0\}$. Notemos que $Z(D^*) = Z(D) \setminus \{0\}$. Por lo tanto, aplicando 1.1.8 obtenemos que:

$$|D^*| = |Z(D^*)| + \sum_{i=0}^{m} [D^* : C_{D^*}(x_i)] = |Z(D^*)| + \sum_{i=0}^{m} \frac{D^*}{C_{D^*}(x_i)}.$$

Notemos que $C_D(x_i) = C_{D^*}(x_i) \cup \{0\}$ es un $Z(D)$-espacio vectorial, luego $|C_D(x_i)| = q^{r_i}$. Además, $C_{D^*} \leq D^*$, por lo que por el teorema 1.1.2 tenemos que $|C_D(x_i)| = q^{r_i} - 1 \mid q^n - 1 = |D^*|$. Con lo que obtenemos:

$$q^n - 1 = q - 1 + \sum_{i=1}^{m} \frac{q^n - 1}{q^{r_i} - 1}.$$

Además se cumple que para todo $i = 0, \ldots, m$, $r_i \mid n$. En efecto, por el algoritmo de Euclides tenemos que $n = Q \cdot r_i + r$ con $0 \leq r < r_i$. Como $q^{r_i} - 1 \mid q^n - 1$, se tiene que cumplir que

$$q^{r_i} - 1 \mid \left(q^n - 1 - (q^{n-r_i} + q^{n-2r_i} + \cdots + q^{n-Qr_i}) \cdot (q^{r_i} - 1)\right) = q^{n-Qr_i} - 1 = q^r - 1.$$

Como $\forall i \ r_i \mid n$, por la proposición 1.2.7 tenemos que $\Phi_n(q) \mid \frac{q^n-1}{q^{r_i}-1}$ y como $\Phi_n(q) \mid q^n - 1$, tenemos entonces que $\Phi_n(q) = \prod_\xi (q - \xi) \mid q - 1$, en particular $q - 1 \geq \prod_\xi |q - \xi|$. Donde las ξ son las n-raíces primitivas de la unidad. Absurdo, pues al ser $n > 1$ y $q \geq 2$, tenemos que para cada ξ, $|q - \xi| > q - 1$. Luego $n = 1$ y, por tanto, D es un cuerpo. □

Corolario 1.2.9. *Todo subanillo finito de D es un cuerpo.*

Demostración. Sea R un subanillo finito de D. Notemos que al ser R subanillo, es cerrado para sumas y productos. Si demostramos que $\forall x \in R$ no nulo, $x^{-1} \in R$, tendremos que R es un anillo de división finito y por el teorema 1.2.8, un cuerpo. Sea $0 \neq x \in R$. Consideremos el conjunto $\{1, x, x^2, \ldots\} \subset R$. Como R es finito, $\exists m, n$, con $m < n$ tales que $x^m = x^n$. Sin embargo, como $x^{-1} \in D$, $x^{n-m} = 1$, luego $x^{n-m-1} = x^{-1}$. Si $x \neq 1$, $n - m - 1 > 0$, luego $x^{-1} = x^{n-m-1} \in R$. Por tanto R es cerrado para inversos y queda demostrado el corolario. □

Un resultado conocido en la teoría de cuerpos es que todo subgrupo del grupo multiplicativo de un cuerpo es cíclico. En la teoría de anillos de división ese resultado es, en general, falso. Consideremos por ejemplo el anillo de división, D, de los cuaternios. El grupo cuaternio $Q_8 = \{\pm 1, \pm i, \pm j, \pm k\} \leq D^*$ y no es cíclico (se puede comprobar fácilmente viendo que no hay elementos de orden 8). Sin embargo, si nos restringimos a los anillos de división de característica positiva, el resultado sí es cierto.

Corolario 1.2.10. *Sea D un anillo de división de característica positiva. Todo subgrupo finito de D^* es cíclico.*

Demostración. Sea $p > 0$ la característica de D, $G \leq D^*$ finito y $\mathbb{F}_p$ el cuerpo primo de D. Consideremos el conjunto:

$$K = \left\{ \sum fg \,|\, f \in \mathbb{F}_p \,,\, g \in G \right\}.$$

Tenemos que K es un subanillo finito de D, por lo tanto es un cuerpo. Además, como G es un subgrupo finito de K^*, G es cíclico □

Proposición 1.2.11. *Sea D un anillo de división no conmutativo. Para todo $d \in D$, su centralizador, $C_D(d)$, es infinito.*

Demostración. Por el teorema 1.2.8, sabemos que D es necesariamente infinito, luego tiene sentido preguntarnos si $C_D(d)$ es finito o infinito. Procedamos por reducción al absurdo. Supongamos que $C_D(d)$ es finito. Por el teorema 1.2.8, tenemos que es un cuerpo con $Z(D) \subset C_D(d)$. Entonces tenemos que $|Z(D)| = q$, $|C_D(D)| = q^n$, con q una potencia de primo y $n \geq 1$. Es inmediato ver que si $d \in Z(D)$, entonces $C_D(d) = D$ que es infinito por el teorema 1.2.8. Si $d \notin Z(D)$, necesariamente $n > 1$ y, debido a resultados conocidos de teoría de Galois [1] tenemos que la aplicación $\tilde{\mu} : C_D(d) \to C_D(D); \tilde{\mu}(a) = a^q$ es un automorfismo de $C_d(d)$ con cuerpo fijo $Z(D)$ que podemos extender a un automorfismo interno $\mu : D \to D$. Es decir, que existe $\sigma \in D$ tal que $\sigma a \sigma^{-1} = a^q \ \forall a \in C_D(d)$. En particular $\sigma d \sigma^{-1} = d^q$ y, componiendo n veces obtenemos que $\sigma^n d \sigma^{-n} = d^{q^n} = d$. Luego $\sigma^n \in C_D(d)$ y por tanto el orden de σ es finito. Consideremos el siguiente conjunto:

$$E := \left\{ \sum_{i,j} \lambda_{i,j} d^i \sigma^j \,|\, \lambda_{i,j} \in Z(D) \right\}.$$

Como tanto $Z(D)$, como $C_D(d)$ y $\{\sigma^i\}_{i \geq 0}$ son conjuntos finitos, E también lo será. Además E es cerrado para sumas por definición y comprobemos que también lo es para productos e inversos. En efecto, sean $x, y \in E$ con $x = \lambda d^i \sigma^j$ e $y = \mu d^k \sigma^l$. Tenemos que

$$xy = (\lambda d^i \sigma^j)(\gamma d^k \sigma^l) \underset{\gamma \in Z(D)}{=} \lambda\gamma d^i \sigma^j d^k \sigma^l \underset{\sigma d = d^q \sigma}{=} (\lambda\gamma) d^{i+kq^j} \sigma^{j+l} \in E.$$

[1] En este trabajo vamos a dar los resultados de teoría de Galois como conocidos. Para más información ver [Lan02] y [Wei09].

Si tomamos ahora dos elementos cualesquiera $a, b \in E$, su producto va a ser un sumatorio de sumandos con la misma forma que xy, que como hemos visto en la ecuación anterior, pertenecen a E. Al ser E cerrado por sumas, tenemos que $ab \in E$, luego es cerrado por productos. Esto nos lleva a ver que $E^* = E \setminus \{0\} \subset D^*$ es un semigrupo finito. Por la proposición 1.1.9, E^* es un grupo, luego E es un anillo de división finito.

Por el teorema 1.2.8 E es un cuerpo, sin embargo, $\sigma d = d^q \sigma \neq d\sigma$, pues $d^q \neq d$ al tener $\tilde{\mu}$ como mucho q puntos fijos (siendo estos puntos las raíces de $x^q - x \in C_D(d)[x]$) y como $d \notin Z(D)$, tenemos que $|Z(D) \cup \{d\}| = q + 1$. Luego llegamos a un absurdo y concluimos que en este caso $C_D(d)$ también es infinito. □

Para demostrar la siguiente propiedad necesitaremos usar el teorema de Jacobson, cuya demostración se puede ver en [Lam01, p.208].

Teorema 1.2.12 (Jacobson)**.** *Sea D un álgebra de división algebraica sobre un cuerpo finito $\mathbb{F}$. Entonces D es conmutativo.*

Proposición 1.2.13. *Sea D un anillo de división no conmutativo centralmente finito. Entonces $|Z(D)| = \infty$.*

Demostración. Ya sabemos que D es una $Z(D)$-álgebra, veamos que es algebraica. Como D es centralmente finito tenemos que $[D : Z(D)] = d < \infty$. Sea $a \in D$, consideremos el conjunto $\{a^i\}_{i=0}^d$. Como $\dim_{Z(D)} D = d$ y en dicho conjunto hay $d+1$ elementos, tenemos que es un conjunto linealmente dependiente. Por tanto existirán $\lambda_0, \ldots, \lambda_d \in Z(D)$ tales que

$$\sum_{i=0}^{d} \lambda_i a^i = 0;$$

sin embargo esta relación nos define un polinomio $p(x) = \sum_{i=0}^d \lambda_i x^i \in Z(D)[x]$ con $p(a) = 0$. Por tanto a es algebraico sobre $Z(D)$ y concluimos que D es un álgebra de división algebraica.

Por el teorema 1.2.12, si $Z(D)$ fuese un cuerpo finito, entonces D sería conmutativo, lo cual contradice nuestras hipótesis y por tanto $Z(D)$ es infinito. □

Capítulo 2

Polinomios sobre anillos de división.

La posible falta de conmutatividad en los anillos de división produce ciertas dificultades a la hora de definir polinomios. Es necesario tratar con cuidado las nociones de *raíz de un polinomio* y del *homomorfismo evaluación*. En este trabajo estudiaremos las dos principales formas de definir los polinomios, dependiendo de si permitimos que la variable x conmute con los coeficientes o no. Las principales referencias para este capítulo son [Lam01] y [GM65].

2.1. Polinomios a izquierda.

Los polinomios a izquierda se definen de forma idéntica a como se definen los polinomios sobre un cuerpo y para esta definición, estamos suponiendo que $ax = xa$, $\forall a \in D$.

Definición 2.1.1. Llamaremos *anillo de polinomios a izquierda* al conjunto

$$D[x] = \left\{ \sum_{i=0}^{n} a_i x^i \mid a_i \in D,\, n \in \mathbb{N} \right\}.$$

Es inmediato comprobar que $D[x]$ es un anillo con la suma y el producto de polinomios usuales. Sin embargo, la evaluación usual $ev_r : D[x] \to D;\ x \mapsto r$ no es un homomorfismo. Esto se puede comprobar fácilmente tomando $a, b \in D$ tales que $ab \neq ba$ y $f(x) = g(x)h(x) = (x-a)(x-b) = x^2 - (a+b)x + ab$. En estas condiciones tenemos que $f(a) = a^2 - (a+b)a + ab = ab - ba \neq 0$, mientras que $g(a)h(a) = (a-a)(b-a) = 0$.

Definición 2.1.2. Diremos que $r \in D$ es una *raíz a derecha* de $f(x) = a_n x^n + \cdots + a_1 x + a_0 \in D[x]$ si se cumple que

$$f(r) := a_n r^n + \cdots + a_1 r + a_0 = 0.$$

Llamaremos raíces a las raíces a derecha siempre y cuando esto no dé lugar a confusión.

Nota: Es importante remarcar que *antes de evaluar* es necesario operar y expresar el polinomio con *todos* sus coeficientes a la izquierda. En el ejemplo anterior, a no es una raíz de $f(x) = g(x)h(x)$, aunque $g(a) = 0$.

Pese a que, como se ha visto en el ejemplo anterior, en general no vamos a poder factorizar los polinomios en productos de factores de la forma $(x - \xi)$ donde $\xi \in D$ son raíces del polinomio en cuestión sí que podemos generalizar el teorema del factor que tenemos para los anillos de polinomios sobre un cuerpo.

Teorema 2.1.3 (Teorema del factor, [GM65, p.220])**.** *Sea D un anillo de división y $f \in D[x]$. Dado $\xi \in D$, tenemos que ξ es una raíz de f si y solo si existe un polinomio $g \in D[x]$ tal que $f(x) = g(x) \cdot (x - \xi)$.*

Demostración. Sea $f(x) = a_n x^n + \cdots + a_1 x + a_0$. Los casos $n = 0, 1$ son trivialmente ciertos, luego supondremos que $n \geq 2$. Supongamos que $\xi \in D$ es una raíz de $f(x)$ y consideremos el polinomio $g(x) = \sum_{i=0}^{n-1} b_i x^i$, donde $\forall i$, $b_i = \sum_{j=i}^{n-1} a_{j+1}\xi^{j-i}$. Es inmediato comprobar que $b_{n-1} = a_n$ y $\forall i \in \{1, \ldots, n-1\}$, $b_{i-1} - b_i\xi = a_i$, por tanto:

$$\begin{aligned} g(x)(x-\xi) =& (\sum_{i=0}^{n-1} b_i x^i)(x-\xi) = b_{n-1}x^n + \sum_{i=1}^{n-1} b_{i-1}x^i - \sum_{i=1}^{n-1} b_i\xi x^i - b_0\xi \\ =& a_n x^n + \sum_{i=1}^{n-1}(b_{i-1} - b_i\xi)x^i - \left(a_1\xi + a_2\xi^2 + \cdots + a_{n-1}\xi^{n-1}\right) \\ =& a_n x^n + \sum_{i=1}^{n-1} a_i x^i + a_0 = f(x), \end{aligned}$$

pues $f(\xi) = 0 \implies a_0 = -\left(a_1\xi + a_2\xi^2 + \cdots + a_{n-1}\xi^{n-1}\right)$. Si suponemos, en cambio, que $f(x) = g(x)(x - \xi)$, con $g(x) = \sum_{i=0}^{n-1} b_i x^i$, operando e igualando coeficientes del mismo grado obtenemos que $a_n = b_{n-1}$, $a_0 = -b_0\,\xi$ y para todo $i = 1, \ldots, n-1$, $a_i = b_{i-1} - b_i\,\xi$. Si, para cada $0 \leq i \leq n$, multiplicamos la ecuación asociada a a_i por ξ^i y sumamos todas obtenemos:

$$f(\xi) = \sum_{i=0}^{n} = a_i\xi^i = b_{n-1}\xi^n + \sum_{i=1}^{n-1}(b_{i-1} - b_i\xi)\xi^i - b_0\xi = \sum_{i=0}^{n-1}(b_i\xi^{i+1} - b_i\xi^{i+1}) = 0,$$

cumpliéndose así la condición necesaria. □

Nota: Este teorema es el motivo por el que a las raíces de los polinomios *a izquierda* se denominan raíces *a derecha*, pues los factores aparecen a la derecha.

Uno de los resultados fundamentales del álgebra conmutativa es el teorema fundamental del álgebra, que afirma que todo polinomio de grado n sobre un cuerpo tiene a lo sumo n raíces (exactamente n, contando multiplicidades, si el cuerpo es algebraicamente cerrado). Esto contrasta con dos de los resultados más sorprendentes de la teoría de polinomios sobre anillos de división, expuestos por Gordon y Motzkin en [GM65]. Sin embargo, para el primer teorema seguiremos la demostración que proporciona Lam en [Lam01], al considerarla más clara y simple. Para ello primero demostraremos un lema auxiliar que necesitaremos para la demostración.

Lema 2.1.4. *Sean D un anillo de división, $f(x) = g(x)h(x) \in D[x]$ y $d \in D$ tales que $\xi := h(d) \neq 0$. Entonces:*

$$f(d) = g(\xi\, d\, \xi^{-1})h(d).$$

En particular, si $d \in D$ es una raíz de f, pero no de h, entonces $\xi\, d\, \xi^{-1}$ es una raíz de g.

Demostración. Si $g(x) = \sum_{i=0}^{n} a_i x^i$, como x conmuta con todo elemento de D, aplicando la propiedad distributiva obtenemos que $f(x) = \sum_{i=0}^{n} a_i h(x) x^i$. Sustituyendo la x por d (aunque la evaluación no es un homomorfismo, como en la expresión de la derecha todas las x están a la derecha y los coeficientes a izquierda, es correcto hacerlo), como $\xi = h(d) \neq 0$, existe ξ^{-1} y obtenemos que:

$$f(d) = \sum_{i=0}^{n} a_i h(d) d^i = \sum_{i=0}^{n} a_i\, \xi\, d^i\, \xi^{-1}\, \xi = \sum_{i=0}^{n} a_i (\xi\, d\, \xi^{-1})^i h(d) = g(\xi\, d\, \xi^{-1})h(d).$$

Como todo elemento no nulo tiene inverso en un anillo de división, D no puede tener divisores de cero, luego si $f(d) = 0$, como $h(d) \neq 0$ necesariamente $g(\xi\, d\, \xi^{-1}) = 0$. □

Con ello ya podemos demostrar el siguiente resultado.

Teorema 2.1.5 (Gordon-Motzkin, [GM65, p.220])**.** *Sea D un anillo de división y $f(x) = \sum_{i=0}^{n} a_i x^i \in D$ (con $a_n \neq 0$). Entonces las raíces de f se encuentran en, a lo sumo, n clases de conjugación de D. Además, si $f(x) = \prod_{i=1}^{n}(x - \xi_i)$, entonces cada raíz de f es conjugada a ξ_i para algún $i = 1, \ldots, n$.*

Demostración. Procedamos por inducción. Los casos $n = 0, 1$ son triviales, luego suponga-mos que $n \geq 2$ y que el resultado es cierto para todo polinomio de grado $k < n$. Sea $\xi \in D$ una raíz de f, por el teorema 2.1.3 tenemos que $f(x) = g(x)(x - \xi)$. Sea $\eta \neq \xi$ otra raíz de f. Como $\eta \neq \xi$, tenemos que $\eta - \xi \neq 0$ y por tanto nos encontramos en las condiciones del lema 2.1.4, luego η es un conjugado de alguna raíz de g. Como $deg\, g(x) = n - 1$, aplicamos la hipótesis de inducción y obtenemos que η pertenece a una unión de a lo sumo $n - 1$ clases de conjugación, como ξ puede que no pertenezca a ninguna de estas clases de conjugación, obtenemos que las raíces de f se encuentran en, a lo sumo, n clases de conjugación. Por inducción, este resultado es válido $\forall n \geq 0$.

La última conclusión del teorema se sigue de forma similar aplicando el lema 2.1.4 de derecha a izquierda en $f(x) = (x - \xi_1) \cdots (x - \xi_n)$. □

Este teorema resulta impactante, porque a diferencia de lo que ocurre con los cuerpos, los polinomios sobre anillos de división pueden tener más raíces que su grado. El teorema 2.1.5 lo único que acota es el número de clases de conjugación en las que se pueden encontrar dichas raíces, no el número total de las mismas. De hecho, el segundo teorema de Gordon y Motzkin va a tratar el número de raíces que se encuentra en cada clase de conjugación y mostrará la facilidad con la que puede suceder que haya infinitas raíces.

Teorema 2.1.6 (Gordon-Motzkin, [GM65, p.221]**.** *Sea $f \in D[x]$. Sea $\Gamma \subset D$ el conjunto de raíces de f y sea A una clase de conjugación de D. Si $|\Gamma \cap A| \geq 2$, entonces $|\Gamma \cap A| = \infty$.*

Demostración. Sean $c \in \Gamma$ y $\sigma \in D$ tales que $c \neq \sigma c \sigma^{-1} \in \Gamma$.

Supongamos que $f(x) = \sum_{i=0}^{n} a_i x^i$ y consideremos la siguiente ecuación en y:

$$f(ycy^{-1}) = a_n y c^n y^{-1} + \cdots + a_1 ycy^{-1} + a_0 = (a_n y c^n + \cdots + a_1 yc + a_0 y)y^{-1} = 0. \quad (2.1.1)$$

Notemos que las raíces de $a_n y c^n + \cdots + a_1 yc + a_0 y = 0$ son las mismas que en 2.1.1 salvo $y = 0$ y de esta forma se ve fácilmente que, si $y_1, y_2 \in D$ son raíces de la ecuación y $z \in C_D(c)$, entonces tanto $y_1 + y_2$ como yz son raíces de la misma.

Notemos ahora que tanto $y = 1$ como $y = \sigma$ son raíces de 2.1.1, con lo que obtenemos que $\sigma + z$ es una raíz $\forall z \in C_D(c)$ y por tanto $(\sigma + z)c(\sigma + z)^{-1}$ será una raíz de f. Comprobemos ahora que $\forall z_1, z_2 \in C_D(c)$ tales que $z_1 \neq z_2$ se cumple que $(\sigma + z_1)c(\sigma + z_1)^{-1} \neq (\sigma + z_2)c(\sigma + z_2)^{-1}$.

Supongamos que $(\sigma + z_1)c(\sigma + z_1)^{-1} = (\sigma + z_2)c(\sigma + z_2)^{-1}$ y procedamos por reducción al absurdo. Tenemos entonces que $z_3 := (\sigma + z_2)^{-1}(\sigma + z_1) \in C_D(c)$. De aquí obtenemos que $\sigma + z_1 = \sigma z_3 + z_2 z_3$. Si $z_3 = 1$, inmediatamente obtenemos que $z_1 = z_2$ y llegamos a un absurdo. Por lo tanto, seguimos bajo la suposición de $z_3 \neq 1$. Notemos en este caso que podemos despejar y obtener $\sigma = (z_2 z_3 - z_1)(1 - z_3)^{-1}$. Pero esto implica que $\sigma \in C_D(c)$ al ser $C_D(c)$ un anillo de división (proposición 1.2.3), lo cual contradice la hipótesis $c \neq \sigma c \sigma^{-1}$. Por lo tanto podemos concluir que los elementos de la forma $(\sigma + z)c(\sigma + z)^{-1}$ con $z \in C_D(c)$ son todos raíces distintas de f. Como por la proposición 1.2.11 sabemos que $C_D(c)$ es infinito, tenemos que $|\Gamma \cap A| = \infty$, donde A es la clase de conjugación de c en D. □

Corolario 2.1.7. *Sea $f \in D[x]$ con $\deg(f) = n$. Entonces f tiene, o bien a lo sumo n raíces, o bien infinitas.*

Demostración. Supongamos que f tiene más de n raíces. Por el teorema 2.1.5 sabemos que esas raíces están repartidas en, a lo sumo n clases de conjugación, luego por el principio del palomar, habrá alguna clase con más de una raíz en ella. Por tanto, por el teorema 2.1.6 tenemos que hay infinitas raíces en esa clase de conjugación. □

2.2. Polinomios generales.

En esta sección definiremos el anillo de polinomios generales sobre un anillo de división y daremos los principales resultados sobre las raíces. La idea básica del anillo de polinomios

generales, al que denotaremos por $D\langle x\rangle$, es que la variable x en general no va a conmutar con los elementos de D (aunque sí con los de $Z(D)$). Por lo tanto, los monomios de $D\langle x\rangle$ en lugar de ser de la forma $a_i x^i$, serán de la forma $a_{i,0}xa_{i,1}x\dots a_{i,i-1}xa_{i,i}$. A continuación vamos a dar una definición formal de $D\langle x\rangle$ que respete las siguientes propiedades "evidentes"que esperamos que cumplan los monomios:

$$(a_0x\dots xa_i)(b_0x\dots xb_j) = a_0x\dots x(a_ib_0)x\dots xb_j. \tag{2.2.1}$$

$$a_0x\dots xa_ix\dots xa_n + a_0x\dots xb_ix\dots xa_n = a_0x\dots x(a_i+b_i)x\dots xa_n. \tag{2.2.2}$$

$$\forall c\in Z(D),\ cx = xc. \tag{2.2.3}$$

Para ello, consideremos el conjunto $S := \bigcup_{i=1}^{\infty} D^i$, de todas las secuencias finitas de elementos de D. Nuestro objetivo es identificar las secuencias $(a_0, a_1, \dots, a_n) \in S$ con los monomios $a_0xa_1\dots xa_n$. Empecemos definiendo un producto en S siguiendo las líneas de la ecuación 2.2.1, es decir, dados dos elementos $a, b \in S$, con $a = (a_i)_{i=0}^n$ y $b = (b_i)_{i=0}^k$:

$$ab = (a_0,\dots,a_n)(b_0,\dots,b_k) = (a_0,\dots,a_nb_0,\dots,b_k) \in S.$$

Como el producto de dos secuencias finitas sigue siendo una secuencia finita, es cerrado en S. Como además es inmediato ver que es asociativo tenemos que S es un semigrupo.

Consideremos el anillo de semigrupo de S, es decir, el conjunto de sumas formales de un número finito de elementos de S, $R := \{\sum_{i=0}^k \lambda\cdot s_i \,|\, k\in\mathbb{N},\ \lambda\in\mathbb{Z},\ \forall i\in\{0,\dots,k\}, s_i\in S\}$. Donde $\forall\lambda\in\mathbb{Z}$ y $\forall s\in S$, definimos:

$$\lambda\cdot s := \begin{cases} s + \overset{\lambda \text{ veces}}{\dots} + s, & \text{si } \lambda > 0;\\ (0), & \text{si } \lambda = 0;\\ (-1)(s + \overset{\lambda \text{ veces}}{\dots} + s), & \text{si } \lambda < 0, \end{cases}$$

y, dados $s\in S$ y $\lambda,\mu\in\mathbb{Z}$, tenemos que $\lambda\cdot s + \mu\cdot s = (\lambda+\mu)\cdot s$. Notemos que podemos ver S como un subconjunto de R identificando, para todo $s\in S$, $1\cdot s$ con s.

Definimos $*: R\times R\to R$ como la aplicación que cumple:

- $\forall a,b\in S\subset R$, $(1\cdot a)*(1\cdot b) = 1\cdot(ab)$.
- $\forall a,b,c\in R$, $a*(b+c) = a*b + a*c$ y $(a+b)*c = a*c + b*c$.

Como $*$ depende del producto definido en S, si no da lugar a confusión, omitiremos el símbolo. Notemos que $(R,+,*)$ es un anillo, sin embargo, no es el anillo que buscamos, pues de momento no terminamos de tener las propiedades equivalentes a 2.2.2 y 2.2.3.

Nota: Es necesario hacer un aparte porque, al ser R un anillo, usando la propiedad distributiva tenemos un resultado extremadamente parecido a 2.2.2 que puede llevar a confusión. En efecto, consideremos los elementos $(a_0,\dots,a_i,\dots,a_n)$,

$(a_0, \ldots, b_i, \ldots, a_n) \in R$ que solo se diferencian en la posición i. Por como hemos definido el producto en R, tenemos que podemos escribir ambos elementos como:

$$(a_0 \ldots, a_{i-1}, 1)(a_i)(1, a_{i+1}, \ldots a_n), \text{ y}$$

$$(a_0 \ldots, a_{i-1}, 1)(b_i)(1, a_{i+1}, \ldots a_n), \text{ respectivamente.}$$

Luego, si aplicamos la propiedad distributiva obtenemos que:

$$\begin{aligned}
&(a_0, \ldots, a_{i-1}, 1)(a_i)(1, a_{i+1}, \ldots a_n) \\
&+ (a_0, \ldots, a_{i-1}, 1)(b_i)(1, a_{i+1}, \ldots a_n) \\
&= (a_0, \ldots, a_{i-1}, 1)((a_i) + (b_i))(1, a_{i+1}, \ldots a_n) \\
&\overset{*}{\neq} (a_0, \ldots, a_{i-1}, 1)(a_i + b_i)(1, a_{i+1}, \ldots a_n) \\
&= (a_0, \ldots, a_i + b_i, \ldots, a_n),
\end{aligned}$$

pues la suma libre que hay en R no tiene (de momento) ningún tipo de relación con la suma definida en D, aunque por familiaridad estemos realizando un abuso de notación al representar ambas con el símbolo $+$; por lo tanto no podemos afirmar que $(a_i) + (b_i) = (a_i + b_i)$.

Para obtener eso vamos a forzar dichas relaciones de la siguiente forma. Consideremos $\forall k \in \mathbb{N}_0$ y $\forall i \in \mathbb{N}_0$ tal que $0 \leq i \leq k$ los conjuntos $A_{i,k}$ y B_k, donde:

$$\begin{aligned}
A_{i,k} := \{&(a_0, \ldots, a_i + b_i, \ldots, a_k) - (a_0, \ldots, a_i, \ldots, a_k) \\
&- (a_0, \ldots, b_i, \ldots, a_k) \,|\, a_0, \ldots, a_k, b_i \in D\}.
\end{aligned} \tag{2.2.4}$$

$$\begin{aligned}
B_k := \{&(a_0, \ldots, a_k) - (c_0 a_0, \ldots, c_k a_k) \,|\, a_0, \ldots, a_k \in D, \\
&c_0, \ldots, c_n \in Z(D),\ c_0 c_1 \cdots c_k = 1\}.
\end{aligned} \tag{2.2.5}$$

Y definimos el ideal $\mathfrak{a}$ como el ideal generado por el conjunto:

$$\bigcup_{k=0}^{\infty} \left(\bigcup_{i=0}^{k} A_{i,k} \cup B_k \right).$$

Consideremos el anillo cociente $R/\mathfrak{a}$ y definimos $D' := \{(a_0) + \mathfrak{a} \,|\, a_0 \in D\} \subset R/\mathfrak{a}$.

Proposición 2.2.1. *Se tiene que $D' \cong D$.*

Demostración. Consideremos la aplicación $\varphi : D' \to D$ tal que $\varphi((a_0) + \mathfrak{a}) = a_0$. Es inmediato ver que está bien definida. Sean $a_0, b_0 \in D$. Tenemos que $(a_0) + \mathfrak{a}, (b_0) + \mathfrak{a} \in D'$ y además se cumple que

$$\varphi((a_0) + \mathfrak{a}) + \varphi((b_0) + \mathfrak{a}) = a_0 + b_0 = \varphi((a_0 + b_0) + \mathfrak{a}) = \varphi((a_0) + (b_0) + \mathfrak{a}),$$

pues, en $R/\mathfrak{a}$, $(a_0+b_0)=(a_0)+(b_0)$, al estar $(a_0+b_0)-(a_0)-(b_0)\in A_{0,0}$. Por otro lado, también se cumple que:

$$\varphi((a_0b_0)+\mathfrak{a})=a_0b_0=\varphi((a_0)+\mathfrak{a})\varphi((b_0)+\mathfrak{a}).$$

Luego φ es un homomorfismo de anillos. Es inmediato comprobar que, además, es biyectivo, pues $\ker\varphi=\{(0)+\mathfrak{a}\}$ y $\forall a_0\in D$, $(a_0)+\mathfrak{a}\in\varphi^{-1}(a_0)$. Luego φ es un isomorfismo. □

Debido a esto identificaremos D con D' y escribiremos los elementos $(a_0)+\mathfrak{a}\in R/\mathfrak{a}$ simplemente como a_0. Por último, definimos $x:=(1,1)+\mathfrak{a}$. Luego podemos escribir los elementos $(a_0,\dots,a_n)+\mathfrak{a}\in R/\mathfrak{a}$ como:

$$(a_0,a_1,\dots,a_n)+\mathfrak{a}=(a_0)(1,1)(a_1)\cdots(1,1)(a_n)+\mathfrak{a}=a_0xa_1\cdots xa_n.$$

Proposición 2.2.2. *Los elementos $a_0x\cdots xa_n\in R/\mathfrak{a}$ cumplen las propiedades 2.2.1-2.2.3 (que reescribimos a continuación por conveniencia):*

1. $(a_0x\dots xa_i)(b_0x\dots xb_j)=a_0x\dots x(a_ib_0)x\dots xb_j$,
2. $a_0x\dots xa_ix\dots xa_n+a_0x\dots xb_ix\dots xa_n=a_0x\dots x(a_i+b_i)x\dots xa_n$,
3. $\forall c\in Z(D),\ cx=xc$.

Demostración. Sean $a_0x\dots xa_n, b_0x\dots xb_m\in R/\mathfrak{a}$ y $c\in Z(D)$. Tenemos que

$$\begin{aligned}(a_0x\dots xa_n)(b_0x\dots xb_m)&=(a_0,\dots,a_n)(b_0,\dots,b_m)+\mathfrak{a}\underset{\text{def. prod.}}{=}(a_0,\dots,a_nb_0,\dots,b_m)+\mathfrak{a}\\&=a_0x\cdots xa_nb_0x\cdots xb_m,\end{aligned}$$

luego se cumple 1. Por otro lado, tenemos que

$$\begin{aligned}a_0x\cdots xa_ix\cdots xa_n+a_0x\cdots xb_ix\cdots xa_n&=(a_0,\dots,a_i,\dots,a_n)+(a_0,\dots,b_i,\dots,a_n)+\mathfrak{a}\\&\underset{A_{i,n}\subset\mathfrak{a}}{=}(a_0,\dots,a_i+b_i,\dots,a_n)+\mathfrak{a}=a_0x\cdots x(a_i+b_i)x\cdots xa_n,\end{aligned}$$

por lo que se cumple 2. Por último, notemos que $c^{-1}\cdot c=1$, luego $(c,1)-(c^{-1}\cdot c,c\cdot 1)+\mathfrak{a}=(c,1)-(1,c)+\mathfrak{a}\in B_1$. Por lo tanto, tenemos que:

$$cx=(c,1)+\mathfrak{a}\underset{B_1}{=}(1,c)+\mathfrak{a}=xc.\ \square$$

Proposición 2.2.3. *$R/\mathfrak{a}$ es un $Z(D)$-espacio vectorial. Además el conjunto*

$$\mathcal{M}:=\{a_0x\cdots xa_k\,|\,k\geq 0,\ \forall i\in\{0,\dots,k\},\ a_i\in D\}\setminus\{0\}\subset R/\mathfrak{a},$$

es un conjunto generador de $R/\mathfrak{a}$.

Demostración. Debido a la proposición 2.2.1, podemos identificar los elementos de $Z(D)$ como elementos de $R/\mathfrak{a}$. De las propiedades de la suma y el producto en $R/\mathfrak{a}$ se ve inmediatamente que es un $Z(D)$-espacio vectorial.

Por otro lado, como todo elemento $v\in R/\mathfrak{a}$ es una suma finita de elementos de $\mathcal{M}$, claramente tenemos que $\operatorname{span}_{Z(D)}\mathcal{M}=R/\mathfrak{a}$. □

Proposición 2.2.4. *Sea $d \in D$. Definimos la aplicación $ev_d : R/\mathfrak{a} \to D$:*

$$\forall a = a_0 x \cdots x a_k \in \mathcal{M},\ ev_d(a) = a_0 d \cdots d a_k,\ \ y$$

$$\forall v = \sum_{i=0}^{k} v_i,\ v_i \in \mathcal{M},\ ev_d(v) = \sum_{i=0}^{k} ev_d(v_i).$$

Tenemos que ev_d es una aplicación lineal bien definida que, además, es un homomorfismo de anillos.

Demostración. Sabemos que la evaluación está bien definida para $e\tilde{v}_d : R \to D$, donde $e\tilde{v}_d(a_0, \ldots, a_k) = a_0 d \cdots d a_k$. Por lo tanto solo nos hace falta ver que $\mathfrak{a} \subset \ker e\tilde{v}_d$ para que al hacer el paso al cociente también esté bien definida.

Sean $i, k \in \mathbb{N}$ y $a, b, c \in R$ tales que $a = (a_0, \ldots, a_i, \ldots a_k)$, $b = (a_0, \ldots, b_i, \ldots, a_k)$ y $c = (a_0, \ldots, a_i + b_i, \ldots a_k)$. Tenemos que

$$e\tilde{v}_d(c-a-b) = a_0 d \cdots d(a_i+b_i) d \cdots d a_k - a_0 d \cdots d a_i d \cdots d a_k - a_0 d \cdots d b_i d \cdots d a_k \underset{\text{prop. dist.}}{=} 0,$$

luego, $\forall i, k \in \mathbb{N}$, $A_{i,k} \subset \ker e\tilde{v}_d$. Por otro lado, sean $k \in \mathbb{N}$, $a = (a_0, \ldots, a_k)$ y $b = (c_0 a_0, \ldots, c_k a_k)$, con $c_i \in Z(D)$, $\forall i \in \{0, \ldots, k\}$ y $\prod_{i=0}^{k} c_i = 1$. Tenemos

$$e\tilde{v}_d(a - b) = a_0 d \cdots d a_k - c_0 a_0 d \cdots d c_k a_k \underset{d \in Z(D)}{=} (1 - \prod_{i=0}^{k} c_i) ev_d(a) = 0,$$

luego, $\forall k \in \mathbb{N}$, $B_k \subset \ker e\tilde{v}_d$. Como $\mathfrak{a}$ está generado por el conjunto

$$\bigcup_{k=0}^{\infty} \left(\bigcup_{i=0}^{k} A_{i,k} \cup B_k \right),$$

y $\forall i, k \in \mathbb{N}$ $A_{i,k} \cup B_k \subset \ker e\tilde{v}_d$, tendremos que $\mathfrak{a} \subset \ker e\tilde{v}_d$, por tanto ev_d está bien definida. Además, ev_d respeta las sumas de monomios por construcción.

Comprobemos que ev_d respeta el producto de monomios. Sean $a_0 x \cdots x a_n,\ b_0 x \cdots x b_m \in R/\mathfrak{a}$. Tenemos que:

$$\begin{aligned} ev_d(a_0 x \cdots x a_n b_0 x \cdots x b_m) &= a_0 d \cdots d a_n b_0 d \cdots d b_m \\ &= (a_0 d \cdots d a_n)(b_0 d \cdots d b_m) = ev_d(a_0 x \cdots x a_n) ev_d(b_0 x \cdots x b_m). \end{aligned}$$

Notemos además que $\forall \lambda \in Z(D) \subset D \subset R/\mathfrak{a}$, $ev_d(\lambda) = \lambda$. Juntando estas dos últimas observaciones, tenemos que $\forall \lambda \in Z(D)$ y $\forall a \in R/\setminus \mathfrak{a}$, $ev_d(\lambda a) = \lambda ev_d(a)$. Concluyendo así que ev_d es una aplicación lineal que respeta el producto, por tanto será un homomorfismo de anillos. □

El uso de esta notación, las propiedades que cumple y el hecho de que podamos definir una evaluación en D nos justifica en la decisión de referirnos a $R/\mathfrak{a}$ como el anillo de polinomios generales.

Definición 2.2.5. Llamaremos a $R/\mathfrak{a}$ *anillo de polinomios generales* y lo denotaremos por $D\langle x\rangle$.

A los elementos $a_0x\cdots xa_n \in \mathcal{M}$, como ya hemos visto, los llamaremos *monomios generales* y los denotaremos por $M_\nu(x)$ (siendo ν un índice genérico). Si $M(x) = a_0x\cdots xa_n$ cumple $\prod_{i=0}^n a_i \neq 0$, diremos entonces que $M(x)$ *es de grado* n y lo denotaremos por $\deg M(x) = n$.

Por último, tenemos que todo elemento de $D\langle x\rangle$ se puede expresar como una suma finita de monomios generales, $f(x) = \sum_{\nu=1}^m M_\nu(x)$. A dichos elementos los llamaremos polinomios generales.

Lema 2.2.6. *Sea $f \in D\langle x\rangle$. La aplicación*

$$\begin{aligned} ev_f : D\langle x\rangle &\to D\langle x\rangle \\ g &\mapsto g(f(x)), \end{aligned}$$

es un endomorfismo de anillos.

Demostración. Sean $f, g, h \in D\langle x\rangle$ con $g = g_0x\cdots xg_n, h = h_0x\cdots xh_m$ monomios. Es inmediato ver que

$$\begin{aligned} ev_f(g+h) &= (g+h)(f(x)) = (g_0x\cdots xg_n + h_0x\cdots xh_m)(f(x)) \\ &= g_0f(x)\cdots f(x)g_n + h_0f(x)\cdots f(x)h_m = ev_f(g) + ev_f(h), \end{aligned}$$

además, dado $\lambda \in Z(D)$, tenemos que $ev_f(\lambda g) = \lambda g_0f(x)\cdots f(x)g_n = \lambda ev_f(g)$. Por tanto, como respeta las sumas de monomios y el producto por escalares, por la proposición 2.2.3 tenemos que ev_f es una aplicación lineal entre espacios vectoriales. Si comprobamos que dados dos monomios cualesquiera $g = g_0x\cdots xg_n, h = h_0x\cdots xh_m \in D\langle x\rangle$, $ev_f(gh) = ev_f(g)ev_f(h)$, tendremos que ev_f será un endomorfismo de anillos. En efecto,

$$\begin{aligned} ev_f(gh) &= (g_0x\cdots xg_nh_0x\cdots xh_m)(f(x)) = (g_0f(x)\cdots f(x)g_n)(h_0f(x)\cdots f(x)h_m) \\ &= ev_f(g)ev_f(h), \end{aligned}$$

por lo que, como los monomios generan $D\langle x\rangle$ y ev_f respeta tanto la suma como el producto de monomios, tendremos que $\forall g, h \in D\langle x\rangle$, $ev_f(g+h) = ev_f(g) + ev_f(h)$ y $ev_f(gh) = ev_f(g)ev_f(h)$. □

Definición 2.2.7. Sean $f(x) = \sum_{M\in\mathcal{M}} M(x) \in D\langle x\rangle$, donde $\mathcal{M}$ es un conjunto de monomios, e $i \in \mathbb{N}$. Llamaremos *componente homogénea de grado* i al polinomio:

$$M_i(x) := \sum_{\substack{M\in\mathcal{M} \\ \deg M = i}} M(x).$$

Nota: Hemos definido las componentes homogéneas $\forall i \in \mathbb{N}$. Notemos sin embargo que, como todo polinomio $f = \sum_{i=0}^\infty M_i(x) \in D\langle x\rangle$ está definido como una suma finita de monomios, $\exists n \in \mathbb{N}$ tal que $\forall i > n$, $M_i(x) = 0$.

Lema 2.2.8. *Todo polinomio en $D\langle x\rangle$ se puede expresar de forma única como suma de componentes homogéneas.*

Demostración. Si D es un cuerpo el resultado es trivialmente cierto, luego supongamos que es un anillo de división no conmutativo. Sean $f \in D\langle x\rangle$ y dos familias de monomios, $\mathcal{M}$ y $\mathcal{N}$, tales que

$$f(x) = \sum_{M\in\mathcal{M}} M(x) = \sum_{N\in\mathcal{N}} N(x),$$

Sean $\{M_i(x)\}_{i\in\mathbb{N}}$ y $\{N_i(x)\}_{i\in\mathbb{N}}$ las familias de componentes homogéneas asociadas a cada sumatorio. Tenemos que $f(x) = \sum_{i=0}^{\infty} M_i(x) = \sum_{j=0}^{\infty} N_j(x)$. Veamos que $\forall i \in \mathbb{N}$, $M_i(x) = N_i(x)$. En efecto, notemos que

$$p(x) = \sum_{i=0}^{\infty} P_i := \sum_{i=0}^{\infty} M_i(x) - \sum_{j=0}^{\infty} N_j(x) = 0.$$

Además, dado $\lambda \in Z(D)$, para todo monomio de grado k $A(x) = a_0x\cdots xa_k$ se cumple $ev_{\lambda x} = A(\lambda x) = a_0\lambda x\cdots\lambda x a_k = \lambda^k A(x)$. Sea $\mathfrak{m} \in \mathbb{N}$ tal que $M_i(x) = 0 = N_i(x), \forall i > \mathfrak{m}$. Entonces podemos escribir $p(x) = \sum_{i=0}^{\mathfrak{m}} P_i(x)$. Por lo tanto, sea $\{\lambda_j\}_{j=0}^{\mathfrak{m}} \subset Z(D)$, distintos dos a dos (podemos tomarlos porque sabemos que $Z(D)$ es infinito mediante la proposición 1.2.13). Tenemos que el sistema de ecuaciones

$$p(\lambda_j x) = \sum_{i=0}^{\mathfrak{m}} P_i(\lambda_j x) = \sum_{i=0}^{\mathfrak{m}} \lambda_j^i P_i(x) = 0, \quad \forall j \in \{0,\dots,\mathfrak{m}\},$$

es un sistema homogéneo compatible determinado, pues la matriz $\Lambda = (\lambda_j^i)_{j,i=0}^{\mathfrak{m}}$ es una matriz de Vandermonde y, como $\lambda_i \neq \lambda_j$, $\forall i \neq j$ tenemos que $\det\Lambda \neq 0$, ([Lan02, p.516]). Con lo cual tenemos que $\forall i \in \{0,\dots,\mathfrak{m}\}, P_i(x) = 0$ y por tanto $\forall i \in \mathbb{N}$, $M_i(x) = N_i(x)$. □

Definición 2.2.9. Sea $f(x) = \sum_{i\in\mathbb{N}} M_i(x) \in D\langle x\rangle$, donde $M_i(x)$ es la componente homogénea de grado i de f. Definimos el *grado del polinomio* f como el grado de la mayor componente homogénea no nula, es decir,

$$\deg f = \max_{i\in\mathbb{N}}\{i \mid M_i(x) \neq 0\}.$$

El máximo existirá porque solo habrá un número finito de componentes homogéneas no nulas, por tanto $\deg f$ está bien definido.

Definición 2.2.10. Como en la construcción anterior solo hemos usado que D es un anillo unitario, definimos el *anillo de polinomios generales en n variables*

$$D\langle x_1,\dots,x_n\rangle := D\langle x_1,\dots,x_{n-1}\rangle\langle x_n\rangle,$$

con el grado, evaluación, etc. definidos de forma análoga al caso de una sola variable.

Definición 2.2.11. Sean $d \in D$ y $f \in D\langle x\rangle$. Diremos que d *es una raíz de* f si $f(d) := ev_d(f) = 0$.

Al conjunto de raíces de f en D lo denotaremos $N(f) = \{d \in D \mid f(d) = 0\}$.

A continuación vamos a estudiar qué propiedades tienen las raíces de los polinomios generales en el caso de que D sea centralmente finito, es decir, cuando $[D : Z(D)] = d < \infty$, lo cual nos permite simplificar la cuestión explotando la estructura de espacio vectorial de D mediante el álgebra lineal.

Definición 2.2.12. Sea $\mathcal{B} = \{1 = e_1, e_2, \ldots, e_d\}$ una $Z(D)$-base de D. Entendemos la x como un elemento genérico de D y trabajamos con sus coordenadas genéricas: $x = \xi_1 e_1 + \cdots + \xi_d e_d$. Sea $f \in D\langle x\rangle$, descomponiendo también cada coeficiente en sus coordenadas, evaluando en $x \in D$ y operando y agrupando convenientemente, podemos escribir $f(x) = f_1(\xi_1, \ldots, \xi_d)e_1 + \cdots + f_d(\xi_1, \ldots, \xi_d)e_d$, donde $f_i \in Z(D)[\xi_1, \ldots, \xi_d]$. Definimos

$$\Phi : D\langle x\rangle \to A := \bigoplus_{i=1}^{d} Z(D)[\xi_1, \ldots, \xi_d]e_i$$

$$f \mapsto \sum_{i=1}^{d} f_i(\xi_1, \ldots, \xi_d)e_i,$$

donde los f_i son los definidos anteriormente.

Nota: Es importante resaltar que si $\deg f = n$, para todo i $\deg f_i \leq n$. En efecto, como $\deg f = n$, cada monomio va a tener, como mucho, n apariciones de x. Por lo tanto, al sustituir por las coordenadas y aplicar la propiedad distributiva, en cada sumando parecerá un producto de a lo sumo n variables χ_i, con $i \in \{1, \ldots, d\}$.

Antes de dar el primer resultado ([GM65, Teorema 6]) es importante explicar que, a nuestro juicio, Gordon y Motzkin no fueron lo suficientemente rigurosos pues tratan en todo momento como si A fuese un subanillo de $D\langle x\rangle$ sin demostrarlo. De hecho, este primer resultado, lo que en esencia demuestra, es que existe una aplicación, $\Phi : D\langle x\rangle \to A := \bigoplus_{i=1}^{d} Z(D)[\xi_1, \ldots, \xi_d]e_i$, que es sobreyectiva y respeta las evaluaciones. Por ello, vamos a actualizar el enunciado y la demostración intentando ser más rigurosos. Una vez demostrado, podremos ver las ecuaciones $f(x) = 0$, como un sistema de ecuaciones en varias variables $f_i(\xi_1, \ldots, \xi_d) = 0$, $\forall i \in \{1, \ldots, d\}$. Aunque a priori pueda parecer que hemos complicado el problema, al tener más ecuaciones y más variables, notemos que los f_i son polinomios conmutativos en varias variables *pues $Z(D)$ es un cuerpo*, permitiéndonos así usar resultados conocidos de álgebra conmutativa.

Nota: De ahora en adelante, a menos que se diga lo contrario, vamos a usar $\mathcal{B}$ siempre como base.

Teorema 2.2.13 ([GM65, p.222])**.** *La aplicación Φ es sobreyectiva y respeta las evaluaciones, entendiendo esto último como*

$$f(a) = \Phi(f)(a) = \sum_{i=1}^{d} f_i(a_1, \ldots, a_d)e_i, \forall a \in D,$$

donde $a = a_1e_1 + \cdots + a_de_d$.

Para dar la demostración necesitaremos dos resultados de Jacobson ([Jac64, p.32]) que no demostraremos en este trabajo por salirse del ámbito del mismo.

Definición 2.2.14. Sea V un D-espacio vectorial por la izquierda, siendo D un anillo de división. Diremos que un conjunto X de endomorfismos lineales de V es *k-veces transitivo* si y solo si, para cualquier conjunto de vectores linealmente independientes $\{v_1, \ldots, v_i\}$ con $i \leq k$ se cumple que para i vectores cualesquiera $w_1, \ldots, w_i$ existe un endomorfismo $f \in X$ tal que $f(v_j) = w_j$, para todo $j \in \{1, \ldots, i\}$.

Teorema 2.2.15 (Teorema de densidad de Jacobson)**.** *Sea D un anillo de división, V un D-espacio vectorial y $\mathcal{L}(V)$ el conjunto de endomorfismos lineales de V. Sea $R \subset \mathcal{L}$ tal que R es un anillo y es 2-transitivo. Entonces R es denso en $\mathcal{L}(V)$ con la topología finita de $\mathcal{L}(V)$.*

Corolario 2.2.16. *Sea V un D-espacio vectorial con $\dim_D V = n < \infty$. Sea R un anillo de endomorfismos lineales de V denso en $\mathcal{L}(V)$. Entonces $R = \mathcal{L}(V)$.*

Demostración del teorema 2.2.13. Como cada polinomio viene determinado de forma única por sus componentes homogéneas es inmediato ver que Φ está bien definida. Veamos ahora que es sobreyectiva.

Sea $v := \sum_{i=1}^{d} f_i(\xi_1, \ldots, \xi_d)e_i \in \bigoplus_{i=1}^{d} Z(D)[\xi_1, \ldots, \xi_d]e_i$. Veamos que $\exists f \in D\langle x\rangle$ tal que $\Phi(f) = v$. Si $d = 1$, tenemos que $D \cong Z(D)$ y $D\langle x\rangle \cong Z(D)[x] \cong Z(D)[\xi_1] = A$, con $\Phi(x) = \xi_1$, por lo que la conclusión es trivialmente cierta. Asumamos ahora que $d > 1$. Para demostrarlo, basta con probar que si todos los f_i son homogéneos de grado n, podemos encontrar un polinomio $f \in D\langle x\rangle$ homogéneo de grado n tal que $\Phi(f) = \sum_i f_ie_i$, pues para el caso general basta con separar los f_i en sus componentes homogéneas, aplicar el caso homogéneo y sumar los polinomios resultantes.

Si $n = 1$, tenemos que los f_i son de la forma $f_i(\xi_1, \ldots, \xi_d) = \sum_{j=1}^{d} \alpha_{ij}\xi_j$ con $\alpha_{ij} \in Z(D)$.

De esta forma, $\tilde{f}(\lambda_1, \ldots, \lambda_d) = \sum_{i=0}^{d} ev_{(\lambda_1,\ldots,\lambda_d)}f_i(\xi_1, \ldots, \xi_d) \cdot e_i$ para $(\lambda_1, \ldots, \lambda_d) \in Z(D)^d$ define un homomorfismo lineal de $Z(D)^d$ en D. Además, como D es un $Z(D)$-espacio vectorial de dimensión d tenemos que, una vez fijada la base $\mathcal{B}$, existe un isomorfismo lineal $\varphi : D \to Z(D)^d$ definido por $\varphi(e_i) := (\delta_{ij})_{j=0}^{d}$, donde δ_{ij} es la delta de

Kronecker. Por tanto $\mathfrak{f} := \tilde{f} \circ \varphi$ es un endomorfismo lineal de D. Ahora lo que queremos ver es que existe un polinomio homogéneo $f(x) \in D\langle x \rangle$ de grado 1 tal que $\forall d \in D,\ f(d) = \mathfrak{f}(d)$.

Sea $R \subset D\langle x \rangle$ el conjunto de polinomios homogéneos de grado 1. Estos polinomios son de la forma $f(x) = \sum_\nu a_\nu x b_\nu$, con $a_\nu, b_\nu \in D$. Como tanto la suma, como la composición de endomorfismos lineales es un endomorfismo lineal, $(R, +, \circ)$ es un anillo. Veamos ahora que R es 2-transitivo.

Sean $a, b \in D$ linealmente independientes y $c, d \in D$ dos elementos cualesquiera. Como $a, b \in D$, tenemos que $ab \neq 0$. Además $ab^{-1} \notin Z(D)$. Pues, en caso contrario tendríamos que $(-1) \cdot a + (ab^{-1}) \cdot b = -a + a = 0$, lo cual contradice la hipótesis de que a y b son linealmente independientes. Por tanto, existe al menos un elemento $r \in D$ tal que $s := rab^{-1} - ab^{-1}r \neq 0$ y $t := ba^{-1}r^{-1} - r^{-1}ba^{-1} \neq 0$. Con ello podemos definir el siguiente polinomio:

$$g(x) = (rxb^{-1} - xb^{-1}r)s^{-1}c + (xa^{-1}r^{-1} - r^{-1}xa^{-1})t^{-1}d,$$

que cumple que g define un endomorfismo lineal con $g(a) = c$ y $g(b) = d$. Por tanto, R es 2-transitivo y por 2.2.15 R es denso en $\mathcal{L}(D)$. Además, como $[D : Z(D)] = d < \infty$, por el corolario 2.2.16 tenemos que $R = \mathcal{L}(D)$, pero recordemos que $\mathfrak{f} \in \mathcal{L}(D)$, luego existirá un polinomio homogéneo de grado 1, $f(x) \in R$, tal que, $\Phi(f) = \sum_{i=1}^{d} f_i(\xi_1, \ldots, \xi_d)e_i$, con $f(d) = \mathfrak{f}(d)$ para todo $d \in D$.

Para estudiar el caso de $n > 1$, que ya no es lineal, vamos a trabajar con polinomios de n variables, para conseguir linealidad en cada una de ellas. Sea $D\langle x_1, \ldots, x_n \rangle$ el anillo de polinomios en n variables. Diremos que un polinomio $p(x_1, \ldots, x_n) \in D\langle x_1, \ldots, x_n \rangle$ es multilineal si p es homogéneo y, para todo $1 \leq i \leq n$, p es lineal con respecto a x_i dejando el resto de variables fijas.

Siguiendo la misma estrategia que anteriormente, expresamos las variables $\{x_i\}_{i=1}^n$ como n elementos generales de D con coordenadas genéricas en la $Z(D)$-base $\mathcal{B}$, obteniendo que $\forall i\ x_i = \sum_{j=1}^{d} \xi_{ij} e_j$. De esta forma, si expresamos los coeficientes de p en coordenadas, operamos y agrupamos convenientemente obtenemos que:

$$p(x_1, \ldots, x_n) = \sum_{i=1}^{d} p_i(\xi_{11}, \ldots, \xi_{1d}, \ldots, \xi_{nd})e_i.$$

Es fácil comprobar que si $p(x_1, \ldots, x_n)$ es multilineal entonces, para todo i, los polinomios $p_i(\xi_{11}, \ldots, \xi_{nd})$ son lineales en cada conjunto de variables $\{\xi_{j1}, \ldots, \xi_{jd}\}$. En efecto, si p es multilineal, será de la forma $p(x_1, \ldots, x_n) = \sum_\nu a_{\nu 0} x_{\sigma(1)} \cdots x_{\sigma(k)} a_{\nu k}$, con $1 \leq k \leq n$ y σ una permutación de $\{1, \ldots, n\}$. Tomemos uno de esos sumandos cualquiera y expresémoslo en términos de las coordenadas con respecto a la base $\{e_i\}_{i=1}^d$.

$$\begin{aligned} a_{\nu 0} x_{\sigma(1)} \cdots x_{\sigma(k)} a_{\nu k} =& (\alpha_{\nu 0}^1 e_1 + \cdots + \alpha_{\nu 0}^d e_d)(\xi_{\sigma(1)1} e_1 + \cdots + \xi_{\sigma(1)d} e_d) \cdot \cdots \\ & \cdot (\xi_{\sigma(k)1} e_1 + \cdots + \xi_{\sigma(k)d} e_d)(\alpha_{\nu k}^1 e_1 + \cdots + \alpha_{\nu k}^d e_d), \end{aligned}$$

lo que, como antes, nos define una aplicación

$$\Phi_n : D\langle x_1 \ldots, x_n \rangle \to A_n := \bigoplus_{i=1}^{d} Z(D)[\xi_{11}, \ldots, \xi_{nd}].$$

Notemos que, al aplicar la propiedad distributiva, las variables $\{\xi_{j1}, \ldots, \xi_{jd}\}$ no interactúan entre sí, cayendo en sumandos distintos, lo cual inmediatamente nos lleva a que, una vez agrupemos los sumandos en los p_i, cada p_i va a ser lineal en cada conjunto de variables $\{\xi_{j1}, \ldots, \xi_{jd}\}$.

Nota: De ahora en adelante, cometeremos un pequeño abuso de notación y como cada variable x_j tiene asociadas las d variables $\{\xi_{jk}\}_{k=1}^{d}$, en vez de decir *el polinomio $p_i(\xi_{11}, \ldots, \xi_{nd})$ es lineal en cada conjunto de variables $\{\xi_{j1}, \ldots, \xi_{jd}\}$*, diremos simplemente que el polinomio p_i es multilineal.

A continuación comprobaremos que dados d polinomios multilineales $p_1, \ldots, p_d \in Z(D)[\xi_{11}, \ldots, \xi_{nd}]$, siempre podremos encontrar un polinomio multilineal $p(x_1, \ldots, x_n) \in D\langle x_1, \ldots, x_n\rangle$ tal que $\Phi_n(p) = \sum p_i e_i$. Para ello sean $g_k(\xi_{k1}, \ldots, \xi_{kd}) \in Z(D)[\xi_{k1}, \ldots, \xi_{kd}]$ con $k = 1, \ldots, n$ polinomios homogéneos de grado 1. Aplicando lo demostrado en el caso $n = 1$, y abusando de la notación con respecto a los nombres de las variables, a los polinomios $g_1 e_i, g_2 e_1, \ldots, g_n e_1$ (con $1 \leq i \leq d$ fijo), existirán polinomios homogéneos de grado 1, $h_k(x_k) \in D\langle x_k\rangle$, tales que $\Phi(h_1(x_1)) = g_1 e_i$ y, si $1 < k \leq n$, $\Phi(h_k(x_k)) = g_k e_1 = g_k$.

Notemos ahora dos cosas. La primera es que $\prod_{j=1}^{n} \Phi(h_j) = g_1 \cdots g_n e_i$. La segunda es que, al ser p_i multilineal, sus sumandos van a ser de la forma $g_1 \cdots g_n$. Juntando estas dos observaciones obtenemos que $p_i e_i$ se podrá expresar como una suma de elementos de la forma $\Phi(h_1(x_1)) \cdots \Phi(h_n(x_n))$. Aplicando este proceso para cada $1 \leq i \leq d$ se sigue inmediatamente que podemos encontrar un polinomio $p \in D\langle x_1, \ldots, x_n\rangle$ tal que $\Phi_n(p(x_1, \ldots, x_n)) = \sum_{i=1}^{d} p_i e_i$. Además, como los polinomios $h_k(x_k)$ que hemos construido son homogéneos de grado 1, p será multilineal.

Para finalizar la demostración lo único que nos queda por hacer es notar que, dados $f_i(\xi_1, \ldots, \xi_d) \in Z(D)[\xi_1, \ldots, \xi_d]$ homogéneos de grado menor o igual a n, mediante polarización podemos construir polinomios $p_i \in Z(D)[\xi_{11}, \xi_{12}, \ldots, \xi_{nd}]$ *multilineales* tales que, $\forall i \in \{1, \ldots, d\}$ $f_i = p_i|_{\substack{\xi_{j1} = \cdots = \xi_{jn} = \xi_j \\ \forall j \in \{1, \ldots, n\}}}$. Por lo tanto, con lo visto anteriormente, obtenemos un polinomio multilineal en n variables, $p \in D\langle x_1, \ldots, x_n\rangle$ tal que $\Phi_n(p) = \sum_{i=1}^{d} p_i e_i$. Por último, definimos $f(x) := p(x, \ldots, x) \in D\langle x\rangle$ y, por construcción, tenemos que f es homogéneo de grado $\leq n$ y $\Phi(f) = \sum_{i=1}^{d} f_i e_i$.

Por tanto hemos visto que, dado cualquier elemento $\tilde{f} \in A$, existe un elemento $f \in D\langle x\rangle$ tal que $\tilde{f} = \Phi(f)$. Además, en la construcción se ha visto que respeta las evaluaciones. □

Estos resultados técnicos nos van a permitir demostrar dos teoremas que, en esencia, nos van a decir que dado un $n \in \mathbb{N}$ vamos a poder encontrar polinomios $f \in D\langle x\rangle$ con $\deg f = n$ con cualquier número de raíces entre 0 y n^d, donde $d := [D : Z(D)]$.

Teorema 2.2.17 (Gordon-Motzkin, [GM65, p.223])**.** *Sea K un cuerpo infinito y sean $\{n_1, \ldots, n_d\} \subset \mathbb{N}$. Sea $h \in \mathbb{N}$ tal que $1 \leq h \leq n_1 n_2 \cdots n_d$. Existen d polinomios $f_1, \ldots, f_d \in$*

$K[\xi_1, \ldots, \xi_d]$ con $\deg f_i = n_i$ tales que el sistema $\forall i \in \{1, \ldots, d\},\ f_i(\xi_1, \ldots, \xi_d) = 0$ tiene exactamente h soluciones.

Además, si $d > 1$, también se cumple para $h = 0$.

Demostración. Vamos a demostrarlo mediante inducción sobre d y, teniendo la d fija, sobre $s = \sum_{i=1}^{d} n_i$. Notemos que como $\forall i,\ n_i > 0$, en particular $s \geq d$. Además, vamos a probar un resultado más fuerte del que nos exige el teorema, al resultar más sencillo de demostrar, que es que podemos tomar los f_i como producto de polinomios lineales que cumplen que, si tomamos p_i un factor lineal de f_i para cada $i \in \{1, \ldots, d\}$ entonces el conjunto $\{p_i\}_{i=1}^{d}$ es K-linealmente independiente.

Empecemos estudiando el caso $d = 1$. Tomemos h elementos distintos $\alpha_1, \ldots, \alpha_h \in K$. Tenemos que el polinomio $f_1(\xi_1) = (\xi_1 - \alpha_1)^{n_1 - h + 1}(\xi_1 - \alpha_2) \cdots (\xi_1 - \alpha_h)$ es de grado n_1 y tiene exactamente h raíces, luego la ecuación $f_1 = 0$ tiene h soluciones.

Si $d > 1$, empecemos probando el caso $s = d \iff n_1 = n_2 = \cdots = n_d = 1$. En este caso la condición del teorema es $0 \leq h \leq 1$. Si $h = 1$, basta con tomar los polinomios $f_i = \xi_i$. Estos polinomios son de la forma que buscamos y la única solución al sistema $f_i = 0, \forall i$ es $\xi_1 = \cdots = \xi_d = 0$. Por otro lado, si $h = 0$, nos basta con tomar $f_1 = \xi_1$, $f_2 = \xi_1 + 1$ y $\forall i > 2,\ f_i = \xi_i$. Notemos que los polinomios f_i siguen siendo linealmente independientes y, además, $f_1 = 0 \iff \xi_1 = 0$, pero si $\xi_1 = 0$, $f_2 \neq 0$, luego el sistema no tiene solución.

Supongamos ahora que el teorema se cumple para todos los conjuntos $\{m_1, \ldots, m_c\} \subset \mathbb{N}$ si $c < d$ y, si $c = d$, para todos los conjuntos $\{m_1, \ldots, m_d\} \subset \mathbb{N}$ si $\sum_{i=1}^{d} m_i < s$. Podemos suponer, sin pérdida de generalidad, que $n_1 > 1$, pues podríamos reordenar las variables.

Por hipótesis de inducción, el teorema es cierto para el conjunto $\{n_1 - 1, n_2, \ldots, n_d\}$, luego, dado $0 \leq h \leq (n_1 - 1)n_2 \cdots n_d$ podemos encontrar polinomios $g_i \in K[\xi_1, \ldots, \xi_d]$ que cumplen el teorema. En particular, como $\deg g_1 = n_1 - 1$, tomemos $p \in K[\xi_1, \ldots, \xi_d]$ un factor de g_1, tenemos que el sistema formado por los polinomios $f_1 = pg_1$ y, para todo $i \geq 2$, $f_i = g_i$ también tiene h soluciones, está formado por polinomios que son producto de factores lineales pero además cumple que $\forall i \geq 1$, $\deg f_i = n_i$.

Luego nos queda estudiar el caso en el que $(n_1 - 1)n_2 \cdots n_d < h \leq n_1 n_2 \cdots n_d$. Para ello notemos que podemos expresar $h = (n_1 - 1)n_2 \cdots n_d + k$, con $1 \leq k \leq n_2 \cdots n_d$. Sin embargo, el conjunto $\{n_2, \ldots, n_d\}$ cumple las hipótesis de inducción, luego podemos encontrar $d - 1$ polinomios $g_2, \ldots, g_d \in K[\xi_2, \ldots, \xi_d]$ tales que $\forall i \geq 2$, $\deg g_i = n_i$, todos ellos están formados por factores lineales, $g_i = p_{i1}p_{i2} \cdots p_{in_i}$, el conjunto $\{p_{2j_2}, \ldots, p_{dj_d}\}$ es linealmente independiente para cualquier combinación de factores lineales escogidos y el sistema $\forall i \geq 2,\ g_i = 0$ tiene exactamente k soluciones.

La idea de la demostración es, en base a los $g_i \in K[\xi_2, \ldots \xi_d]$ construir d polinomios, $f_1, \ldots, f_d \in K[\xi_1, \ldots \xi_d]$ del grado necesario que sean producto de los factores lineales apropiados para obtener las h soluciones al sistema. La ventaja de intentar la condición más fuerte de que todos los polinomios involucrados sean producto de factores lineales es que, al ser la evaluación un homomorfismo, podemos ir buscando las soluciones de forma combinatoria, en base a qué factores anulan. Para ello consideremos los siguientes

polinomios:

$$f_1 = \xi_1 \prod_{j=2}^{n_1} (\beta_i \xi_1 + p_{1j}),$$
$$f_i = \prod_{j=1}^{n_i} (\alpha_{ij} \xi_i + p_{ij}),\ \forall i \geq 2.$$

Donde $\beta_i, \alpha_{ij} \in K$ y $p_{1j} \in K[\xi_2, \ldots, \xi_d]$ son elementos a determinar para que se cumpla que:

1. El conjunto $\{q_1, \ldots, q_d\}$, donde q_i es factor lineal de f_i, es linealmente independiente.
2. El sistema $f_i = 0, \forall i$ tiene exactamente h soluciones.

Las otras dos condiciones no hace falta demostrarlas pues, por construcción, los f_i ya son producto de factores lineales y tienen grado n_i, respectivamente. Empecemos probando la independencia lineal. Para ello estudiaremos en dos casos separados.

Empecemos suponiendo que el factor lineal escogido para el polinomio f_1 es $q_1 = \xi_1$. Tenemos que probar que, independientemente de qué factores lineales q_j escojamos, :

$$\lambda_1 \xi_1 + \lambda_2 q_2 + \cdots + \lambda_d q_d = 0 \iff \lambda_1 = \cdots = \lambda_d = 0.$$

Notemos que los q_i, con $i \geq 2$, son de la forma $q_i = \alpha_{ij} \xi_i + p_{ij}$, para algún $j \in \{1, \ldots, n_i\}$. Como los p_{ij} son polinomios en $\xi_2, \ldots, \xi_d$, no se van a cancelar con los términos que tienen ξ_1, luego se tienen que cancelar entre ellos. Sin embargo, por hipótesis, tenemos que son linealmente independientes, luego la única forma de que se anulen es si $\lambda_2 = \cdots = \lambda_d = 0$. Pero eso nos lleva a que $\lambda_1 \xi_1 = 0$ y por tanto también tenemos que $\lambda_1 = 0$.

Notemos que el conjunto de polinomios lineales $p \in K[\xi_2, \ldots, \xi_d]$ es un espacio vectorial generado por $\{1, \xi_2, \ldots, \xi_d\}$ como base, luego es de dimensión d. Esto nos va a permitir escoger los p_{1j} de tal manera que el conjunto siga siendo linealmente independiente, independientemente de cuáles sean los polinomios p_{il_i} para todo $i \geq 2$. En efecto, cada conjunto $\{p_{il_i}\}_{i=2}^d$ genera un subespacio de dimensión $d-1$ (al ser linealmente independientes) y solo hay un número finito de dichos subespacios que podemos formar, por tanto basta con coger $n_1 - 1$ polinomios fuera de dichos subespacios. Por tanto, siguiendo un razonamiento similar a cuando $q_1 = \xi_1$, tenemos que los polinomios $\beta_j \xi_1 + p_{1j}, \alpha_{2j_2} \xi_1 + p_{2j_2}, \ldots, \alpha_{dj_d} \xi_1 + p_{dj_d}$ son linealmente independientes al serlo el conjunto $\{p_{ij_i}\}_{i=1}^d$.

Además, si tomamos los coeficientes de los polinomios $\{p_{ij_i}\}_{i=1}^d$, al ser estos linealmente independientes, tendremos que es una matriz de d filas y $d-1$ columnas de rango máximo, es decir, rango $d-1$. Al estudiar los coeficientes de los factores q_i, estamos añadiendo la columna de los ξ_1 y como el rango de la matriz anterior es $d-1$ y tiene d filas, al añadir una columna podemos aumentar el rango hasta rango d. Con un razonamiento similar al anterior, podemos escoger los β y los α de tal forma que ninguna combinación de ellos caiga en la unión finita de subespacios $d-1$ dimensionales independientemente de los factores lineales escogidos. De esta forma, nos aseguramos de que la matriz del sistema

lineal $q_i = 0, \forall i \in \{1, \ldots, d\}$ es cuadrada y de rango máximo. Por tanto, el sistema va a tener una única solución para cada elección de los q_i. Además, siguiendo una vez más el razonamiento de evitar un número finito de subespacios, podemos escoger las α y β asegurándonos de que no haya $d+1$ factores lineales de los f_i tales que el sistema lineal que formen tenga solución. Lo cual lleva inmediatamente a que todas las soluciones únicas encontradas para cada posible elección de los q_i son distintas entre sí.

Ahora, notemos que para que se cumpla nuestro sistema de ecuaciones polinómicas, $f_i = 0, \forall i$, al menos uno de los factores lineales, q_i, de cada polinomio, f_i se tiene que anular. Si el factor anulado $q_1 \neq \xi_1$, ya hemos visto que va a haber tantas soluciones como posibles elecciones haya para cada q_i, pero ese número es $n_i, \forall i \geq 2$ y $n_1 - 1$ para q_1. Por tanto el número de soluciones en este caso es $(n_1 - 1)n_2 \cdots n_d$. Por otro lado, si $q_1 = \xi_1 = 0$, tenemos que $f_i(0, \xi_2, \ldots, \xi_d) = g_i, \forall i \geq 2$. Por hipótesis, ese sistema tenía k soluciones. Con todo ello concluimos que el sistema de ecuaciones polinómicas $f_i = 0, \forall i$ tiene exactamente $(n_1 - 1)n_2 \cdots n_d + k$ soluciones. Por definición, ese número es exactamente h, quedando así demostrado el teorema. □

Corolario 2.2.18 (Gordon-Motzkin, [GM65, p.225])**.** *Sea D un anillo de división no conmutativo centralmente finito con $[D : Z(D)] = d < \infty$. Dados $n \geq 1$ y $h \in \mathbb{N}$ tales que $h \leq n^d$. Existe un polinomio $f \in D\langle x \rangle$ con* $\deg f = n$ *y exactamente h raíces.*

Demostración. Vamos a comprobar que este teorema es un caso concreto del teorema anterior. Primero notemos que $Z(D)$ es necesariamente un cuerpo infinito, pues de no serlo, como $[D : Z(D)] < \infty$, D también sería finito y, por el teorema 1.2.8, un cuerpo; sin embargo, estamos suponiendo que D es no conmutativo. Por tanto, por el teorema anterior (2.2.17), tomando $K = Z(D)$ y $n_1 = n_2 = \cdots = n_d = n$, como $d > 1$, tenemos que para cada $0 \leq h \leq n^d$ podemos encontrar d polinomios $f_i \in Z(D)[\xi_1, \ldots, \xi_d]$, con $\deg f_i = n$, tales que el sistema de ecuaciones que forman tiene exactamente h soluciones. Por el teorema 2.2.13, tenemos que dichos f_i nos determinan un $f \in D\langle x \rangle$ tal que $f = \sum_{i=1}^{d} f_i e_i$, con $\deg f \leq n$ y cada solución del sistema, $\lambda_1, \ldots, \lambda_d \in Z(D)$, nos determina de forma única un elemento $a_\lambda = \lambda_1 e_1 + \cdots + \lambda_d e_d \in D$ con $f(a_\lambda) = \sum_{i=1}^{d} f_i(\lambda_1, \ldots, \lambda_d) e_i = 0$. Como hemos visto anteriormente, dicho sistema tiene exactamente h soluciones, luego f tendrá h raíces. □

Nota: Notemos que, en particular, este corolario nos dice que un anillo de división centralmente finito no puede ser algebraicamente cerrado para los polinomios generales.

Capítulo 3

Polinomios sobre cuaternios.

A lo largo del trabajo hemos estado estudiando los anillos de división y los polinomios definidos sobre ellos de forma general. En cambio, cerraremos el trabajo tomando como caso concreto a los cuaternios. La primera sección tratará de una introducción a las propiedades básicas que tienen los cuaternios; en la segunda veremos, mediante los teoremas de Niven y Baer, que para los polinomios a izquierda "esencialmente" el único anillo de división centralmente finito y algebraicamente cerrado que existe es el de los cuaternios; en la tercera desarrollaremos el método de Janovská-Opfer para el cálculo de raíces de estos polinomios; y en la cuarta daremos la clasificación que dieron Janovská y Opfer para las raíces de un subconjunto de los polinomios generales para los cuaternios.

3.1. Introducción a los cuaternios.

Unos de los ejemplos de anillos de división no conmutativos más reconocidos son los cuaternios. El concepto de cuaternio fue introducido por William Rowan Hamilton en 1844 [Ham44] donde buscaba generalizar al espacio tridimensional la representación geométrica del plano modelada por los números complejos. Usando notación moderna, Hamilton los definió de la siguiente manera. Tomemos el $\mathbb{R}$-espacio vectorial $\mathbb{H} = \mathbb{R}1+\mathbb{R}i+\mathbb{R}j+\mathbb{R}k$ donde $\{1, i, j, k\}$ son linealmente independientes, 1 es la unidad de $\mathbb{R}$ y definimos el *producto de Hamilton* en base a las siguientes normas que afectan a la base $\{1, i, j, k\}$ y a los escalares $\lambda \in \mathbb{R}$:

- $\forall \lambda \in \mathbb{R}$, $\lambda i = i\lambda$, $\lambda j = j\lambda$ y $\lambda k = k\lambda$
- $i^2 = j^2 = k^2 = -1$.
- $ij = k$, $jk = i$ y $ki = j$.
- $ji = -k$, $kj = -i$ e $ik = -j$.

Con estas normas, dados dos elementos, $v = a + bi + cj + dk, v' = a' + b'i + c'j + d'k \in \mathbb{H}$ cualesquiera definimos su producto siguiendo la propiedad distributiva y las normas

anteriores:

$$\begin{aligned}(a+bi+cj+dk)(a'+b'i+c'j+d'k) &= aa'+ab'i+ac'j+ad'k+a'bi\\ &+bb'i^2+bc'ij+bd'ik+a'cj+b'cji+cc'j^2+cd'jk\\ &+a'dk+b'dki+c'dkj+dd'k^2\\ &=(aa'-bb'-cc'-dd')+(ab'+a'b+cd'-c'd)i\\ &+(ac'-bd'+a'c+b'd)j+(ad'+bc'-b'c+da')k\in\mathbb{H}.\end{aligned}$$

Con esto es inmediato ver que $\mathbb{H}$ con la suma componente a componente y el producto de Hamilton forman un anillo no conmutativo, pues las unidades imaginarias, i, j y k, son anticonmutativas. Para ver que, además, forma un anillo de división, definimos el conjugado de $v = a + bi + cj + dk$ como $\overline{v} = a - bi - cj - dk$. Es inmediato probar que $0 \leq v\overline{v} = a^2 + b^2 + c^2 + d^2 \in \mathbb{R}$ y la igualdad se da si y solo si $a = b = c = d = 0$. A la raíz cuadrada de este elemento la denominamos *norma* y se denota por $\|v\| = \sqrt{v\overline{v}}$. Ahora, sea $0 \neq v \in \mathbb{H}$. Como $\|v\| \neq 0$, podemos definir $v^{-1} := \frac{\overline{v}}{\|v\|}$. Es inmediato ver que $vv^{-1} = 1 = v^{-1}v$, por lo que $\mathbb{H}$ es un anillo de división no conmutativo.

Completemos ahora la definición de norma, conjugado e inverso con otras dos definiciones bastante útiles a la hora de trabajar con cuaternios y veamos unas propiedades al respecto.

Definición 3.1.1. Sea $a = a_1 + a_2i + a_3j + a_4k \in \mathbb{H}$ un cuaternio.

- Llamaremos *parte real de* a a $\mathcal{R}(a) = a_1 \in \mathbb{R}$.
- Llamaremos *parte imaginaria de* a a $\mathcal{I}(a) = a_2 \in \mathbb{R}$.

Nota: Dado $a = a_1 + a_2i + a_3j + a_4k \in \mathbb{H}$, lo habitual es definir $\mathcal{I}(a) = a_2i + a_3j + a_4k$, en este trabajo hemos tomado la decisión de usar la definición anterior, introducida por Janovská y Opfer en [JO10a], pues va a resultar muy útil a la hora de desarrollar su teoría en las siguientes secciones.

Proposición 3.1.2. *Sean $a, b \in \mathbb{H}$. Se cumple que $\mathcal{R}(ab) = \mathcal{R}(ba)$*

Demostración. Se sigue inmediatamente de aplicar el producto de Hamilton, pues:

$$\mathcal{R}(ab) = a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4 = \mathcal{R}(ba)$$ □

Proposición 3.1.3. *Sean $a, b \in \mathbb{H}$. Tenemos que $\overline{ab} = \overline{b}\overline{a}$.*

Demostración. Usando el producto de Hamilton por un lado obtenemos que:

$$\begin{aligned}\overline{ab} =&(a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4) - (a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3)i\\ &- (a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2)j - (a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1)k.\end{aligned}$$

Mientras que por el otro:

$$\begin{aligned}\overline{b}\overline{a} =&(a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4) + (-b_1a_2 - b_2a_1 + b_3a_4 - b_4a_3)i \\ &+ (-b_1a_3 - b_2a_4 - b_3a_1 + b_4a_2)j + (-b_1a_4 + b_2a_3 - b_3a_2 - b_4a_1)k\end{aligned}$$

Reordenando factores y sacando un (-1) factor común en los términos en i, j y k es inmediato ver que ambos productos son iguales. □

Proposición 3.1.4. *Sean $a, b \in \mathbb{H}$. Tenemos que:*

$$\|ab\| = \|ba\| = \|a\| \, \|b\| .$$

Demostración. Como $\forall a \in \mathbb{H}$, $\|a\| \in \mathbb{R}$, basta con demostrar que $\|ab\| = \|a\| \, \|b\|$ y aplicar conmutatividad.

$$\|ab\| = \sqrt{ab\overline{ab}} \underset{3.1.3}{=} \sqrt{a(b\bar{b})\overline{a}} \underset{b\bar{b}\in\mathbb{R}}{=} \sqrt{a\overline{a}}\sqrt{b\bar{b}} = \|a\| \, \|b\| . \quad \square$$

Proposición 3.1.5. *Sean $x, y \in \mathbb{H}$. Tenemos que:*

$$\overline{x}y + \overline{y}x = 2\mathcal{R}(xy).$$

Demostración. Aplicando el producto de Hamilton es fácil observar lo siguiente:

- $\mathcal{R}(\overline{x}y) = \mathcal{R}(\overline{y}x)$ y,
- $\overline{x}y - \mathcal{R}(\overline{x}y) = -\left(\overline{y}x - \mathcal{R}(\overline{y}x)\right)$.

De estas dos observaciones se sigue inmediatamente el resultado. □

Si sustituimos $\mathbb{R}$ por cualquier cuerpo $\mathbb{F}$, podemos generalizar este concepto al de *álgebra de cuaternios*, aunque este álgebra, en general, no será un anillo de división. Sin embargo, en el caso concreto de que estemos trabajando con un cuerpo real cerrado, R, el álgebra de los cuaternios resultante sí será un anillo de división. Empecemos con la definición.

Definición 3.1.6 (Cuerpo real cerrado)**.** Un cuerpo real cerrado, R, es un cuerpo en el que se cumple cualquiera de las siguientes propiedades equivalentes:

1. R es un cuerpo ordenado en el que se cumple que para todo $a > 0$ existe existe $b \in R$ tal que $a = b^2$ y que todo polinomio $f \in R[x]$ de grado impar tiene al menos una raíz en R.
2. R no es algebraicamente cerrado pero su clausura algebraica es una extensión finita.
3. R no es algebraicamente cerrado pero $R\left(\sqrt{-1}\right) = R[i]$ sí lo es.
4. R satisface las mismas sentencias de primer orden que los reales. En cuyo caso diremos que R es *elementalmente equivalente* a $\mathbb{R}$.

Nota: Es importante resaltar que se suelen definir los cuerpos reales cerrados mediante la primera propiedad, después se demuestra que esos cuerpos son elementalmente equivalentes a los reales y de ahí se deducen las otras dos equivalencias. Dicha demostración excede los objetivos del trabajo. Sin embargo, vamos a dar en el apéndice A la idea básica de por qué los cuerpos real cerrados son elementalmente equivalentes a los reales siguiendo los resultados de [TZ12] sin dar demostraciones.

Definición 3.1.7. Sea R un cuerpo real cerrado. Definimos el *álgebra de cuaternios sobre* R como el R-espacio vectorial, $H := R1 + Ri + Rj + Rk$, donde $\{1, i, j, k\}$ con el producto de Hamilton.

Proposición 3.1.8. *Sea R un cuerpo real cerrado y H su álgebra de cuaternios. Entonces H es un anillo de división no conmutativo.*

Demostración. Como R es un cuerpo ordenado, $\forall a \in R$, $a^2 \geq 0$ y $\sum a_i^2 = 0 \iff \forall i$, $a_i = 0$. Por tanto, dado $v = a + bi + cj + dk \in H \setminus \{0\}$ podemos definir el conjugado $\overline{v} = a - bi - cj - dk$ y la norma $\|v\| = \sqrt{v\overline{v}}$, que está bien definida pues $a^2 + b^2 + c^2 + d^2 > 0$. Por tanto, definimos $v^{-1} = \frac{\overline{v}}{\|v\|}$ y es inmediato comprobar que $vv^{-1} = 1 = v^{-1}v$. □

Nota: Como tanto el producto, como la norma, como el inverso están definidos de la misma forma para el álgebra de cuaternios H sobre R que para los cuaternios de Hamilton $\mathbb{H}$, tenemos que todos los resultados de esta sección tienen su análogo para H.

3.2. Clausura algebraica de los polinomios a izquierda.

En esta sección daremos la caracterización de los anillos de división algebraicamente cerrados *para polinomios a izquierda* (resp. polinomios a derecha) demostrada por Niven en [Niv41], apoyándonos en las demostraciones actualizadas y simplificadas expuestas por Lam en [Lam01]. Empecemos definiendo qué entendemos como anillo de división algebraicamente cerrado.

Definición 3.2.1. Sea D un anillo de división. Diremos que D es *algebraicamente cerrado a derecha* si todo polinomio, $f(x) \in D[x]$, tiene al menos una raíz a derecha en D.

Notemos que por el teorema 2.1.3 esta afirmación es equivalente a decir que $f(x)$ descompone como producto de factores lineales.

Podemos definir de forma análoga el concepto de *algebraicamente cerrado a izquierda* con polinomios $f(x) \in [x]D$, raíces a izquierda y el equivalente al teorema 2.1.3 para polinomios a derecha. Si D es algebraicamente cerrado a derecha y a izquierda, diremos que D es algebraicamente cerrado.

Teorema 3.2.2 (Niven-Jacobson, [Lam01, p.255]**.** *Sea R un cuerpo real cerrado y D el anillo de división de cuaternios sobre R. Entonces D es algebraicamente cerrado.*

Demostración. Consideremos la aplicación $\varphi : D \to D$; $\varphi(q) = \overline{q}$. Es inmediato que $\forall \lambda \in R$, $\varphi(\lambda q) = \lambda \varphi(q)$. Tomemos ahora dos elementos $q = a + bi + cj + dk, r = a' + b'i + c'j + d'k \in D$. Notemos que:

$$\begin{aligned}\varphi(q) + \varphi(r) &= a - bi - cj - dk + a' - b'i - c'j - d'k \\ &= (a + a') - (b + b')i - (c + c')j - (d + d')k = \varphi(q + r).\end{aligned}$$

Por otro lado, tenemos que:

$$\begin{aligned}\varphi(r)\varphi(q) &= (a' - b'i - c'j - d'k)(a - bi - cj - dk) \\ &= (a'a - b'b - c'c - d'd) + (-a'b - b'a + c'd - d'c)i \\ &\quad + (-a'c - b'd - c'a + d'b)j + (-a'd + b'c - c'b - d'a)k \\ &= (aa' - bb' - cc' - dd') - (ab' + a'b + cd' - c'd)i \\ &\quad - (ac' - bd' + a'c + b'd)j - (ad' + bc' - b'c + da')k = \varphi(qr).\end{aligned}$$

Luego φ es un anti-isomorfismo R-lineal que deja fijo a R.

Es inmediato extender φ a $D[x]$ mediante la conjugación de los coeficientes. Es decir, dado $f = \sum q_r x^r \in D[x]$, definimos $\overline{f} = \sum \overline{q_r} x^r \in D[x]$. Notemos que, de forma análoga a como lo hemos demostrado anteriormente en la proposición 3.1.3, tenemos que $\overline{fg} = \overline{g}\,\overline{f}$ y, en particular, $\overline{f\overline{f}} = \overline{\overline{f}f} = f\overline{f}$. Como la conjugación deja fijo a $f\overline{f}$, tenemos que $f\overline{f} \in R[x]$.

Con estos preliminares demostrados, procederemos a demostrar que f tiene al menos una raíz mediante inducción sobre su grado.

Si $\deg f = 1$, tenemos que $f = qx + r$, con $q \neq 0$, luego $x = -\frac{r}{q}$ es claramente una raíz de f. Supongamos ahora que $\deg f = n \geq 2$. Teníamos que $\overline{f}f \in R[x]$. Como R es real cerrado, $R(i)$ es algebraicamente cerrado y por tanto $\exists \alpha \in R(i) \subset D$ raíz de $\overline{f}f$.

Por el lema 2.1.4, o bien α es una raíz de f, en cuyo caso hemos terminado, o bien $\beta \in D$, elemento conjugado de α, es una raíz de $\overline{f}$. Es decir, $\sum_{r=0}^{n} \overline{q_r}\beta^r = 0$. Conjugando a ambos lados de la ecuación, obtenemos que $\sum_{r=0}^{n} \overline{\beta}^r q_r = 0$. Luego $\overline{\beta}$ es una raíz a izquierda de f. Usando el análogo del teorema 2.1.3 para raíces a izquierda, tenemos que $f(x) = (x - \overline{\beta})g(x)$ con $g \in D[x]$ y $\deg g = n - 1$. Por hipótesis de inducción, existe $\gamma \in D$, tal que $g(\gamma) = 0$. Por el teorema 2.1.3, $\exists h \in D[x]$ tal que $g(x) = h(x)(x - \gamma)$, luego $f(x) = (x - \overline{\beta})h(x)(x - \gamma)$ y volviendo a aplicar el teorema 2.1.3, vemos que γ es una raíz (a derecha) de $f(x)$. Por tanto D es algebraicamente cerrado a derecha.

De forma análoga, usando en cada caso los teoremas "complementarios" cuando se hace referencia a raíces a izquierda o a derecha, se demuestra que D es algebraicamente cerrado a izquierda y por tanto, algebraicamente cerrado. □

Para la demostración del recíproco del teorema necesitaremos una versión generalizada del teorema de Frobenius extraída de [Bre14], que a su vez necesitará de varios lemas técnicos. Para estos lemas y el teorema de Frobenius denotaremos por R un cuerpo real cerrado y por D una R-álgebra de división con $\dim_R D = n < \infty$.

Lema 3.2.3. *Para todo $d \in D$ existe $\lambda \in R$ tal que $d^2 + \lambda d \in R$.*

Demostración. Como $\dim_R D = n$, el conjunto $\{1, d, \ldots, d^n\}$ es linealmente dependiente. Es decir, podemos afirmar que existe un polinomio mónico $f \in R[x]$ tal que $f(\alpha) = 0$. Como $R[i]$ es la clausura algebraica de R, tenemos que todo polinomio en $R[x]$ va a descomponer en un producto de factores de grado 1 o grado 2, es decir:

$$f(x) = (x - a_1) \cdots (x - a_m)(x^2 + b_1 x + c_1) \cdots (x^2 + b_l x + c_l)$$

Como D es un álgebra de división, en particular no hay divisores de cero. Por tanto, como $f(\alpha) = 0$, $\exists i$ tal que, o bien $\alpha = a_i \in R$ o bien $\alpha^2 + b_i \alpha = -c_i \in R$. En ambos casos se cumple la conclusión. □

Lema 3.2.4. *El conjunto $V = \{v \in D \mid v^2 \in R, v^2 \leq 0\}$ es un subespacio vectorial de D y $D = R \oplus V$.*

Demostración. Es inmediato ver que $R \cap V = \{0\}$, pues $\forall a \in R, a^2 \geq 0$; y que V es cerrado por multiplicación de escalares, pues si $v \in V$ y $\lambda \in R$, como $\lambda^2 \geq 0$ y $v^2 \leq 0$, $(\lambda v)^2 = \lambda^2 v^2 \leq 0$, por tanto $\lambda v \in V$.

A continuación veamos que dados $u, v \in V$, $u + v \in V$. Al estar trabajando en espacios vectoriales, podemos asumir sin pérdida de generalidad que u, v son linealmente independientes. Comprobemos que $\{u, v, 1\}$ es un conjunto linealmente independiente mediante reducción al absurdo. Si fuesen linealmente dependientes, existirían $\lambda, \mu \in R$ tales que $1 + \lambda u = \mu v$. Elevando ambos lados de la igualdad al cuadrado obtenemos $1 + \lambda^2 u^2 + 2\lambda u = \mu^2 v^2$, lo cual obligaría a que $2\lambda u \in R$, pero esto es solo posible si y solo si $\lambda = 0$, pues $2\lambda u \in V$ y $V \cap R = \{0\}$, y entonces $\mu v = 1$ implica $\mu = 0$ al forzar $\mu v \in V \cap R = \{0\}$, llegando a una contradicción, pues $0 \neq 1$. Luego $\{1, u, v\}$ son linealmente independientes.

Por el lema 3.2.3 existen $\lambda, \mu \in R$ tales que $(u+v)^2 + \lambda(u+v) \in R$ y $(u-v)^2 + \mu(u-v) \in R$. Además, operando obtenemos que $(u + v)^2 + (u - v)^2 = 2u^2 + 2v^2$ que está en R pues $u, v \in V$. Sustituyendo el valor de $(u + v)^2$ y $(u - v)^2$ de las dos primeras expresiones en la tercera y operando adecuadamente obtenemos:

$$(\lambda + \mu)u + (\lambda - \mu)v + \gamma 1 = 0, \quad \text{para un cierto } \gamma \in R.$$

Pero u, v y 1 eran linealmente independientes, por lo que $\lambda + \mu = \lambda - \mu = 0$. Esto sucede si y solo si $\lambda = \mu = 0$, sin embargo, recordemos que los λ y μ nos los daba el lema 3.2.3, por lo que $(u + v)^2, (u - v)^2 \in R$. Ahora bien, supongamos que $(u + v)^2 > 0$. Como R es un cuerpo real cerrado, tenemos que existirá un $\alpha \in R$ tal que $(u+v)^2 = \alpha^2$, factorizando y despejando obtenemos que, necesariamente, $u + v = \pm\alpha \in R$, luego $u + v \mp \alpha 1 = 0$.

Absurdo, pues como se ha demostrado anteriormente $\{1, u, v\}$ es un conjunto linealmente independiente, luego $(u+v)^2 \leq 0$ y, por tanto, $(u+v) \in V$.

Con esto hemos terminado de demostrar que V es un subespacio vectorial de D, quedando únicamente por demostrar que dado $d \in D$, existen $v \in V$ y $\lambda \in R$ tales que $d = \lambda + v$. Sea $x \in D \setminus R$, por el lema 3.2.3 tenemos que existe $\lambda \in R$ tal que $x^2 + \lambda x = \alpha \in R$. Completando cuadrados obtenemos que $(x + \frac{\lambda}{2})^2 = \alpha - \frac{\lambda^2}{4} \in R$. Si suponemos que $(x+\frac{\lambda}{2})^2 > 0$, como R es real cerrado, al seguir el mismo razonamiento que en el párrafo anterior, obtenemos que $x + \frac{\lambda}{2} \in R$, es decir $x \in R$, con lo que llegamos a una contradicción. Por tanto, $(x + \frac{\lambda}{2})^2 \leq 0$ y $x + \frac{\lambda}{2} \in V$. Es decir, existirá $v \in V$ tal que $x = -\frac{\lambda}{2} + v$. Esto, junto a que $V \cap R = \{0\}$ nos lleva a concluir que $D = R \oplus V$. □

Dados $u, v \in D$, definimos el *producto de Jordan* $u \circ v = uv + vu$. Como D es un anillo de división, es inmediato ver que el producto de Jordan es conmutativo, $\forall u, v, w \in D$, $u \circ (v + w) = u \circ v + u \circ w$ y R-bilineal. Además, se cumple el siguiente lema:

Lema 3.2.5. *Para todo $u, v \in V$, $u \circ v \in R$.*

Demostración. Al igual que antes, supongamos sin pérdida de generalidad que u, v son linealmente independientes. Como hemos visto en la demostración del lema 3.2.4, teníamos que $(u+v)^2, (u-v)^2 \in R$. Por tanto:

$$\begin{aligned} u \circ v = uv + vu &= \frac{1}{2}\left[(u^2 + uv + vu + v^2) - (u^2 - uv - vu + v^2)\right] \\ &= \frac{1}{2}\left[(u+v)^2 - (u-v)^2\right] \in R. \end{aligned}$$ □

Lema 3.2.6. *Si $n = dim_R D > 2$, entonces existen tres elementos $i, j, k \in D$ tales que:*

$$i^2 = j^2 = k^2 = -1, \tag{1}$$

$$ij = -ji = k,\ jk = -kj = i,\ ki = -ik = j. \tag{2}$$

Además, $1, i, j, k$ son linealmente independientes.

Demostración. Por el lema 3.2.4 tenemos que $\dim_R V = n - 1 > 1$. Por tanto, tomamos $v, w \in V$ linealmente independientes. Definimos $u := w - \frac{w \circ v}{v \circ v} v \in V$. Notemos que $u \neq 0$, pues v y w son linealmente independientes, y:

$$u \circ v = \left(w - \frac{w \circ v}{v \circ v} v\right) \circ v = w \circ v - \frac{w \circ v}{v \circ v}(v \circ v) = 0,$$

por lo que $uv = -vu$.

Visto esto definimos $i := \frac{1}{\sqrt{-u^2}} u$, $j := \frac{1}{\sqrt{-v^2}} v$ y $k := ij$. Es inmediato ver que $i^2 = -\frac{u^2}{u^2} = -1$, de forma análoga, $j^2 = -1$ y $k^2 = (ij)^2 = ijij = -i^2 j^2 = -1$ (pues i y j, anticonmutan), comprobando que cumplen las ecuaciones (1). Además:

$$ij = k \implies jij = jk \implies -j^2 i = i = jk, \text{ y}$$

$$ij = k \implies iji = ki \implies -i^2 j = j = ki.$$

Obteniendo las ecuaciones que faltan mediante la anticonmutividad de i y de j, comprobamos que también se cumplen las ecuaciones (2).

Por último, dados $\lambda_1, \ldots, \lambda_4 \in R$, consideremos la siguiente igualdad:

$$(\lambda_1 + \lambda_2 i + \lambda_3 j + \lambda_4 k)(\lambda_1 - \lambda_2 i - \lambda_3 j - \lambda_4 k) = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_4^2. \tag{3}$$

Notemos que (3) es un número real no negativo que es igual a cero si y solamente si $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$. Esto en particular nos dice que $1, i, j, k$ son linealmente independientes, pues en caso contrario existirían unos λ_i, no todos nulos, que anularían la parte izquierda de (3) pero no la derecha, llegando a una contradicción. □

Teorema 3.2.7 (Frobenius, generalizado)**.** *Sea R un cuerpo real cerrado, $C := R[\sqrt{-1}]$ su clausura algebraica y $H := R \oplus Ri \oplus Rj \oplus Rk$ el álgebra de cuaternios de R. Sea D un álgebra de división sobre R de dimensión finita. Se cumple que $\dim_R D = 1, 2$ ó 4 y entonces D es isomorfo, respectivamente, a R, a C o a H.*

Demostración. Si $n = 1$, $D = R$. Si $n = 2$, por el lema 3.2.4 $V \neq \{0\}$, luego existirá un $v \in V$ tal que $v^2 < 0$, luego existirá un elemento $0 \neq \alpha \in R$ tal que $-v^2 = \alpha^2$. Sea $i := \frac{v}{\alpha}$, notemos que $i^2 = -1$ y que $V = \langle i \rangle$. Por tanto $D = R + Ri = R[i] = R[\sqrt{(}-1)] = C$, pues R es un cuerpo real cerrado.

Por el lema 3.2.6, tenemos que si $\dim_R D > 2$, al existir 4 elementos linealmente independientes, necesariamente $\dim_R D \geq 4$. Concretamente, si $n = 4$, $D = \langle 1, i, j, k \rangle = R \oplus Ri \oplus Rj \oplus Rk$. Esto, junto al hecho de que i, j, k cumplan las ecuaciones (1, 2), nos lleva a que $D \cong H$.

Supongamos entonces que $n > 4$ y procedamos por reducción al absurdo. Sean i, j, k los elementos del lema 3.2.6. En este caso, como $\dim_R V > 3$, existirá un elemento $v \in V$ linealmente independiente de i, j, k. Con v, construimos el elemento $e = v + \frac{i \circ v}{2} i + \frac{j \circ v}{2} j + \frac{k \circ v}{2} k \in V$. Notemos que e anticonmuta con i, j y k. En efecto:

$$i \circ e = i \circ v + \frac{i \circ v}{2} i \circ i + \frac{j \circ v}{2} i \circ j + \frac{k \circ v}{2} i \circ k = i \circ v + \frac{i \circ v}{2}(-2) + \frac{j \circ v}{2} 0 + \frac{k \circ v}{2} 0 = i \circ v - i \circ v = 0.$$

Por tanto i y e anticonmutan. El resto de casos, $j \circ e = k \circ e = 0$, se demuestran de forma análoga.

Sin embargo, esto nos lleva a afirmar que $ek = eij = -iej = ije = ke = -ek$. Esto solo es posible si $ek = 0$, sin embargo $e \neq 0$ y $k \neq 0$ y, al ser D un álgebra de división, no puede haber divisores de 0.

Por tanto, $n = \dim_R D$ solo puede tomar los valores de $1, 2$ ó 4, cuyos casos ya hemos estudiado anteriormente. □

Teorema 3.2.8 (Baer, [Lam01, p.255])**.** *Sea D un anillo de división no conmutativo centralmente finito con centro R tal que todo polinomio no constante de $R[x]$ tiene una raíz en D. Entonces R es un cuerpo real cerrado y D es el anillo de división de los cuaternios sobre R.*

Demostración. Sea $n = \dim_R D$. Como para cada polinomio irreducible $f(x) \in R[x]$ no constante hay una raíz $\alpha \in D$ tenemos que $R \subset R(\alpha) \subset D$ y por tanto:

$$\deg f = \dim_R R(\alpha) \leq \dim_R D = n \tag{4}$$

Veamos que R es un cuerpo perfecto usando la siguiente caracterización: *"o bien char* $R = 0$*, o bien si char* $R = p > 0$*, entonces todo elemento de* R *es una potencia p-ésima".*

Supongamos entonces que *char* $R = p > 0$, con p primo. Sean $\alpha \in R$ y C una clausura algebraica de R. Para todo $m \in \mathbb{N}$, sea $\alpha^{\frac{1}{p^m}}$ la única raíz p^m-ésima de α en C. Notemos que, debido a (4), dado $\beta \in C$, existirá un polinomio irreducible $m_\beta \in R[x]$ de grado menor o igual a n con $m_\beta(\beta) = 0$, con lo que $[R(\beta) : R] \geq n$. Por tanto las dimensiones de las extensiones $R\left(\alpha^{\frac{1}{p^m}}\right)$ están acotadas y tenemos que la cadena $R\left(\alpha^{\frac{1}{p}}\right) \subset R\left(\alpha^{\frac{1}{p^2}}\right) \subset \cdots$ estabiliza. Luego existe m tal que $\alpha^{\frac{1}{p^{m+1}}} \in R\left(\alpha^{\frac{1}{p^m}}\right)$. Pero esto implica que $\alpha^{\frac{1}{p}} \in R^{p^m}(\alpha) \subset R$, por tanto R es perfecto.

Veamos ahora que R es real cerrado. Para ello usaremos la caracterización de que los cuerpos reales cerrados son exactamente aquellos que no son algebraicamente cerrados pero su clausura es una extensión finita. De todas las extensiones simples de R en C, tomamos como K la extensión de mayor dimensión, que existe pues como hemos dicho anteriormente, están acotadas. Supongamos que $\exists \beta \in C \setminus K$ y procedamos por reducción al absurdo. Como R es un cuerpo perfecto, $K(\beta)$ es una extensión finita y separable de R. Por tanto, podemos aplicar el teorema del elemento primitivo [1] y afirmar que $K(\beta)$ es una extensión simple de R. Sin embargo, necesariamente $\dim_R K(\beta) > \dim_R K$, lo cual contradice la maximalidad de K supuesta anteriormente. Por reducción al absurdo, tenemos que $C = K$ y por tanto C es una extensión finita de R. Además, $R \neq C$, de lo contrario $C = D$ pero en las hipótesis del enunciado tenemos que D no es conmutativo. Luego tenemos que R no es algebraicamente cerrado y que $\dim_R C < \infty$, por tanto, R es un cuerpo real cerrado.

Para terminar la demostración, veamos que D es el álgebra de los cuaternios sobre R. Por el teorema 3.2.7 tenemos que $D \cong R, C$ o H, sin embargo D es no conmutativo, por lo que, necesariamente, $D \cong H$. □

Combinando los teoremas 3.2.2 y 3.2.8 obtenemos el siguiente teorema de clasificación.

Nota: Notemos que aunque en 3.2.8 los polinomios que aparecen son sobre $Z(D)$ y no sobre D, como $Z(D)[x]$ está contenido en $D[x] \cup [x]D$ esto es suficiente para establecer que un anillo de división centralmente finito y algebraicamente cerrado sólo puede ser un anillo de cuaternios sobre un cuerpo real cerrado.

---

[1] Demostración en [Wei09]

Teorema 3.2.9 (Clasificación de anillos de división algebraicamente cerrados.)**.** *Los anillos de división centralmente finitos no conmutativos que son algebraicamente cerrados son únicamente los anillos de división de cuaternios sobre cuerpos reales cerrados.*

3.3. Cálculo de raíces de polinomios simples en $\mathbb{H}$.

En la sección anterior hemos demostrado el teorema de Niven-Jacobson que nos permite afirmar que todo polinomio a izquierda y todo polinomio a derecha (no constantes) sobre los cuaternios, $\mathbb{H}$, tienen raíz. Sin embargo este teorema es un teorema de existencia, no da ninguna información sobre cómo son dichas raíces o cómo calcularlas. En una sección posterior de [Niv41], Niven proporciona un método para su cálculo, sin embargo, a lo largo de los más de 80 años desde que se publicó el artículo, ese método se ha ido adaptando y actualizando. En esta sección expondremos el método de Janovská-Opfer publicado en [JO10a] y la generalización que dieron en [JO10b] para el cálculo de raíces en polinomios que denominaron "biláteros".

Recordemos que, para la clasificación de las raíces en polinomios a izquierda sobre un anillo de división, en los teoremas 2.1.5 y 2.1.6 eran fundamentales las clases de conjugación. Por ello, empecemos explicitándolas en el caso de los cuaternios y dando una propiedad fundamental de las mismas.

Definición 3.3.1. Sean $a, b \in \mathbb{H}$. Diremos que a *y* b *son equivalentes* si pertenecen a la misma clase de conjugación. Es decir, $a \sim b \iff \exists h \in \mathbb{H} \setminus \{0\}$, $a = h^{-1}bh$. Denotaremos la clase de conjugación de un elemento $a \in \mathbb{H}$ por $[a]$.

Proposición 3.3.2. *Sean* $a, b \in \mathbb{H}$.

$$a \sim b \text{ si y solo si } \mathcal{R}(a) = \mathcal{R}(b) \text{ y } |a| = |b|.$$

En particular $a \sim \overline{a}$.

Demostración. Empecemos demostrando que $a \sim b$ implica $\mathcal{R}(a) = \mathcal{R}(b)$ y $|a| = |b|$.

Por definición de $b \in [a]$ tenemos que $\exists h \in \mathbb{H} \setminus \{0\}$ tal que $a = h^{-1}bh$. Por tanto $\mathcal{R}(a) = \mathcal{R}\left(h^{-1}bh\right)$, sin embargo, por la proposición 3.1.2, tenemos que $\mathcal{R}\left(h^{-1}(bh)\right) = \mathcal{R}\left(bhh^{-1}\right) = \mathcal{R}(b)$. Por otro lado, aplicando la proposición 3.1.4, tenemos que:

$$\|a\| = \left\|h^{-1}bh\right\| = \left\|bh^{-1}h\right\| = \|b\|.$$

Pongámonos en las hipótesis del recíproco y comprobemos que existe un $h \in \mathbb{H} \setminus \{0\}$ tal que $ha = bh$. Definimos $u := a - b$ y $v := a + b$. Como $\mathcal{R}(a) = \mathcal{R}(b)$, tenemos que $\mathcal{R}(u) = u_1 = 0$.

Por otro lado, aplicando el producto de Hamilton con $h \in \mathbb{H}$ arbitrario:

$$\begin{aligned} 0 = ha - bh = {} & (h_1(a_1 - b_1) - h_2(a_2 - b_2) - h_3(a_3 - b_3) - h_4(a_4 - b_4)) \\ & + (h_1(a_2 - b_2) + h_2(a_1 - b_1) + h_3(a_4 + b_4) - h_4(a_3 + b_3))i \\ & + (h_1(a_3 - b_3) - h_2(a_4 + b_4) + h_3(a_1 - b_1) + h_4(a_2 + b_2))j \\ & + (h_1(a_4 - b_4) + h_2(a_3 + b_3) - h_3(a_2 + b_2) + h_4(a_1 - b_1))k \end{aligned}$$

Notemos, sin embargo, que resolver esta ecuación para h es lo mismo que resolver el sistema lineal homogéneo $M\tilde{h} = 0$, donde:

$$M = \begin{pmatrix} 0 & -u_2 & -u_3 & -u_4 \\ u_2 & 0 & v_4 & -v_3 \\ u_3 & -v_4 & 0 & v_2 \\ u_4 & v_3 & -v_2 & 0 \end{pmatrix}, \text{ y}$$

$$\tilde{h} = \begin{pmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{pmatrix}.$$

Al ser un sistema homogéneo, tenemos que $\tilde{h} = \overline{0}$ siempre va a ser una solución, pero nos interesa que existan soluciones con $h \neq 0$. Esto solo ocurrirá si el sistema es compatible indeterminado, es decir, si y solo si $\det M = 0$. Notemos que:

$$\det M = u_2^2v_2^2 + u_3^2v_3^2 + u_4^2v_4^2 + 2\,u_2v_2u_3v_3 + 2\,u_2v_2u_4v_4 + 2\,u_3v_3u_4v_4 = (u_2v_2 + u_3v_3 + u_4v_4)^2.$$

Sin embargo, la condición $\|a\| = \|b\|$ implica que $u_2v_2 + u_3v_3 + u_4v_4 = 0$. En efecto, tenemos que $\|a\| = \|b\| \iff \|a\|^2 = \|b\|^2 \iff 0 = \|a\|^2 - \|b\|^2$. Por tanto:

$$0 = \sum_{i=1}^{4} a_i^2 - b_i^2 = \sum_{i=1}^{4} (a_i - b_i)(a_i + b_i) = \sum_{i=1}^{4} u_iv_i \underset{u_1=0}{=} u_2v_2 + u_3v_3 + u_4v_4.$$

Luego concluimos que el sistema lineal homogéneo tiene más soluciones además de la trivial y por tanto $\exists h \in \mathbb{H} \setminus \{0\}$ tal que $ha = bh$. □

Con esta proposición podemos reescribir cómo son las clases de conjugación en $\mathbb{H}$ de la siguiente forma:

$$[a] = \{h \in \mathbb{H} \,|\, \mathcal{R}(z) = \mathcal{R}(a) \wedge \|z\| = \|a\|\}.$$

Esto nos permite dar dos corolarios con respecto a la geometría de las clases de conjugación.

Lema 3.3.3. *Si $a \in \mathbb{R} \subset \mathbb{H}$ entonces $[a] = \{a\}$.*

Demostración. Sea $a \in \mathbb{R}$. Como $\mathbb{R} = Z(\mathbb{H})$, para todo $h \in \mathbb{H}$ se tiene

$$hah^{-1} = ahh^{-1} = a,$$

de donde $[a] = \{a\}$. □

Corolario 3.3.4. *Si $a \in \mathbb{H} \setminus \mathbb{R}$, entonces podemos interpretar $[a]$ como una esfera en $\mathbb{R}^4$.*

Demostración. Sea $a \in \mathbb{H} \setminus \mathbb{R}$. Por la proposición 3.3.2 tenemos que $\forall z \in [a]$, $\mathcal{R}(z) = \mathcal{R}(a)$ y $\|z\| = \|a\|$. La primera igualdad nos da la ecuación $z_1 - a_1 = 0$, que es una ecuación de un hiperplano de $\mathbb{R}^4$; mientras que la segunda igualdad nos da la ecuación $z_1^2 + z_2^2 + z_3^2 + z_4^2 = \|a\|$, que es precisamente la ecuación de una hiperesfera de dimensión 3 en $\mathbb{R}^4$. Además, el hiperplano y la hiperesfera no son tangentes, pues $\overline{a} \neq a$ debido a que $a \notin \mathbb{R}$ y $\overline{a} \in [a]$ y por tanto hay al menos dos puntos en la intersección. Luego su intersección nos va a dar una esfera en $\mathbb{R}^4$. □

Proposición 3.3.5. *Sea $a \in \mathbb{H} \setminus \mathbb{R}$. Hay exactamente dos números complejos $z_1, z_2 \in \mathbb{C}$ tales que $z_1, z_2 \in [a]$. Además, $z_1 = \overline{z_2}$.*

Demostración. Sean $a = a_1 + a_2 i + a_3 j + a_4 k \in \mathbb{H} \setminus \mathbb{R}$ y $z = x + yi \in \mathbb{C}$ tales que $z \in [a]$. Por la proposición 3.3.2, tenemos que $a_1 = \mathcal{R}(a) = \mathcal{R}(z) = x$ y $\|a\| = \|z\|$. Esto último es cierto si y solo si $\|a\|^2 = a_1^2 + a_2^2 + a_3^2 + a_4^2 = x^2 + y^2 = \|z\|^2$. Sustituyendo y despejando la y obtenemos que, necesariamente:

$$y = \pm\sqrt{a_2^2 + a_3^2 + a_4^2}.$$

Por tanto, solo pueden existir dos números complejos en $[a]$:

$$z_1 = a_1 + i\sqrt{a_2^2 + a_3^2 + a_4^2}$$
$$z_2 = \overline{z_1} = a_1 - i\sqrt{a_2^2 + a_3^2 + a_4^2}.$$

Por el razonamiento inicial, tenemos que $z_1, z_2 \in [a]$ y además $z_1 \neq z_2$, pues $a_2^2 + a_3^2 + a_4^2 = 0$ si y solo si $a_2 = a_3 = a_4 = 0$, lo que contradice $a \notin \mathbb{R}$. □

Definición 3.3.6. Dado $a \in \mathbb{H} \setminus \mathbb{R}$, llamaremos *representante complejo de* $[a]$ al z_1 definido en la proposición anterior.

De ahora en adelante denotaremos por $p_n = \sum_{i=0}^{n} a_i z^i \in \mathbb{H}[z]$ a un polinomio genérico de grado n. Supondremos sin pérdida de generalidad que $a_n = 1$ y también supondremos que 0 no es una raíz de p_n, es decir, $a_0 \neq 0$.

Definición 3.3.7. Sea $z_0 \in \mathbb{H}$ una raíz de p_n. Si $z_0 \neq \mathbb{R}$ y se cumple que $\forall z \in [z_0]$, $p_n(z) = 0$ diremos que z_0 genera una raíz esférica. Por el contrario, si $z_0 \in \mathbb{R}$ o $\exists z \in [z_0]$ tal que $p_n(z) \neq 0$, diremos que z_0 es una raíz aislada.

Nota: Como se ha visto en los teoremas de Gordon-Motzkin (2.1.5 y 2.1.6) las raíces van a estar agrupadas en a lo sumo n clases de conjugación, que dentro de cada clase puede tener una o infinitas raíces. Como vamos a ver a continuación, las clases con una única raíz se corresponden con las clases de las raíces aisladas, mientras que las clases con infinitas raíces se corresponden con las clases de elementos que generan raíces esféricas. Por ello, cometeremos un abuso de notación; diremos que z_0 *es* una raíz esférica, en lugar de decir que la genera; y contaremos el número de raíces de un polinomio como el número de clases de equivalencia en las que tiene raíces, haciendo referencia a raíces esféricas o aisladas para distinguir entre los casos de una raíz frente a infinitas.

Ya sabemos que calcular potencias enteras sucesivas en $\mathbb{C}$ es extremadamente sencillo. El siguiente resultado nos dará una forma de calcularlas para elementos $z \in \mathbb{H}$.

Proposición 3.3.8. *Sea $z \in \mathbb{H}$. Para todo $l \in \mathbb{Z}$ existen $\alpha_l, \beta_l \in \mathbb{R}$ tales que $z^l = \alpha_l z + \beta_l$. Concretamente, $\forall l \geq 0$, α_l y β_l vienen dados por la siguiente relación de recurrencia:*

$$\alpha_0 = 0, \beta_0 = 1, \tag{a}$$

$$\alpha_{l+1} = 2\mathcal{R}(z)\alpha_l + \beta_l, \tag{b}$$

$$\beta_{l+1} = -\|z\|^2 \alpha_l. \tag{c}$$

Demostración. Notemos que solo hace falta demostrarlo para todo $l \geq 0$, pues $z^{-l} = \left(z^{-1}\right)^l$. Por tanto, procedamos mediante inducción sobre $l \geq 0$ usando la relación de recurrencia.

- Caso base. Si $l = 0$, tenemos que $z^0 = \alpha_0 z + \beta_0 = 0z + 1 = 1$.
- Hipótesis de inducción (HI). Sea $0 \leq m \in \mathbb{Z}$. Vamos a suponer que $\forall l$ tal que $0 \leq l \leq m$ tenemos que $z^l = \alpha_l z + \beta_l$.
- Consideremos el caso $l = m + 1$:

$$z^{m+1} = zz^m \overset{HI}{=} z(\alpha_m z + \beta_m) = \alpha_m z^2 + \beta_m z. \tag{d}$$

Notemos que, si $z = a+bi+cj+dk$, entonces $z^2 = a^2-b^2-c^2-d^2+2a(bi+cj+dk) = -(a^2+b^2+c^2+d^2)+2a(a+bi+cj+dk) = -\|z\|^2+2\mathcal{R}(z)z$. Por tanto, sustituyendo en (d) obtenemos:

$$z^{m+1} = \cdots = \alpha_m \left(-\|z\|^2 + 2\mathcal{R}(z)z\right) + \beta_m z = \left(2\mathcal{R}(z)\alpha_m + \beta_m\right) z - \|z\|^2 \alpha_m$$
$$\underset{\text{(b) y (c)}}{=} \alpha_{m+1} z + \beta_{m+1}.$$

□

Debido a esta proposición, podemos evaluar el polinomio p_n en $z \in \mathbb{H}$ de la siguiente forma:

$$p_n(z) = \sum_{i=0}^{n} a_i z^i = \sum_{i=0}^{n} a_i(\alpha_i z + \beta_i) = \left(\sum_{i=0}^{n} \alpha_i a_i\right) z + \sum_{i=0}^{n} \beta_i a_i = A(z)z + B(z)$$

Donde $A(z)$ y $B(z)$ son cuaternios que dependen de z.

Teorema 3.3.9 (Janovská-Opfer, [JO10b, p.248])**.** *Fijemos un elemento $z_0 \in \mathbb{H}$. Entonces $A(z) = A_{z_0} \in \mathbb{H}$ y $B(z) = B_{z_0} \in \mathbb{H}$ para todo $z \in [z_0]$, donde $A(z)$ y $B(z)$ son los definidos anteriormente. Si z_0 es una raíz de p_n, entonces:*

$$p_n(z_0) = A(z)z_0 + B(z) = 0, \quad \forall z \in [z_0].$$

Además, se cumple que $\forall z \in [z_0]$, $A(z) = 0 \iff B(z) = 0$. Por último si $z_0 \notin \mathbb{R}$ y $A(z_0) = 0$, z_0 es una raíz esférica, mientras que si $A(z_0) \neq 0$ entonces z_0 es una raíz aislada.

Demostración. Notemos que los coeficientes α_j, β_j dependen únicamente de $\mathcal{R}(z)$ y de $\|z\|$. Por tanto, $\forall z \in [z_0]$ tanto dichos coeficientes como $A(z)$ y $B(z)$ se van a mantener constantes.

Como $p_n(z_0) = 0$ y $z_0 \neq 0$, al ser $a_0 \neq 0$, tenemos que $A(z_0) = 0 \iff B(z_0) = 0$. Por tanto, $\forall z \in [z_0], A(z) = 0 \iff B(z) = 0$. Esto nos lleva a afirmar que, si $A(z_0) = 0$, entonces $p_n(z) \equiv 0, \quad \forall z \in [z_0]$. Luego si $z_0 \notin \mathbb{R}$, estamos en las condiciones de afirmar que z_0 es una raíz esférica.

Por otro lado, si $A(z_0) \neq 0 \neq B(z_0)$, tenemos que $z_0 = -A(z_0)^{-1}B(z_0)$ es la única raíz de p_n en $[z_0]$. □

En el teorema anterior hemos visto que, en el caso $z_0 \notin \mathbb{R}$ si $A(z_0) = 0$, toda la clase de conjugación, $[z_0]$, es una raíz; mientras que, si $A(z_0) \neq 0$, solo hay una. Por el teorema 2.1.6, ya sabemos que si en una misma clase de conjugación hay más de una raíz, necesariamente hay infinitas, pero en el caso de $\mathbb{H}$ podemos demostrarlo de forma más sencilla obteniendo un resultado más fuerte, como vamos a ver a continuación. Además, notemos que en la definición que hemos dado de raíz aislada, a priori, si $z_0 \notin \mathbb{R}$ solo hemos exigido que haya algún $z \in [z_0]$ que no sea raíz. Por lo tanto, este teorema también nos dice que ser una raíz aislada es una condición más fuerte de la que habíamos exigido en un primer momento.

Por último, notar que, una vez hemos fijado z, como $A(z)$ y $B(z)$ no dependen del representante escogido de la clase $[z]$, de ahora en adelante, si no da lugar a confusión, los expresaremos directamente como A, B y reservaremos las expresiones $A(z)$ y $B(z)$ para cuando estemos hablando de la z como variable del polinomio p_n.

Corolario 3.3.10 (Gordon-Motzkin, versión cuaternios)**.** *Sean $z_0, z_1 \in \mathbb{H}$ raíces distintas de p_n con $z_1 \in [z_0]$. Entonces z_0 es una raíz esférica.*

Demostración. Tenemos que, por un lado $p_n(z_0) = Az_0 + B = 0$ y, por otro, $p_n(z_1) = Az_1 + B = 0$. Restando ambas ecuaciones obtenemos que $A(z_0 - z_1) = 0$. Como en $\mathbb{H}$ no hay divisores de cero y $z_0 \neq z_1$, obtenemos que $A = 0$. Por el teorema 3.3.9, z_0 es una raíz esférica. □

Corolario 3.3.11. *Sea $z_0 \in \mathbb{H}$ una raíz aislada de p_n. Entonces z_0 es la única raíz de $[z_0]$.*

Demostración. Si z_0 es una raíz aislada pueden suceder dos cosas, que $z_0 \in \mathbb{R}$ o no. El primer caso es trivialmente cierto pues $\forall z \in \mathbb{R}, [z] = \{z\}$. El segundo, por otro lado, está estudiado en el teorema 3.3.9 y hemos llegado a la conclusión que z_0 era la única raíz de $[z_0]$. □

Al igual que usando el método de Niven, para el método de Janovská-Opfer es necesario definir un polinomio con coeficientes en $\mathbb{R}$ que nos ayudará a sondear las posibles soluciones del polinomio p_n. A continuación daremos la definición de dicho polinomio y un lema que nos permite relacionarlo con los valores de $A(z)$ y $B(z)$ de p_n.

Definición 3.3.12 (Polinomio acompañante)**.** Sea $p_n \in \mathbb{H}[z]$ definido anteriormente. Llamaremos *polinomio acompañante* de p_n al polinomio, $q_{2n} \in \mathbb{R}[z]$ de grado $2n$ con raíces en $\mathbb{C}$ definido de la siguiente manera:

$$q_{2n}(z) := \sum_{j,k=0}^{n} \overline{a_j} a_k z^{j+k}.$$

Lema 3.3.13. *Sea $p_n(z) = A(z)z + B(z)$. Entonces:*

$$q_{2n}(z) = \|A(z)\|^2 z^2 + 2\mathcal{R}(\overline{A(z)}B(z))z + \|B(z)\|^2.$$

Demostración. Tenemos que:

$$\begin{aligned} q_{2n}(z) &= \sum_{j=0}^{n} \overline{a_j} \left(\sum_{k=0}^{n} a_k z^k \right) z^j = \sum_{j=0}^{n} \overline{a_j} \left(A(z)z + B(z) \right) z^j \underset{3.3.8}{=} \sum_{j=0}^{n} \overline{a_j} \left(A(z)z + B(z) \right) (\alpha_j z + \beta_j) \\ &= \sum_{j=0}^{n} (\alpha_j \overline{a_j}) A(z) z^2 + \sum_{j=0}^{n} (\beta_j \overline{a_j}) A(z) z + \sum_{j=0}^{n} (\alpha_j \overline{a_j}) B(z) z + \sum_{j=0}^{n} (\beta_j \overline{a_j}) B(z). \end{aligned}$$

Notemos que por la definición de $A(z)$ y $B(z)$ y que $\overline{a+b} = \overline{a} + \overline{b}, \forall a, b \in \mathbb{H}$, tenemos que $\overline{A(z)} = \sum_{i=0}^{n} \alpha_i \overline{a_i}$ y $\overline{B(z)} = \sum_{i=0}^{n} \beta_i \overline{a_i}$. Por tanto

$$\begin{aligned} q_{2n}(z) &= \overline{A(z)} A(z) z^2 + \left(\overline{A(z)} B(z) + \overline{B(z)} A(z) \right) z + \overline{B(z)} B(z) \\ &\underset{3.1.5}{=} \|A(z)\|^2 z^2 + 2\mathcal{R}(\overline{A(z)} B(z)) z + \|B(z)\|^2. \end{aligned}$$

□

Nota: Es importante recordar que, a menos que hayamos fijado una clase de conjugación en $\mathbb{H}$, $A(z)$ y $B(z)$ *no* son constantes, aunque, para cada $z \in \mathbb{H}$ siempre van a ser un cuaternio, luego estamos justificados en usar las propiedades de la conjugación y la norma en cuaternios.

La idea de los siguientes resultados es que *todas las raíces de p_n están "asociadas" a, al menos, una raíz de q_{2n}*, para el cual, al ser un polinomio en $\mathbb{C}$ con coeficientes en $\mathbb{R}$, sabemos encontrar las raíces. Es decir, las raíces de q_{2n} no tienen por qué ser raíces de p_n, pero a partir de ellas vamos a encontrar *todas las de p_n*. Sin embargo, en el caso de las raíces reales de hecho se cumple una condición más fuerte, pues las raíces reales de q_{2n} son exactamente todas las raíces reales de p_n.

Teorema 3.3.14. *Sea $z_0 \in \mathbb{R}$. Entonces:*

$$q_{2n}(z_0) = 0 \iff p_n(z_0) = 0.$$

Demostración. Notemos que $\forall z \in \mathbb{R}$ se cumple que:

$$\|p_n(z)\|^2 = p_n(z)\overline{p_n(z)} = (Az+B)(\overline{A}\overline{z}+\overline{B}) = \|A\|^2\|z\|^2 + Az\overline{B} + B\overline{A}\overline{z} + \|B\|^2$$
$$\underset{z\in\mathbb{R}}{\overset{3.1.5}{=}} \|A\|^2 z^2 + 2\mathcal{R}(\overline{A}B)z + \|B\|^2 \underset{3.3.13}{=} q_{2n}(z).$$

De aquí se sigue inmediatamente que $\forall z_0 \in \mathbb{R}, q_{2n}(z_0) = 0 \iff p_n(z_0) = 0$. □

A continuación vamos a centrarnos en las raíces esféricas, que también son un caso muy sencillo con las herramientas que tenemos ya demostradas.

Teorema 3.3.15. *Sea $z_0 \in \mathbb{C}\setminus\mathbb{R}$ una raíz de q_{2n} con $A(z_0) = 0$. Entonces z_0 es una raíz esférica de p_n.*

Demostración. Sea $z_0 \in \mathbb{C}$ en las hipótesis del teorema. Tenemos que si $A = 0$, por el lema 3.3.13, entonces $0 = q_{2n}(z_0) = B$. Por tanto, $p_n(z_0) = Az_0 + B = 0 \cdot z_0 + 0 = 0$. Con lo que z_0 es una raíz no real de p_n con $A = 0$. Por el teorema 3.3.9, tenemos que z_0 es una raíz esférica de p_n. □

En siguiente lugar, centrémonos en las raíces $z \notin \mathbb{R}$ de q_{2n} cuyo $A(z)$ asociado no se anule. El elemento z puede que no sea raíz de p_n, pero si existiese un $z_0 \in [z]$ tal que $p_n(z_0) = 0$, entonces sabemos por el teorema 3.3.9 que $z_0 = -A^{-1}B$. El primer resultado nos garantizará que si construimos así el z_0, entonces $z_0 \in [z]$ y el segundo nos permitirá afirmar que ese z_0 es, de hecho, una raíz aislada de p_n.

Lema 3.3.16. *Sea $z \in \mathbb{C}\setminus\mathbb{R}$ una raíz de q_{2n} con $A(z) \neq 0$. Entonces $z_0 := -A^{-1}B \in [z]$.*

Demostración. Sea $z \in \mathbb{C}$ en las condiciones del enunciado. Por el lema 3.3.13 tenemos que z cumple la siguiente ecuación:

$$\|A(z)\|^2 z^2 + 2\mathcal{R}(\overline{A(z)}B(z))z + \|B(z)\|^2 = 0.$$

Para mejorar la legibilidad, eliminamos el argumento de A y de B e identificamos $\overline{A}B =: v = v_1 + v_2 i + v_3 j + v_4 k$. Además, como $z \notin \mathbb{R}$ tenemos que $z = z_1 + z_2 i$ con $z_2 \neq 0$. De esta forma, si sustituimos y separamos la ecuación anterior en su parte real e imaginaria se nos queda en las dos siguientes ecuaciones:

$$\|A\|^2 (z_1^2 - z_2^2) + 2v_1 z_1 + \|B\|^2 = 0, \tag{e}$$

$$2z_2\left(\|A\|^2 z_1 + v_1\right) = 0 \underset{z_2\neq 0}{\iff} \|A\|^2 z_1 + v_1 = 0. \tag{f}$$

Tenemos que $z_0 = -A^{-1}B = -\frac{\overline{A}B}{\|A\|^2} = -\frac{v}{\|A\|^2}$. Por tanto, tenemos que:

$$\mathcal{R}(z_0) = -\frac{v_1}{\|A\|^2} \underset{\text{Eq. (f)}}{=} z_1 = \mathcal{R}(z).$$

Además, tenemos que:

$$\|z_0\| = \left\|-\frac{v}{\|A\|^2}\right\| = \frac{\left\|\overline{A}B\right\|}{\|A\|^2} = \frac{\|B\|}{\|A\|}.$$

Por otro lado, si despejamos v_1 en (f) y sustituimos en (e), obtenemos que:

$$0 = \|A\|^2 (z_1^2 - z_2^2) - \|A\|^2 z_1^2 + \|B\|^2 = -\|A\|^2 (z_1^2 + z_2^2) + \|B\|^2 = -\|A\|^2 \|z\|^2 + \|B\|^2 .$$

De donde se deduce inmediatamente que $\|z\|^2 = \frac{\|B\|^2}{\|A\|^2} \iff \|z\| = \frac{\|B\|}{\|A\|} = \|z_0\|$. Por la proposición 3.3.2 tenemos que $z_0 \in [z]$. □

Para el siguiente teorema, para todo $a = a_1 + a_2 i + a_3 j + a_4 k \in \mathbb{H}$ denotaremos por $\vec{a} := (a_2, a_3, a_4) \in \mathbb{R}^3$ y definimos la norma de un vector como la norma euclídea, $\|\vec{a}\| := \sqrt{a_2^2 + a_3^2 + a_4^2}$.

Teorema 3.3.17 (Janovská-Opfer, [JO10a, p.250]). *Sea $p_n \in \mathbb{H}[z]$ y q_{2n} su polinomio acompañante. Sea $z \in \mathbb{C} \setminus \mathbb{R}$ con $A(z) \neq 0$ una raíz de q_{2n}. Entonces el z_0 definido en el lema 3.3.16 es una raíz aislada de p_n.*

Además, podemos expresar z_0 en función de z_1, z_2 y $v := \overline{A}B$ de la siguiente manera:

$$z_0 = z_1 - \lambda (v_2 i + v_3 j + v_4 k) ,$$

donde $\lambda = \frac{|z_2|}{\|\vec{v}\|} \in \mathbb{R}$.

Demostración. Pongámonos en las hipótesis del teorema, por el lema 3.3.16, tenemos que $z_0 \in [z]$, por tanto $A(z_0) = A$ y $B(z_0) = B$. Por construcción tenemos que $z_0 = -A^{-1}B$, luego $p_n(z_0) = A\left(-A^{-1}B\right) + B = -B + B = 0$ y por tanto concluimos que z_0 es una raíz de p_n.

Notemos ahora que $z_0 \in \mathbb{R} \iff \vec{z_0} = \vec{0} \iff \|\vec{z_0}\| = 0$. Sin embargo, notemos que $\vec{z_0} = -\frac{1}{\|A\|^2}\vec{v}$. Por tanto, como $\|z_0\|^2 = z_1^2 + \|\vec{z_0}\|^2 = z_1^2 + \frac{1}{\|A\|^4}\|\vec{v}\|^2 = \|z\|^2 = z_1^2 + z_2^2$ y $z_2 \neq 0$, obtenemos que:

$$\|\vec{z_0}\| = \frac{1}{\|A\|^2}\|\vec{v}\| = |z_2| \neq 0.$$

Luego nos encontramos en las condiciones del teorema 3.3.9 y tenemos que z_0 es una raíz aislada de p_n.

Además, al despejar obtenemos que $\frac{1}{\|A\|^2} = \frac{|z_2|}{\|\vec{v}\|} = \lambda$. Por tanto:

$$z_0 = -\frac{1}{\|A\|^2}v = -\frac{v_1}{\|A\|^2} - \lambda (v_1 i + v_2 j + v_3 k) \underset{\text{Eq. (f)}}{=} z_1 - \lambda (v_2 i + v_3 j + v_4 k) . \qquad \square$$

Por último nos falta de comprobar que, con las raíces complejas de q_{2n} y los teoremas anteriores, realmente hemos encontrado todas las raíces de p_n.

Teorema 3.3.18 (Janovská-Opfer, [JO10a, p.250]). *Sea $p_n \in \mathbb{H}$ y z_0 una raíz de p_n. Entonces existe un $z \in \mathbb{C}$ tal que $z \in [z_0]$ y $q_{2n}(z) = 0$.*

Demostración. Por el teorema 3.3.14 sabemos que p_n y q_{2n} comparten raíces reales. Supongamos entonces que $z_0 \in \mathbb{H} \setminus \mathbb{R}$.

Empecemos suponiendo que $A := A(z_0) = 0$. Como $A = 0 = B$, tenemos que $\forall z \in [z_0]$, $q_{2n}(z) = 0$. Por la proposición 3.3.5 sabemos que hay dos números $z, \overline{z} \in \mathbb{C}$ que son raíces de q_{2n}.

Por tanto solo nos queda estudiar el caso $A \neq 0 \neq B$. Notemos entonces que, si escribimos q_{2n} en términos de A y B usando el lema 3.3.13, tenemos que, $\forall z \in [z_0]$, como A y B son constantes, $q_{2n}(z)$ es un polinomio de segundo grado con coeficientes reales. Aplicando la fórmula para el cálculo de soluciones de ecuaciones polinómicas de segundo grado obtenemos que:

$$z_{1,2} = \frac{-2\mathcal{R}(\overline{A}B) \pm \sqrt{4\mathcal{R}(\overline{A}B)^2 - 4\left\|A\right\|^2\left\|B\right\|^2}}{2\left\|A\right\|^2} = \frac{-\mathcal{R}(\overline{A}B) \pm i\sqrt{\left\|\overline{A}B\right\|^2 - \mathcal{R}(\overline{A}B)^2}}{\left\|A\right\|^2}.$$

Además, como $\forall h \in \mathbb{H}$, $|\mathcal{R}(h)| \leq \|h\|$, tenemos que el radicando no es negativo. Además, como hemos visto en la demostración del teorema 3.3.17, $z_0 \in \mathbb{R} \iff \overline{A}B \in \mathbb{R}$. Por tanto, $\left\|\overline{A}B\right\|^2 - \mathcal{R}(\overline{A}B)^2 > 0$ y z_1, z_2 son raíces complejas no reales conjugadas.

Por último, falta de comprobar que ambas raíces son conjugadas de z_0, pues, al haber utilizado que $A(z)$ y $B(z)$ son constantes en toda la clase de conjugación, si no perteneciesen a la misma, el razonamiento no sería válido. Para ello aplicaremos la proposición 3.3.2 a z_1 y z_0, pues al ser $z_2 = \overline{z_1}$, tenemos que $z_2 \in [z_1]$. Notemos que, como $z_0 = -\frac{\overline{A}B}{\|A\|^2}$, tenemos que $\mathcal{R}(z_1) = -\frac{\mathcal{R}(\overline{A}B)}{\|A\|^2} = \mathcal{R}(z_0)$. De igual forma, para la norma obtenemos que:

$$\|z_1\|^2 = \frac{\mathcal{R}(\overline{A}B)^2}{\|A\|^4} + \frac{\|A\|^2\|B\|^2 - \mathcal{R}(\overline{A}B)^2}{\|A\|^4} = \frac{\|B\|^2}{\|A\|^2} = \|z_0\|^2.$$

Por tanto $\|z_1\| = \|z_0\|$ y concluimos que $z_1, z_2 \in [z_0]$. □

Teorema 3.3.19. *Sean* $\tilde{p}_n = \sum_{i=0}^n z^i a_i \in \mathbb{H}$ *y* $p_n = \sum_{i=0}^n \overline{a_i} z^i$. *Entonces:*

$$p_n(z) = 0 \iff \tilde{p_n}(\overline{z}) = 0.$$

Además, ambos polinomios tienen las mismas raíces reales y esféricas.

Demostración. Notemos que, para todo $z \in \mathbb{H}$ $p_n(z) = \overline{\tilde{p}_n(\overline{z})}$. Por tanto las raíces de un polinomio van a ser las raíces conjugadas del otro.

Además, como $\forall z \in \mathbb{R}$, $\overline{z} = z$ y $\forall z \in \mathbb{H}$, $\overline{z} \in [z]$, tenemos que las raíces reales y esféricas coinciden en ambos polinomios. □

Como hemos visto, al igual que otros resultados del trabajo, el método de Janovská-Opfer no solo es aplicable a los polinomios a izquierda, sino a los polinomios a derecha. Para simplificar la notación, Janovská y Opfer llaman *polinomios simples* en [JO10a] a $\mathbb{H}[z] \cup [z]\mathbb{H}$. De esta forma, también los distinguen de los *polinomios biláteros* que definen y estudian en [JO10b] y nosotros trataremos en la siguiente sección. Este es el motivo de que esta sección este titulada "Cálculo de raíces de polinomios simples en $\mathbb{H}$".

3.4. Clasificación de raíces de polinomios biláteros en $\mathbb{H}$

Como hemos visto en la sección 2.2, podemos definir lo que es un polinomio sobre $\mathbb{H}$ de una forma más general, centrándonos en el anillo de polinomios generales $\mathbb{H}\langle z\rangle$. En general no va a ser cierto que todo polinomio $f \in \mathbb{H}\langle z\rangle$ tiene al menos una raíz en $\mathbb{H}$, basta con aplicar el corolario 2.2.18 para $h = 0$. Sin embargo, en [EN44], Eilenberg y Niven llegan al siguiente resultado:

Teorema 3.4.1 (Eilenberg-Niven, [EN44, p.1])**.** *Sea $f = a_0 z a_1 z \cdots z a_n + \phi \in \mathbb{H}\langle z\rangle$, donde $\phi \in \mathbb{H}\langle z\rangle$ es un polinomio general con $\deg \phi < n$. Entonces f tiene al menos una raíz en $\mathbb{H}$.*

Esquema de la demostración. La demostración de este teorema se basa en que, en estas condiciones, si entendemos $\mathbb{H}$ como $\mathbb{R}^4$ con la topología euclídea y realizamos la compactificación de Alexandroff para obtener la esfera 4-dimensional, S^4, podemos ver f como una aplicación continua de S^4 en sí mismo de grado n en sentido Brouwer. Por tanto, la ecuación $f(z) = 0$ tiene, en esencia, 4 soluciones.

Sin embargo, los pormenores de esta demostración, al ser de corte topológico, se escapan a los objetivos de este trabajo, que hemos querido enfocar desde un punto de vista puramente algebraico. □

En esta sección vamos a dar una generalización del método Janovská-Opfer visto en la sección anterior, para un caso concreto de polinomios que cumplen las condiciones del teorema 3.4.1, que publicaron ellos mismos en [JO10b].

Definición 3.4.2. Sea $p \in \mathbb{H}\langle z\rangle$. Diremos que p es un polinomio bilátero de grado n si es de la forma

$$p(z) = \sum_{i=0}^{n} a_i z^i b_i,$$

con $a_j, b_j \in \mathbb{H}$ tales que $a_n b_n \neq 0$, para asegurar que tiene término de grado n y $a_0 b_0 \neq 0$. Además, podemos suponer sin pérdida de generalidad que $a_n = 1 = b_n$.

De forma similar a la sección anterior, podemos aplicar la proposición 3.3.8 para reescribir p de la siguiente forma:

$$p(z) = \sum_{i=0}^{n} a_i(\alpha_i z + \beta_i) b_i = \sum_{i=0}^{n} \alpha_i a_i z b_i + \sum_{i=0}^{n} \beta_i a_i b_i = C(z) + B(z). \tag{a}$$

Lema 3.4.3. *Sea $z_0 \in \mathbb{H} \setminus \mathbb{R}$. Entonces $B(z)$ es constante $\forall z \in [z_0]$. Además, si $\exists z \in \mathbb{H}$ tal que $p(z) = 0$, entonces $C(z) = 0 \iff B(z) = 0$.*

Demostración. Recordemos que los α_i y β_i, por la proposición 3.3.8, únicamente dependían de $\mathcal{R}(z_0)$ y $\|z_0\|$. Por tanto, por la proposición 3.3.2 van a ser los mismos $\forall z \in [z_0]$, con lo que, en una misma clase de conjugación, $A(z)$ y $B(z)$ van a ser constantes.

La última afirmación es obvia pues si $p(z) = C(z) + B(z) = 0$, tenemos que

$$C(z) = -B(z).$$

□

Nota: $C(z)$ depende de la propia z, no solo de su parte real o su norma. Por tanto, no podemos asegurar que sea constante dentro de la propia clase de conjugación. Para obtener el $A(z)$ correspondiente al teorema 3.3.9, necesitaremos trabajar con los cuaternios en forma matricial.

Definición 3.4.4. Sean $a = a_1 + a_2 i + a_3 j + a_4 k, b = b_1 + b_2 i + b_3 j + b_4 k \in \mathbb{H}$. Definimos las siguientes aplicaciones:

$$\iota_1 : \mathbb{H} \to \mathcal{M}_4(\mathbb{R}) \qquad \iota_2 : \mathbb{H} \to \mathcal{M}_4(\mathbb{R})$$

$$a \mapsto \begin{pmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{pmatrix}. \qquad a \mapsto \begin{pmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & a_4 & -a_3 \\ a_3 & -a_4 & a_1 & a_2 \\ a_4 & a_3 & -a_2 & a_1 \end{pmatrix}.$$

$$\iota_3 : \mathbb{H}^2 \to \mathcal{M}_4(\mathbb{R}) \qquad \mathrm{col}(a) : \mathbb{H} \to \mathbb{R}^4$$

$$(a, b) \mapsto \iota_1(a)\iota_2(b). \qquad a \mapsto \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}.$$

Donde $\mathcal{M}_4(\mathbb{R})$ denota el anillo de matrices cuadradas 4×4 con coeficientes en $\mathbb{R}$.

Estas aplicaciones nos serán útiles a la hora de clasificar los distintos tipos de raíces que aparecerán en los polinomios biláteros. Empecemos viendo diversas propiedades de las mismas.

Proposición 3.4.5. *Las aplicaciones ι_1 e ι_2 cumplen las siguientes propiedades:*

- *Son aplicaciones $\mathbb{R}$-lineales.*
- *$\forall a, b \in \mathbb{H}$ tenemos que $\iota_1(ab) = \iota_1(a)\iota_1(b)$ y $\iota_2(ab) = \iota_2(b)\iota_2(a)$.*
- *$\forall a, b \in \mathbb{H}$, $\iota_1(a) = \iota_2(b) \iff a = b \in \mathbb{R}$.*
- *$\forall a \in \mathbb{H}$, $\forall i \in \{1, 2\}$, $\iota_i(a)\iota_i(a)^T = \iota_i(a)^T\iota_i(a) = \|a\|^2 I_4$. Donde A^T denota la matriz traspuesta de A e I_4 la matriz identidad 4×4.*

Demostración. Sean $a, b \in \mathbb{H}$ y $\lambda, \mu \in \mathbb{R}$. Notemos que:

$$\begin{aligned}
\iota_1(\lambda a + \mu b) &= \begin{pmatrix} \lambda a_1 + \mu b_1 & -\lambda a_2 - \mu b_2 & -\lambda a_3 - \mu b_3 & -\lambda a_4 - \mu b_4 \\ \lambda a_2 + \mu b_2 & \lambda a_1 + \mu b_2 & -\lambda a_4 - \mu b_4 & \lambda a_3 + \mu b_3 \\ \lambda a_3 + \mu b_3 & \lambda a_4 + \mu b_4 & \lambda a_1 + \mu b_1 & -\lambda a_2 + \mu b_2 \\ \lambda a_4 + \mu b_4 & -\lambda a_3 - \mu b_3 & \lambda a_2 + \mu b_2 & \lambda a_1 + \mu b_1 \end{pmatrix} \\
&= \lambda \begin{pmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{pmatrix} + \mu \begin{pmatrix} b_1 & -b_2 & -b_3 & -b_4 \\ b_2 & b_1 & -b_4 & b_3 \\ b_3 & b_4 & b_1 & -b_2 \\ b_4 & -b_3 & b_2 & b_1 \end{pmatrix} \\
&= \lambda \iota_1(a) + \mu \iota_1(b).
\end{aligned}$$

Por tanto, análogamente con ι_2, obtenemos que ambas aplicaciones son lineales.

Denotemos por $v = v_1 + v_2 i + v_3 j + v_4 k = ab = (a_1 b_1 - a_2 b_2 - a_3 b_3 - a_4 b_4) + (a_1 b_2 + a_2 b_1 + a_3 b_4 - a_4 b_3)i + (a_1 b_3 - a_2 b_4 + a_3 b_1 + a_4 b_2)j + (a_1 b_4 + a_2 b_3 - a_3 b_2 + a_4 b_1)k$. Es fácil ver que:

$$\begin{aligned}
\iota_1(a)\iota_1(b) &= \begin{pmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{pmatrix} \begin{pmatrix} b_1 & -b_2 & -b_3 & -b_4 \\ b_2 & b_1 & -b_4 & b_3 \\ b_3 & b_4 & b_1 & -b_2 \\ b_4 & -b_3 & b_2 & b_1 \end{pmatrix} \\
&= \begin{pmatrix} v_1 & -v_2 & -v_3 & -v_4 \\ v_2 & v_1 & -v_4 & v_3 \\ v_3 & v_4 & v_1 & -v_2 \\ v_4 & -v_3 & v_2 & v_1 \end{pmatrix} = \iota_1(v) = \iota_1(ab).
\end{aligned}$$

De forma análoga tenemos que:

$$\begin{aligned}
\iota_2(b)\iota_2(a) &= \begin{pmatrix} b_1 & -b_2 & -b_3 & -b_4 \\ b_2 & b_1 & b_4 & -b_3 \\ b_3 & -b_4 & b_1 & b_2 \\ b_4 & b_3 & -b_2 & b_1 \end{pmatrix} \begin{pmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & a_4 & -a_3 \\ a_3 & -a_4 & a_1 & a_2 \\ a_4 & a_3 & -a_2 & a_1 \end{pmatrix} \\
&= \begin{pmatrix} v_1 & -v_2 & -v_3 & -v_4 \\ v_2 & v_1 & v_4 & -v_3 \\ v_3 & -v_4 & v_1 & v_2 \\ v_4 & v_3 & -v_2 & v_1 \end{pmatrix} = \iota_2(v) = \iota_2(ab).
\end{aligned}$$

Supongamos ahora que $\iota_1(a) = \iota_2(b)$. Tenemos que:

$$\begin{pmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{pmatrix} = \begin{pmatrix} b_1 & -b_2 & -b_3 & -b_4 \\ b_2 & b_1 & b_4 & -b_3 \\ b_3 & -b_4 & b_1 & b_2 \\ b_4 & b_3 & -b_2 & b_1 \end{pmatrix}.$$

Como dos matrices son iguales si y solo si son iguales componente a componente, tenemos que $a_1 = b_1$ y $\forall i \in \{2, 3, 4\}$, $a_i = b_i = 0$. Por tanto $a = a_1 = b_1 = b \in \mathbb{R}$.

Por último, notemos que:

$$\iota_1(a)^T = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ -a_2 & a_1 & a_4 & -a_3 \\ -a_3 & -a_4 & a_1 & a_2 \\ -a_4 & a_3 & -a_2 & a_1 \end{pmatrix} = \iota_1(\overline{a}).$$

Definamos $c \in \mathbb{H}$ por $c = c_1 + c_2 i + c_3 j + c_4 k := \|a\|^2 \in \mathbb{R}$, tenemos que $c_1 = \|a\|^2$ y $c_2 = c_3 = c_4 = 0$. Por lo tanto, obtenemos:

$$\iota_1(a)\iota_1(a)^T = \iota_1(a)\iota_1(\overline{a}) = \iota_1(a\overline{a}) \overset{*}{=} \iota_1(c) = c_1 I_4 = \|a\|^2 I_4.$$

Además, como $a\overline{a} = \overline{a}a$, a partir de $(*)$ obtenemos que:

$$\iota_1(a)\iota_1(a)^T = \iota_1(a\overline{a}) = \iota_1(\overline{a}a) = \iota_1(a)^T\iota_1(a)$$

La demostración para ι_2 es completamente análoga. □

Proposición 3.4.6. *La aplicación col es $\mathbb{R}$-lineal.*

Demostración. Sean $a, b \in \mathbb{H}$ y $\lambda, \mu \in \mathbb{R}$. Sea $v := \lambda a + \mu b$. La demostración se sigue inmediatamente de que $\forall i \in \{1, \ldots, 4\}$, $v_i = \lambda a_i + \mu b_i$. □

Lema 3.4.7. *Sean $a, b, c \in \mathbb{H}$. Tenemos que:*

- *$col(ab) = \iota_1(a)col(b) = \iota_2(b)col(a)$.*
- *$col(abc) = \iota_3(a, c)col(b)$.*

Demostración. Sean $a, b, c \in \mathbb{H}$. Al igual que en la demostración de la proposición 3.4.5, denotaremos por $v := ab = (a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4) + (a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3)i + (a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2)j + (a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1)k$. Tenemos que:

$$\iota_1(a)\mathrm{col}(b) = \begin{pmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} = \begin{pmatrix} a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4 \\ a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3 \\ a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2 \\ a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1 \end{pmatrix} = \mathrm{col}(v).$$

Repitiendo el procedimiento llegamos a que $\iota_2(b)\mathrm{col}(a) = \mathrm{col}(v)$.

Ahora, para demostrar la úlitma igualdad, nos vamos a aprovechar de la que acabamos de demostrar, pues:

$$\mathrm{col}(abc) = \mathrm{col}(a(bc)) = \iota_1(a)\mathrm{col}(bc) = \iota_1(a)\iota_2(c)\mathrm{col}(b) = \iota_3(a, c)\mathrm{col}(b).$$ □

Lema 3.4.8. *Sean $a, b \in \mathbb{H}$. La matriz $\iota_3(a, b)$ es normal y todos sus autovalores tienen la misma norma, $\|a\| \|b\|$.*

Demostración. Como $\forall A \in \mathcal{R}_4(\mathbb{R})$ tenemos que $A^T = \overline{A}^T$, la condición de normalidad para matrices cuadradas con coeficientes reales se reduce a $AA^T = A^TA$. Entonces, como $\iota_3(a,b) \in \mathcal{R}_4(\mathbb{R})$, al usar la proposición 3.4.5, obtenemos que:

$$\begin{aligned}\iota_3(a,b)\iota_3(a,b)^T &= \iota_1(a)\iota_2(b)\left(\iota_1(a)\iota_2(b)\right)^T = \iota_1(a)\iota_2(b)\iota_2(b)^T\iota_1(a)^T \\ &= \iota_1(a)\left\|b\right\|^2 I_4\iota_1(a)^T = \left\|b\right\|^2\iota_1(a)\iota_1(A)^T = \left\|b\right\|^2\left\|a\right\|^2 I_4 \\ &= \left\|a\right\|^2\left\|b\right\|^2 I_4 = \left\|a\right\|^2\iota_2(b)^T\iota_2(b) = \iota_2(b)^T\left\|a\right\|^2 I_4\iota_2(b) \\ &= \iota_2(b)^T\iota_1(a)^T\iota_1(a)\iota_2(b) = \left(\iota_1(a)\iota_2(b)\right)^T\iota_1(a)\iota_2(b) = \iota_3(a,b)^T\iota_3(a,b).\end{aligned}$$

Además, como hemos visto en el párrafo anterior (resaltado en rojo), $\iota_3(a,b)\iota_3(a,b)^T = \left\|a\right\|^2\left\|b\right\|^2 I_4$. Por tanto, tenemos que $A := \frac{1}{\|a\|\|b\|}\iota_3(a,b)$ es ortogonal. Por el teorema espectral para matrices unitarias (ver [Lan02]), de las cuales las ortogonales son un caso particular, tenemos que existirán matrices $P, D \in \mathcal{M}_4(\mathbb{C})$ tales que $A = PDP^{-1}$, donde P es invertible y D es una matriz diagonal cuyas entradas son los autovalores $\lambda_i \in \mathcal{C}, \forall i \in \{1,\dots,4\}$ de A, todos ellos de norma 1. Además, como $A \in \mathcal{M}_4(\mathbb{R})$, tenemos que dichos autovalores deberán ser complejos conjugados dos a dos (si en particular son reales, serán una pareja de $+1$ o una pareja de -1).

De aquí obtenemos inmediatamente que:

$$\iota_3(a,b) = \left\|a\right\|\left\|b\right\|A = \left\|a\right\|\left\|b\right\|PDP^{-1} = P\tilde{D}P^{-1}.$$

Donde $\tilde{D} = \|ab\|\,D$. Por tanto, $\iota_3(a,b)$ es una matriz diagonalizable y todos sus autovalores tienen norma $\|ab\|$. □

Teorema 3.4.9. *Sean $z \in \mathbb{H}$ y $p \in \mathbb{H}\langle z\rangle$ un polinomio bilátero. Si escribimos $p(z) = C(z) + B(z)$ conforme a la ecuación (a) tenemos:*

$$col(p(z)) = A(z)col(z) + col(B(z)),$$

donde $A(z) := \sum_{i=0}^{n}\alpha_i\iota_3(a_i,b_i) \in \mathcal{M}_4(\mathbb{R})$ y $col(B(z)) = \sum_{i=0}^{n}\beta_i col(a_ib_i)$.

Demostración. Recordemos que, por la proposición 3.4.6, col es una aplicación lineal. Por tanto, $\mathrm{col}(p(z)) = \mathrm{col}(C(z)) + \mathrm{col}(B(z))$ y $\mathrm{col}(B(z)) = \mathrm{col}(\sum_{i=0}^{n}\beta_i a_ib_i) = \sum_{i=0}^{n}\beta_i\mathrm{col}(a_ib_i)$. Por tanto solo nos queda comprobar que $\mathrm{col}(C(z)) = A(z)\mathrm{col}(z)$.

Recordemos que $C(z) = \sum_{i=0}^{n}\alpha_i a_i z b_i$. Entonces tenemos que:

$$\begin{aligned}\mathrm{col}(C(z)) &\underset{3.4.6}{=} \sum_{i=0}^{n}\alpha_i\mathrm{col}(a_izb_i) \underset{3.4.7}{=} \sum_{i=0}^{n}\alpha_i\iota_3(a_i,b_i)\mathrm{col}(z) = \left(\sum_{i=0}^{n}\alpha_i\iota_3(a_i,b_i)\right)\mathrm{col}(z) \\ &= A(z)\mathrm{col}(z).\end{aligned}$$

□

Lema 3.4.10. *Sea $z_0 \in \mathbb{H}$. Entonces la matriz $A(z)$ es constante $\forall z \in [z_0]$.*

Demostración. Notemos que, fijado un polinomio bilátero p de grado n, $A(z)$ sólo depende de $\{\alpha_i\}_{i=0}^n$. Pero estos valores sabemos por la proposición 3.3.8, a su vez, sólo dependían de $\mathcal{R}(z)$ y de $\|z\|$. Por tanto, fijado $z_0 \in \mathbb{H}\backslash\mathbb{R}$, dichos valores van a ser constantes $\forall z \in [z_0]$. □

Teorema 3.4.11 (Janovská-Opfer, [JO10b, p.88])**.** *Sean p un polinomio bilátero y $z \in \mathbb{H}$ tales que:*

$$P(z) := col(p(z)) = A(z)col(z) + col(B(z)) = col(0) =: \vec{0}.$$

Entonces la ecuación sigue siendo válida si sustituimos $A(z)$ y $B(z)$ con $A(z_0)$ y $B(z_0)$, donde z_0 es el representante complejo de $[z]$.

Demostración. Situémonos en las hipótesis del teorema. Por la proposición 3.3.5 $z_0 \in [z]$ y por lo tanto, usando los lemas 3.4.3 y 3.4.10, $B(z_0) = B(z)$ y $A(z_0) = A(z)$. □

Nota: Notemos que, por definición, dado $z \in \mathbb{H}$:

$$p(z) = 0 \iff P(z) = \vec{0}.$$

Pues $P(z) = \text{col}(p(z))$ y $\text{col}(z) = \vec{0} \iff z = 0$.

Corolario 3.4.12. *Para encontrar las raíces no reales, z, de un polinomio bilátero p, basta con encontrar los representantes complejos, $z_0 \in [z]$. Puede que z_0 no sea raíz de $p(z)$.*

Demostración. Situémonos en las hipótesis del corolario. Si averiguamos los posibles representantes complejos z_0 de las posibles clases de equivalencia, $[z]$, donde exista solución, por el teorema 3.4.11, basta con resolver para $\text{col}(z)$ el sistema lineal $A(z_0)z + \text{col}(B(z_0)) = \vec{0}$. Obteniendo así todas las raíces que haya en $[z]$. □

Teorema 3.4.13. *Sean $z_1, z_2 \in \mathbb{H}$ dos raíces distintas de un polinomio bilátero p tales que $z_1 \in [z_2]$. Entonces la matriz $A := A(z_1) = A(z_2)$ es singular.*

Demostración. Sean z_1, z_2 y p en las condiciones del teorema. Sabemos por el lema 3.4.10 que $A := A(z_1) = A(z_2)$ y por el lema 3.4.3 que $B(z_1) = B(z_2)$. Por tanto, tenemos que:

$$P(z_1) - P(z_2) = A\text{col}(z_1 - z_2) = \vec{0}.$$

Como $z_1 \neq z_2$, tenemos que $v = (v_1, v_2, v_3, v_4)^T = \text{col}(z_1 - z_2) \neq \vec{0}$. Por tanto, si denotamos por $a_i \in \mathbb{R}^4$ la i-ésima columna de A, obtenemos que:

$$\sum_{i=0}^{4} v_i a_i = \vec{0}, \quad \exists i \in \{1, \ldots, 4\},\ v_i \neq 0.$$

Por tanto, las columnas de A son linealmente dependientes y A es una matriz singular. □

En este caso hemos obtenido un resultado que generaliza al teorema 2.1.6 a los polinomios $p \in \mathbb{H}\langle z\rangle$ que cumplen la propiedad de ser biláteros. En efecto, si tenemos que en una misma clase de conjugación tenemos al menos dos raíces distintas, el teorema anterior afirma que la matriz A es singular. Por tanto, el sistema lineal establecido en 3.4.11 no puede ser compatible determinado, pero, por hipótesis, es un sistema compatible. Luego será un sistema compatible indeterminado que tendrá infinitas soluciones. Además, estas soluciones formarán un subespacio vectorial de $\mathbb{R}^4$ de dimensión $k = 4 - \text{rango}(A)$. Esta última observación nos lleva a hacer la siguiente clasificación.

Definición 3.4.14. Sean $p \in \mathbb{H}\langle z\rangle$ un polinomio bilátero y $z \in \mathbb{H}$ una raíz de p.

Si $z \notin \mathbb{R}$, sea $z_0 \in \mathbb{C}$ el representante complejo de $[z]$. Diremos que z es una raíz de tipo k, si $\text{rango}(A(z_0)) = 4 - k$, con $0 \leq k \leq 4$. Las raíces de tipo 4, cumplen que $\text{rango}(A(z_0)) = 0$, por tanto $A(z_0) = (0)_{i,j=0}^4$ y $\forall z \in [z_0]$, $p(z_0) = 0$. Las raíces de tipo 0, cumplen que $\text{rango}(A(z_0)) = 4$, por tanto, A es invertible y el sistema $P(z) = \vec{0}$ tendrá como solución única a $\text{col}(z) = -A^{-1}(z_0)B(z_0)$, que a su vez define un único $z \in \mathbb{H}$ tal que $p(z) = 0$. Debido a las similitudes con el la clasificación realizada en la sección anterior, llamaremos a las primeras raíces esféricas y, a las segundas, raíces aisladas.

Si $z \in \mathbb{R}$, como $Z(\mathbb{H}) = \mathbb{R}$, tenemos que podemos reescribir $p(z) = \sum_{i=0}^n a_i z^i b_i$, como $p(z) = \sum_{i=0}^n a_i b_i z^i$. Pero este caso está estudiado ya en la sección anterior, luego se puede resolver mediante el método de Janovská-Opfer y obtener todas las raíces reales de p. Además, como $\forall z \in \mathbb{R}$, $[z] = \{z\}$, también diremos que las raíces reales de los polinomios biláteros son raíces aisladas.

Dado un polinomio bilátero $p \in \mathbb{H}\langle z\rangle$, encontrar sus raíces es, en general, un problema complicado. El principal problema de este método es que para conocer el representante complejo $z_0 \in \mathbb{C}$ de una clase de conjugación $[z]$ que tenga una raíz de p, tenemos que conocer dicha clase a priori, pues no podemos calcularlos por separado resolviendo otra ecuación, a diferencia de lo que ocurría con el polinomio acompañante en la sección anterior.

Podríamos pensar que clasificarlas, al menos, será un problema más sencillo. Sabemos por la sección anterior, pues los polinomios simples son aquellos polinomios biláteros tales que para todo i o bien $a_i = 1$ o bien $b_i = 1$, que las raíces aisladas y las raíces esféricas existen; pero no sabemos nada sobre las raíces de tipo k donde $k = 1, 2$ o 3.

Podríamos pensar en estudiar la matriz $A(z)$ para un polinomio genérico y obtener así información sobre sus posibles rangos en función de z. Sin embargo, la recurrencia con la que están definidos $\forall i \geq 0, \alpha_i, \beta_i \in \mathbb{R}$ provoca que dicha matriz sea excesivamente complicada de trabajar en el caso general. Tomemos como ejemplo un polinomio bilátero de grado 3 genérico:

$$p(z) = z^3 + a_2 z^2 b_2 + a_1 z b_1 + c, \quad \forall i,\ a_i, b_i, c \in \mathbb{H}$$

Aplicando la proposición 3.3.8 obtenemos los siguientes valores:

Lo cual nos lleva a que:

$$A(z) = \iota_3(a_1, b_1) + 2\mathcal{R}(z)\iota_3(a_2, b_2) + \left(4\mathcal{R}(z)^2 - \|z\|^2\right) I_4.$$

$$\begin{aligned}
&\alpha_0 = 0, && \beta_0 = 1,\\
&\alpha_1 = 1, && \beta_1 = 0,\\
&\alpha_2 = 2\mathcal{R}(z), && \beta_2 = -\|z\|^2,\\
&\alpha_3 = 4\mathcal{R}(z)^2 - \|z\|^2, && \beta_3 = -2\|z\|^2\mathcal{R}(z).
\end{aligned}$$

De donde no podemos sacar ninguna información acerca del rango de $A(z)$. Janovská y Opfer, en [JO10b], afirman haber encontrado ejemplos de polinomios biláteros de grado 4 donde aparecen raíces de todos los tipos, por tanto la clasificación tiene sentido en, al menos, los polinomios biláteros de grado ≥ 4 y, además, demuestran que en el caso cuadrático solo pueden existir raíces de tipo par, demostración que explicaremos a continuación.

3.4.1. Polinomios biláteros cuadráticos.

Sea $p(z) = z^2 + azb + c$ un polinomio bilátero de grado dos genérico. Usando los valores $\{\alpha_i, \beta_i\}_{i=0}^2$ obtenidos anteriormente, tenemos que:

$$A(z) = \iota_3(a,b) + 2\mathcal{R}(z)I_4. \tag{b}$$

Proposición 3.4.15. *La matriz $A := A(z)$ definida en (b) es diagonalizable.*

Demostración. Por el lema 3.4.8 sabemos que $\iota_3(a,b)$ es diagonalizable. Es decir, existen matrices $P, D \in \mathcal{M}_4(\mathbb{C})$, con P invertible y D una matriz diagonal, tales que:

$$\iota_3(a,b) = PDP^{-1}.$$

Sustituyendo $\iota_3(a,b)$ en (b) y teniendo en cuenta que $\forall M \in \mathcal{M}_4(\mathbb{C})$ y $\forall \lambda \in \mathbb{R}$, tenemos que $I_4 M = M I_4$ y $\lambda M = M\lambda$, obtenemos:

$$\begin{aligned}
A &= \iota_3(a,b) + 2\mathcal{R}(z)I_4 = PDP^{-1} + 2\mathcal{R}(z)I_4 PP^{-1} = PDP^{-1} + P\left(2\mathcal{R}(z)I_4\right)P^{-1}\\
&= P\left(D + 2\mathcal{R}(z)I_4\right)P^{-1} = P\tilde{D}P^{-1}.
\end{aligned}$$

Tanto D como $2\mathcal{R}(z)I_4$ son matrices diagonales, luego $\tilde{D}$ también lo será y por tanto A es diagonalizable. □

Lema 3.4.16. *El rango de la matriz $A := A(z)$ definida en (b) es par.*

Demostración. Por la proposición 3.4.15 sabemos que A es diagonalizable y, usando la notación de la demostración de dicha proposición, sus autovalores serán las entradas de la diagonal principal de $\tilde{D}$.

Denotemos por $\Lambda := \{\lambda_i\}_{i=0}^4$ y por $M := \{\mu_i\}_{i=0}^4$ los autovalores de A e $\iota_3(a,b)$, respectivamente. Por $\tilde{D} = D + 2\mathcal{R}(z)I_4$ y por el lema 3.4.8 sabemos lo siguiente:

- $\forall i \in \{1,\ldots,4\}$, $\lambda_i = \mu_i + 2\mathcal{R}(z)$.
- $\forall i \in \{1,\ldots,4\}$, $\|\mu_i\| = \|ab\|$ y dichos autovalores están emparejados dos a dos con su conjugado.

Todo ello nos lleva a estudiar tres casos:

(a) En M hay dos parejas de autovalores no reales conjugados. En este caso, $\forall i, \lambda_i \neq 0$, por tanto rango$(A) = 4$.

(b) En M hay una pareja de autovalores no reales y una pareja de autovalores reales iguales. En este caso, los autovalores de A correspondientes a la primera pareja nunca se anularán y los correspondientes a la segunda pareja se anularán siempre a la vez. Por tanto o bien rango$(A) = 4$ o bien rango$(A) = 2$, dependiendo del valor de z.

(c) En M hay dos parejas de autovalores reales iguales. Además, como $\forall i$, $\|\mu_i\| = \|ab\|$, los valores de las dos parejas como mucho se diferencian en un signo. En este caso es inmediato ver que, si los cuatro autovalores son iguales, entonces rango$(A) = 4$ ó 0; mientras que, si ambas parejas tienen valores opuestos, entonces rango$(A) = 4$ ó 2.

En cualquiera de los tres casos, tenemos que rango$(A) = 0, 2$ ó 4. □

Del lema anterior se deduce inmediatamente el siguiente teorema:

Teorema 3.4.17 (Janovská-Opfer, [JO10b, p.90])**.** *Sea p un polinomio bilátero cuadrático tal y como se ha definido anteriormente. Se pueden dar los siguientes tres casos en función de los autovalores de $\iota_3(a, b)$:*

1. *Todos los autovalores son no reales. Entonces solo habrá raíces aisladas.*

2. *Hay autovalores reales y no reales. Entonces puede haber raíces aisladas o raíces de tipo 2.*

3. *Todos los autovalores son reales. Entonces puede haber raíces aisladas, esféricas o de tipo 2.*

Conclusiones y trabajo futuro.

En este trabajo hemos buscado dar una recopilación de los principales resultados en la literatura acerca de la teoría de polinomios sobre anillos de división en general y, más concretamente de polinomios sobre cuaternios, de una forma didáctica, dando las bases necesarias para entender las demostraciones, así como profundizar y desarrollar las mismas con resultados que en los artículos originales se dan por sabidos. También hemos visto cómo, al no poder asegurar la conmutatividad, los resultados son mucho más complejos y sorprendentes que a los que estamos acostumbrados cuando los coeficientes están en un cuerpo.

Usando este trabajo como base, se puede seguir el estudio explorando más en detalle la demostración de carácter topológico del teorema 3.4.1 y estudiar la complejidad computacional de la implementación del método de Janovská-Opfer, así como las diversas consideraciones numéricas necesarias para la resolución de las ecuaciones del tipo $P(z) = \vec{0}$ vistas en la sección anterior. Por ejemplo, Janovská y Opfer usan el método de Newton en una sección de [JO10b].

Pero, sin duda alguna, la continuación más natural de este trabajo, tanto por el desequilibrio de resultados sobre polinomios simples y polinomios generales que hay en el mismo como por seguir trabajando sobre estructuras algebraicas, es un estudio sobre la clausura algebraica de los anillos de división en los polinomios generales. Por el corolario 2.2.18 sabemos que un anillo de división algebraicamente cerrado debe ser centralmente infinito, pero, al menos con la investigación hecha para este trabajo, no se sabe mucho más; ni siquiera si todo anillo de división se puede incluir en un anillo de división algebraicamente cerrado. Sin embargo, en [Mak85], Makar-Limanov construyó el único ejemplo conocido de un anillo de división algebraicamente cerrado, así que al menos se conoce su existencia.

Apéndice A

Los cuerpos reales cerrados son elementalmente equivalentes a los reales.

Empecemos viendo las definiciones de teoría de modelos de primer orden necesarias.

Definición A.0.1. Un *lenguaje* L es un conjunto de símbolos de constantes, símbolos de función y símbolos de predicado. Los símbolos de constantes tienen aridad 0, mientras que los de función y predicado tienen aridad ≥ 1.

Los lenguajes, como tal, no tienen significado, para ello hay que interpretarlos dentro de una estructura.

Definición A.0.2. Sea L un lenguaje. Una L-estructura es un par $\mathfrak{U} = \left(A, (Z^{\mathfrak{U}})_{Z\in L}\right)$, donde:

- A es un conjunto no vacío denominado *universo de* $\mathfrak{U}$.
- Si $Z \in L$ es un símbolo de constante, entonces $Z^{\mathfrak{U}} \in A$.
- Si $Z \in L$ es un símbolo de función n-ario, entonces $Z^{\mathfrak{U}}$ es una aplicación de A^n en A, es decir, $Z^{\mathfrak{U}} : A^n \to A$.
- Si $Z \in L$ es un símbolo de predicado n-ario, entonces $Z^{\mathfrak{U}} \subset A^n$.

Diremos que $Z^{\mathfrak{U}}$ es la interpretación de Z en $\mathfrak{U}$.

Definición A.0.3. Sean $\mathfrak{U} = \left(A, (Z^{\mathfrak{U}})_{Z\in L}\right)$ y $\mathfrak{V} = \left(B, (Z^{\mathfrak{V}})_{Z\in L}\right)$ dos L-estructuras. Diremos que $\mathfrak{U}$ *es una subestructura de* $\mathfrak{V}$ si se cumple que:

- $A \subset B$.
- Para cada símbolo de función n-ario, $f \in L$, se cumple que $f^{\mathfrak{U}} = f^{\mathfrak{V}}|_{A^n}$.
- Para cada símbolo de predicado n-ario, $p \in L$, se cumple que $f^{\mathfrak{U}} = f^{\mathfrak{V}} \cap A^n$.

Este hecho lo denotaremos por $\mathfrak{U} \subset \mathfrak{V}$.

Definición A.0.4. Usando la notación anterior, definimos las L-fórmulas de primer orden como aquellas palabras construidas a partir de los símbolos de función y de constante de L y a partir de unas variables $v_1, v_2, \ldots$ conforme a las siguientes normas:

- Las variables y las constantes son L-fórmulas.
- Dados un símbolo de función $n-$ario, f, y n L-fórmulas, $t_1, \ldots, t_n$, entonces $f(t_1, \ldots, t_n)$ es una L-fórmula.

Definición A.0.5. Sea $\mathfrak{U}$ una L-estructura de universo A. Sean una L-fórmula t y $\vec{a} = (\tilde{a_i})_{i \geq 1}$, con $\tilde{a_i} \in L$ símbolos de constante y $\tilde{a_i}^{\mathfrak{U}} = a_i \in A$. Diremos que $\vec{a}$ es una asignación y llamaremos *interpretación de t en $\vec{a}$* al resultado de sustituir en t cada aparición de la variable v_i por $\tilde{a_i}$, que denotaremos por $t^{\mathfrak{U}}[\vec{a}]$.

Definición A.0.6. Un L-enunciado es un L-fórmula donde todas las variables aparecen de forma ligada mediante los cuantificadores $\exists$ y $\forall$.

Definición A.0.7. Sean $\sigma, t_1, t_2, \ldots$ L-fórmulas y x, y variables. Diremos que σ es cierto en $\mathfrak{U}$ o que $\mathfrak{U}$ modela σ, y lo denotamos por $\mathfrak{U} \models \sigma$, si sucede lo siguiente de forma recursiva:

- Si $\sigma = Pt_1, \ldots, t_n$, con P un símbolo de predicado n-ario, entonces $\mathfrak{U} \models \sigma \iff (t_1^{\mathfrak{U}}, \ldots, t_n^{\mathfrak{U}}) \in A^n$.
- Si $\sigma = (t_1 = t_2)$, entonces $\mathfrak{U} \models \sigma \iff t_1^{\mathfrak{U}} = t_2^{\mathfrak{U}}$.
- Si $\sigma = t_1 \vee t_2$, entonces $\mathfrak{U} \models \sigma \iff \mathfrak{U} \models t_1$ o $\mathfrak{U} \models t_2$.
- Si $\sigma = t_1 \wedge t_2$, entonces $\mathfrak{U} \models \sigma \iff \mathfrak{U} \models t_1$ y $\mathfrak{U} \models t_2$.
- Si $\sigma = \neg t_1$, entonces $\mathfrak{U} \models \sigma \iff \mathfrak{U} \not\models t_1$.
- Si $\sigma = \forall x\, t_1$, entonces $\mathfrak{U} \models \sigma \iff \mathfrak{U} \models t_1^{\mathfrak{U}}[\vec{a}]$ para toda asignación $\vec{a}$.
- Si $\sigma = \exists x\, t_1$, entonces $\mathfrak{U} \models \sigma \iff$ existe al menos una asignación $\vec{a}$ tal que $\mathfrak{U} \models t_1^{\mathfrak{U}}[\vec{a}]$.

Definición A.0.8. Fijado un lenguaje L, a los conjuntos de L-fórmulas los llamaremos *L-teorías.*

Sean Σ una L-teoría y $\mathfrak{U}$ una L-estructura. Si $\forall \sigma \in \Sigma$, $\mathfrak{U} \models \sigma$ diremos que $\mathfrak{U}$ *es un modelo de* Σ.

Diremos que una L-teoría Σ es *consistente* si existe un modelo $\mathfrak{U}$ de Σ.

Sean φ una L-fórmula y Σ una L-teoría. Si $\forall \mathfrak{U}$ modelo de Σ se cumple que $\mathfrak{U} \models \varphi$, entonces diremos que Σ prueba φ y lo denotaremos por $\Sigma \vdash \varphi$.

Diremos que una L-teoría consistente Σ es *completa* si toda L-fórmula, φ, cumple:

$$\Sigma \vdash \varphi \text{ o } \sigma \vdash \neg\varphi.$$

Definición A.0.9. Dos L-estructuras, $\mathfrak{U}, \mathfrak{V}$, se dicen *elementalmente equivalentes* si, para toda L-fórmula de primer orden φ:

$$\mathfrak{U} \models \varphi \iff \mathfrak{V} \models \varphi.$$

Definición A.0.10. Diremos que una L-teoría Σ admite eliminación de cuantificadores si para toda L-fórmula con variables, $\varphi(x_1, \ldots, x_n)$, existe una L-fórmula sin cuantificadores, $\psi(x_1, \ldots, x_n)$, tal que:

$$\Sigma \models (\forall x_1 \forall x_2 \cdots \forall x_n \, (\varphi(x_1, \ldots, x_n) \iff \psi(x_1, \ldots, x_n))) \, .$$

Definición A.0.11. Sean una L-teoría Σ y una L-estructura $\mathfrak{U}$. Diremos que $\mathfrak{U}$ es una estructura prima de Σ si para todo modelo $\mathfrak{V}$ de Σ tenemos que $\mathfrak{U} \subset \mathfrak{V}$.

Con todas las definiciones anteriores ya podemos dar los resultados necesarios.

Lema A.0.12 (1.3.9 en [TZ12])**.** *Si la teoría Σ es completa entonces todos los modelos de Σ son elementalmente equivalentes los unos a los otros.*

Lema A.0.13 (3.2.2 en [TZ12])**.** *Toda teoría con eliminación de cuantificadores que contiene una estructura prima es completa.*

Teorema A.0.14 (Tarski-Seidenberg, 3.3.15 en [TZ12])**.** *La teoría de los cuerpos reales cerrados tiene eliminación de cuantificadores y tiene como estructura prima al cuerpo ordenado de los racionales.*

Juntando estos tres resultados obtenemos que todos los cuerpos reales cerrados son elementalmente equivalentes y, por tanto, elementalmente equivalentes a $\mathbb{R}$.

Bibliografía

[Bre14] Matej Brešar. *Introduction to noncommutative algebra*. Universitext. Springer, Cham, 2014, págs. xxxviii+199. ISBN: 978-3-319-08692-7; 978-3-319-08693-4. DOI: 10.1007/978-3-319-08693-4.

[Coh95] P. M. Cohn. *Skew fields*. Vol. 57. Encyclopedia of Mathematics and its Applications. Theory of general division rings. Cambridge University Press, Cambridge, 1995, págs. xvi+500. ISBN: 0-521-43217-0. DOI: 10.1017/CBO9781139087193.

[EN44] Samuel Eilenberg e Ivan Niven. «The "fundamental theorem of algebra" for quaternions». En: *Bull. Amer. Math. Soc.* 50 (1944), págs. 246-248. ISSN: 0002-9904. DOI: 10.1090/S0002-9904-1944-08125-1.

[GM65] B. Gordon y T. S. Motzkin. «On the zeros of polynomials over division rings». En: *Trans. Amer. Math. Soc.* 116 (1965), págs. 218-226. ISSN: 0002-9947. DOI: 10.2307/1994114.

[Ham44] William Rowan Hamilton. «II. On quaternions; or on a new system of imaginaries in algebra». En: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 25.163 (1844), págs. 10-13. DOI: 10.1080/14786444408644923.

[Jac64] Nathan Jacobson. *Structure of rings*. Revised. American Mathematical Society Colloquium Publications, Vol. 37. American Mathematical Society, Providence, RI, 1964, págs. ix+299.

[JO10a] Drahoslava Janovská y Gerhard Opfer. «A note on the computation of all zeros of simple quaternionic polynomials». En: *SIAM J. Numer. Anal.* 48.1 (2010), págs. 244-256. ISSN: 0036-1429. DOI: 10.1137/090748871.

[JO10b] Drahoslava Janovská y Gerhard Opfer. «The classification and the computation of the zeros of quaternionic, two-sided polynomials». En: *Numer. Math.* 115.1 (2010), págs. 81-100. ISSN: 0029-599X. DOI: 10.1007/s00211-009-0274-y.

[Lam01] T. Y. Lam. *A first course in noncommutative rings*. Second. Vol. 131. Graduate Texts in Mathematics. Springer-Verlag, New York, 2001, págs. xx+385. ISBN: 0-387-95183-0. DOI: 10.1007/978-1-4419-8616-0.

[Lan02] Serge Lang. *Algebra*. third. Vol. 211. Graduate Texts in Mathematics. Springer-Verlag, New York, 2002, págs. xvi+914. ISBN: 0-387-95385-X. DOI: 10.1007/978-1-4613-0041-0.

[Mak85] L. Makar-Limanov. «Algebraically closed skew fields». En: *J. Algebra* 93.1 (1985), págs. 117-135. ISSN: 0021-8693. DOI: 10.1016/0021-8693(85)90177-2.

[Niv41] Ivan Niven. «Equations in quaternions». En: *Amer. Math. Monthly* 48 (1941), págs. 654-661. ISSN: 0002-9890. DOI: 10.2307/2303304.

[TZ12] Katrin Tent y Martin Ziegler. *A course in model theory*. Vol. 40. Lecture Notes in Logic. Association for Symbolic Logic, La Jolla, CA; Cambridge University Press, Cambridge, 2012, págs. x+248. ISBN: 978-0-521-76324-0. DOI: 10.1017/CBO9781139015417.

[Wei09] Steven H. Weintraub. *Galois theory*. Second. Universitext. Springer, New York, 2009, págs. xiv+211. ISBN: 978-0-387-87574-3. DOI: 10.1007/978-0-387-87575-0.