Ecuaciones Algebraicas. Curso 2002/2003.

El Teorema Fundamental del Álgebra.

Profesor José Gómez Torrecillas

Esta lección está dedicada a proporcionar una demostración algebraica, debida fundamentalmente a Gauss, del Teorema Fundamental del Álgebra. En términos modernos, el único argumento no puramente algebraico que usaremos será el hecho de que una función polinómica de grado impar tiene al menos un cero real (lo cual se deduce del Teorema de Bolzano). Esto, en tiempos de Gauss, era «auto-evidente».

1. Polinomios simétricos, brevemente Sea F un cuerpo y $t_1, \ldots, t_m$ indeterminadas. A cada polinomio $f(t_1, \ldots, t_m) \in F[t_1, \ldots, t_m]$ y cada permutación $\sigma \in S_m$, le podemos asignar el polinomio $\sigma \cdot f(t_1, \ldots, t_m) = f(t_{\sigma(1)}, \ldots, t_{\sigma(m)})$. El polinomio f se dice *simétrico* si $\sigma \cdot f = f$ para todo $\sigma \in S_m$. Los polinomios simétricos aparecen de manera natural en la teoría de ecuaciones algebraicas, ya que si tomo el polinomio en X

$$f(X) = (X - t_1)(X - t_2) \cdots (X - t_m) = X^m + a_{m-1}X^{m-1} + \cdots a_1 X + a_0,$$

entonces tenemos las relaciones

$$\begin{aligned}
a_0 &= (-1)^m t_1 t_2 \cdots t_m \\
a_1 &= (-1)^{m-1}(t_1 t_2 + t_1 t_3 + \cdots + t_{m-1} t_m) \\
&\vdots \quad\quad \vdots \\
a_{n-r} &= (-1)^r \sum_{i_1 < \cdots < i_r} t_{i_1} \cdots t_{i_r} \\
&\vdots \quad\quad \vdots \\
a_{n-1} &= -\sum_{i=1}^{m} t_i
\end{aligned}$$

Al polinomio $s_r = \sum_{i_1 < \cdots < i_r} t_{i_1} \cdots t_{i_r}$ se le llama *r–ésimo polinomio simétrico elemental en* $t_1, \ldots, t_m$. De manera que $a_{m-r} = (-1)^r s_r$. La forma de expresar estos polinomios simétricos hace uso del orden natural $1 < 2 < \cdots < m$. Para nuestros propósitos, el conjunto de índices de las variables no siempre va a ser el de los m primeros números naturales, por lo que vamos a dar una expresión alternativa de los polinomios simétricos elementales. Así, si I es un conjunto finito con m elementos, sin presuponer que son números naturales, tomamos una indeterminada t_i para cada $i \in I$. Entonces, el r–ésimo polinomio simétrico elemental en las variables t_i con $i \in I$ viene dado por

$$s_r = \sum_{\substack{J \subseteq I \\ |J| = r}} \prod_{i \in I} t_i,$$

donde $\prod_{i \in J} t_i$ representa el monomio $t_{i_1} \cdots t_{i_r}$ si $J = \{i_1, \ldots, i_r\}$ (y, a fin de cuentas, hemos elegido el orden que nos ha parecido oportuno).

Lema 1. *Sean $x_1, \ldots, x_n$ indeterminadas sobre un cuerpo F y $c \in F$. Para cada $i, j = 1, \ldots, n$ con $i \neq j$, definimos $y_{\{i,j\}}(x_1, \ldots, x_n) = x_i + x_j + c x_i x_j$. Si $\mathcal{S}_r$ es el r–ésimo polinomio simétrico elemental para los $y_{\{i,j\}}$ donde $\{i, j\}$ recorre el conjunto de los subconjuntos de $\{1, \ldots, n\}$ con exactamente dos elementos, entonces $\mathcal{S}_r$ es una función simétrica de $x_1, \ldots, x_n$.*

Demostración. Sea $I = \{\{i,j\}\,;\, i,j = 1,\ldots,n\,;\, i \neq j\}$ el conjunto de todos los subconjuntos de $\{1,\ldots,n\}$ con exactamente dos elementos. Este conjunto tiene obviamente $n(n-1)/2$ elementos. Cada permutación $\sigma \in S_n$ de n símbolos da una permutación $\widehat{\sigma} : I \to I$ definida por $\widehat{\sigma}(\{i,j\}) = \{\sigma(i),\sigma(j)\}$ para todo $\{i,j\} \in I\}$. Dado un número real c, definimos la función $y_{\{i,j\}}(x_1,\ldots,x_n) = x_i + x_j + cx_ix_j$. Si hacemos actuar $\sigma \in S_n$ sobre esta función obtenemos

$$\begin{aligned}
\sigma \cdot y_{\{i,j\}}(x_1,\ldots,x_n) &= y_{\{i,j\}}(x_{\sigma(1)},\ldots,x_{\sigma(n)}) && (1)\\
&= x_{\sigma(i)}x_{\sigma(j)} + cx_{\sigma(i)}x_{\sigma(j)} && (2)\\
&= y_{\{\sigma(i),\sigma(j)\}})(x_1,\ldots,x_n) && (3)\\
&= (\widehat{\sigma}\cdot y_{\{i,j\}})(x_1,\ldots,x_n) && (4)
\end{aligned}$$

Si tomamos ahora el r–ésimo polinomio simétrico elemental $\mathcal{S}_r$ en $y_{\{1,2\}}, y_{\{1,3\}},\ldots,y_{\{n-1,n\}}$, tenemos

$$\mathcal{S}_r = \sum_{\substack{J\subseteq I\\ |J|=r}} \prod_{\{i,j\}\in J} y_{\{i,j\}}$$

Viendo $\mathcal{S}_r$ como función de $x_1,\ldots,x_n$, podemos hacer actuar $\sigma \in S_n$, y obtenemos

$$\begin{aligned}
\sigma\cdot\mathcal{S}_r(x_1,\ldots,x_n) &= \sigma\cdot\sum_{\substack{J\subseteq I\\ |J|=r}} \prod_{\{i,j\}\in J} y_{\{i,j\}}(x_1,\ldots x_n) && (5)\\
&= \sum_{\substack{J\subseteq I\\ |J|=r}} \prod_{\{i,j\}\in J} \sigma\cdot y_{\{i,j\}}(x_1,\ldots x_n) && (6)\\
&= \sum_{\substack{J\subseteq I\\ |J|=r}} \prod_{\{i,j\}\in J} (\widehat{\sigma}\cdot y_{\{i,j\}})(x_1,\ldots x_n) && (7)\\
&= (\widehat{\sigma}\cdot \sum_{\substack{J\subseteq I\\ |J|=r}} \prod_{\{i,j\}\in J} y_{\{i,j\}})(x_1,\ldots x_n) && (8)\\
&= (\widehat{\sigma}\cdot\mathcal{S}_r)(x_1,\ldots,x_n) && (9)\\
&= \mathcal{S}_r(x_1,\ldots,x_n), && (10)
\end{aligned}$$

ya que, a fin de cuentas, $\mathcal{S}_r$ es una función simétrica de los $y_{\{i,j\}}$. Por tanto, la función $\mathcal{S}_r$ es simétrica en $x_1,\dot{,}sx_n$. □

Teorema 1. *Todo polinomio* $f(X) \in \mathbb{R}[X]$ *no constante tiene una raíz compleja.*

Demostración. Desde luego, podemos suponer que $f(X)$ es mónico y que $\mathtt{grado}(f(X)) = 2^e m$, para $e \geq 0$ y m impar. La demostración procede por inducción sobre el exponente e. Si $e = 0$, entonces $f(X)$ es de grado impar y, por el Teorema de Bolzano, tiene al menos una raíz real (y, por ende, compleja). Para $e > 0$, tomamos un cuerpo de descomposición K de $f(X)$ sobre $\mathbb{C}$. De esta manera, existen $x_1,\ldots,x_n \in K$ tales que $f(X) = (X - x_1)\cdots(X - x_n)$, para $n = 2^e m$. Se trata de demostrar que alguna de las raíces x_i es un número complejo. Sea I el conjunto de los subconjuntos de $\{1,\ldots,n\}$ con exactamente dos elementos. Para cada $c \in \mathbb{R}$ y cada $\{i,j\} \in I$, definimos $y_{\{i,j\}}(c) = x_i + x_j + cx_ix_j$ y tomemos $F_c(Y) = \prod_{\{i,j\}\in I}(Y - y_{\{i,j\}}(c)) \in K[Y]$. Este polinomio tiene grado $n(n-1)/2$. Por las relaciones de Cardano-Vieta, el coeficiente de Y^r en K de $F_c(Y)$ son de la forma $(-1)^r\mathcal{S}_r(x_1,\ldots,x_n)$, donde $\mathcal{S}_r$ es el r–ésimo polinomio elemental en los $y_{\{i,j\}}$ con $\{i,j\} \in I$. Por el Lema 1, dichos coeficientes son funciones polinómicas simétricas con coeficientes en $\mathbb{R}$ de $x_1,\ldots,x_n$. El Teorema Fundamental de los polinomios simétricos nos dice ahora que cada coeficiente de $F_c(Y)$ es un polinomio con coeficientes en $\mathbb{R}$ de las funciones simétricas elementales en $x_1,\ldots,x_n$. Como éstas son las raíces de un polinomio $f(X) \in \mathbb{R}[X]$, deducimos que $F_c(Y) \in \mathbb{R}[Y]$. Ahora bien, el grado de $F_c(Y)$ es $n(n-1)/2 = 2^e m(2^e m - 1)/2 = 2^{e-1}m(2^e m - 1) = 2^{e-1}m'$, con $m' = m(2^e m - 1)$ impar. Por

hipótesis de inducción, alguna de las raíces $y_{\{i_c,j_c\}}$ es compleja. Ahora, como el conjunto $\{\{i_c,j_c\}\ ;\ c \in \mathbb{R}\}$ es finito, existen seguro $c, d \in \mathbb{R}$ distintos tales que $\{i_c,j_c\} = \{i_d,j_d\}$ y, por tanto, $y_{\{i_c,j_c\}} = y_{\{i_d,j_d\}}$. Su diferencia es un número complejo que calculamos seguidamente, tomando r, s tales que $\{i_c,j_c\} = \{i_d,j_d\} = \{r,s\}$:

$$y_{\{i_c,j_c\}} - y_{\{i_d,j_d\}} = (c-d)x_r x_s.$$

De aquí se sigue que $x_r x_s \in \mathbb{C}$. Como $x_r + x_s + cx_r x_s \in \mathbb{C}$, obtenemos que $x_r + x_s \in \mathbb{C}$, también. En conclusión, el polinomio $X^2 - (x_r + x_s)X + x_r x_s$ tiene coeficientes en $\mathbb{C}$ y sus raíces, que son x_r y x_s, son números complejos. □

Teorema 2 (Teorema Fundamental del Álgebra). *Todo polinomio no constante* $f(X) \in \mathbb{C}[X]$ *tiene alguna raíz en* $\mathbb{C}$. *Como consecuencia,* $f(X) = a(X-z_1)\cdots(X-z_n)$, *para* $a, z_1, \ldots, z_n \in \mathbb{C}$, *donde* n *es el grado de* $f(X)$.

Demostración. Si $f(X) = a_0 + a_1X + \cdots + a_nX^n$, tomamos $\widehat{f}(X) = \overline{a}_0 + \overline{a}_1X + \cdots + \overline{a}_nX^n$ y tenemos que $f(X)\overline{f}(X) \in \mathbb{R}[X]$. En realidad, lo que hemos hecho es extender el $\mathbb{R}$–automorfismo conjugación $\sigma : \mathbb{C} \to \mathbb{C}$ a un homomorfismo de anillos $\sigma : \mathbb{C}[X] \to \mathbb{C}[X]$, y denotar f^σ por $\overline{f}$. Pero entonces $\overline{f}\overline{\overline{f}} = \overline{f}\overline{\overline{f}} = \overline{f}f$. De aquí que $g = f\overline{f}$ tenga coeficientes reales. Por el Teorema 1, $g(X)$ tiene al menos una raíz compleja, digamos $z \in \mathbb{C}$. Pero entonces z es raíz o bien de $f(X)$, o bien de $\overline{f}(X)$. En el segundo caso, $\overline{z}$ ha de ser raíz de $f(X)$. Así, $f(X)$ tiene alguna raíz compleja, digamos z_1, con lo que $f(X) = (x - z_1)f_1(X)$, para cierto $f_1(X) \in \mathbb{C}[X]$. Un claro argumento inductivo sobre el grado de $f(X)$ permite su factorización completa. □

Corolario 1. *Si* $f(X) \in \mathbb{R}$ *es no constante, entonces* $f(X)$ *se descompone completamente como un producto de polinomios lineales y cuadráticos en* $\mathbb{R}[X]$.

Demostración. Por el Teorema Fundamental del Álgebra, $f(X) = a(X-z_1)\cdots(X-z_n)$, para $a \in \mathbb{R}$ y $z_1, \ldots, z_n \in \mathbb{C}$. Ahora, $f(X) = \overline{f}(X) = a(X-\overline{z}_1)\cdots(X-\overline{z}_n)$, con lo que para raíz z_i de $f(X)$ se tiene que $\overline{z}_i$ es asímismo raíz de $f(X)$. Los factores lineales corresponden por tanto a las raíces reales, y cada raíz compleja z_i de $f(X)$ la emparejamos con su raíz conjugada $\overline{z}_i$, obteniendo el factor cuadrático $(X-z_i)(X-\overline{z}_i) = X^2 - 2\Re z_i + |z_i|^2$, que es un polinomio cuadrático con coeficientes reales. □