Contents

0.1 Polynomial Roots. Elimination. Resultants 0
0.1.1 Polynomial roots 0
0.1.2 Elimination theory. Resultants 3
0.1.3 The abridged method of Bézout 4
0.1.4 Jacobi's version 5
0.1.5 Cauchy's contribution 6
0.1.6 The companion matrix 7

0.1 Polynomial Roots. Elimination. Resultants

> Despite its age, the subject of resultants is an interesting one. It can be used to illuminate several areas of matrix theory, and we relate to problems on the location of roots of polynomials; moreover, there are many relevant applications in the theory of linear control systems, including important extensions to polynomial matrices.
>
> (S. Barnett)

Throughout this section A is an integral domain.

0.1.1 Polynomial roots

Definition: An element $a \in A$ is called a *root* of the polynomial $P \in A[X]$ if $P(a) = 0$.

Proposition 0.1.1 (M. Bézout) *Let $P \in A[X]$ and $a \in A$. Then a is a root of P if and only if $X - a$ divides $P(X)$.*

Proof: By Euclidean division there exist $Q \in A[X]$ and $r \in A$ such that

$$P(X) = (X - a)Q(X) + r\,.$$

Therefore a is a root of P if and only if $r = 0$ or equivalently if and only if $X - a$ divides $P(X)$. □

Remark: This proposition is true for any ring A: the proof does not use any hypothesis on A.

Corollary 0.1.2 *If $a_1, \ldots, a_m \in A$ are distinct roots of $P \in A[X] \setminus A$, then $(X - a_1) \ldots (X - a_m)$ divides $P(X)$ in $A[X]$. Moreover, the number of roots of $P(X)$ in A is at most equal to* $\deg(P)$.

Proof: We use induction on m. For $m = 1$ the result follows from Proposition 0.1.1. Assume the statement is true for at most $m-1$ roots. We may write

$$P(X) = (X - a_1)\dots(X - a_{m-1})Q(X), \qquad Q \in A[X].$$

Then $P(a_m) = (a_m - a_1)\dots(a_m - a_{m-1})Q(a_m)$. Since A has not zero divisors it follows that $Q(a_m) = 0$. Again by Proposition 0.1.1 we have $Q(X) = (X - a_m)S(X)$, with $S \in A[X]$, which proves that $(X - a_1)\dots(X - a_{m-1})(X - a_m)$ divides $P(X)$ in $A[X]$.

On the other hand, from the relation

$$P(X) = (X - a_1)\dots(X - a_m)S(X)$$

it follows that $m \leq \deg(X - a_1)\dots(X - a_m)S(X) = \deg P$. □

Remark: If A has zero divisors, then there exist polynomials $P \in A[X]$ which have more than $\deg(P)$ distinct roots (cf. Exercise 1).

Definition: Let $P \in A[X]$. An element $a \in A$ is called a *root of order* $k \geq 1$ *of* P if $(X - a)^k$ divides P in $A[X]$ and $(X - a)^{k+1}$ does not divide P in $A[X]$. The integer k is called the *multiplicity* of the root a. Clearly, if $P \neq 0$ then $k \leq \deg(P)$ even if A is not an integral domain.

Proposition 0.1.3 *Let $P \in A[X]$ and $a \in A$. Then a is a root of order $k \geq 1$ if and only if there exists $Q \in A[X]$ such that*

$$P = (X - a)^k Q \quad \textit{and} \quad Q(a) \neq 0.$$

Proof: Suppose a is a root of order k. Then $P = (X - a)^k Q$, with $Q \in A[X]$. If a were a root of Q, then $(X - a)^{k+1}$ divides P, a contradiction. Therefore $Q(a) \neq 0$. The other implication is obvious. □

Proposition 0.1.4 *If $P \in A[X] \setminus \{0\}$, then the sum of the multiplicities of the roots of P that belong to A is at most equal to* $\deg(P)$.

Proof: Let $a_1, \dots, a_m$ be the roots of P in A and let $s_1, \dots, s_m$ be their multiplicities. Therefore there exists $Q \in A[X]$ such that

$$P = (X - a_1)^{s_1} \cdots (X - a_m)^{s_m} Q.$$

It follows that

$$s_1 + \cdots + s_m \leq \deg(P).$$

□

Corollary 0.1.5 *If there exist $a_1, \dots, a_s \in A$, with $s > \deg(P)$, such that $P(a_i) = 0$ for all $i = 1, \dots, s$, then $P = 0$.*

Definition: Let $P(X) = a_0 + a_1 X + \ldots + a_n X^n \in A[X]$. The polynomial

$$P'(X) = a_1 + 2a_2 X + \ldots + n a_n X^{n-1}$$

is called the *formal derivative* of the polynomial P.

Definition: For $k \in \mathbb{N}$ let $D_k : A[X] \longrightarrow A[X]$ be defined by the relations

$$\begin{aligned} D_k(X^n) &= \tbinom{n}{k} X^{n-k}, \\ D_k(\lambda P + \mu Q) &= \lambda D_k(P) + \mu D_k(Q), \end{aligned}$$

for any $n \in \mathbb{N}$, $\lambda, \mu \in A$, $P, Q \in A[X]$.

The operator D_k is called the *hyperderivative of order* k.

Remark: If the characteristic of A is zero[1], then $D_k(X^n) = \frac{1}{k!}\big(X^n\big)^{(k)}$.

Proposition 0.1.6 *If $a \in A$ is a root of order $k \geq 1$ of $P \in A[X]$, then $D_i(P)(a) = 0$ for $i = 0, 1, \ldots, k-1$ but $D_k(P)(a) \neq 0$.*

Proof: We first observe that

$$D_k\big((X-a)^j\big)(a) \;=\; \begin{cases} 0, & \text{if } j < k, \\ 1, & \text{if } j = k\ . \end{cases}$$

Then we notice that Leibniz formula remains true for hyperderivatives. Indeed, since the map $(F, G) \mapsto D_k(F \cdot G)$ is bilinear, we have just to verify it when F and G are both monomials, say X^m and Y^n respectively, and - in this case - it is a direct consequence of the trivial relation $(1+X)^{m+n} = (1+X)^m \cdot (1+X)^n$. Now, if we apply Leibniz formula to the product $P = (X-a)^k Q(X)$ we get

$$D_k(P)(a) \;=\; D_k\big((X-a)^k \cdot Q(X)\big)(a) \;=\; D_0(Q)(a) = Q(a) \neq 0\,,$$

hence the result. □

Corollary 0.1.7 *If $a \in A$ is a root of order $k \geq 1$ of the polynomial $P \in A[X]$, then $P^{(i)}(a) = 0$ for all $i \in \{0, 1, \ldots, k-1\}$. Moreover, if the the characteristic of A is zero, then $P^{(k)}(a) \neq 0$.*

If the integral domain A is a field of characteristic zero, then a refinement of Corollary 0.1.7 which uses the Taylor expansion is valid (Corollary 0.1.10).
With the operators D_k we obtain the following generalized Taylor expansion for polynomials:

Proposition 0.1.8 *Let A be a ring and let*

$$P(X) \;=\; a_n X^n + \cdots + a_1 X + a_0 \in A[X]\,.$$

Then

$$P(X+Y) \;=\; P(X) + Y D_1 P(X) + \cdots + Y^n D_n P(X)\,. \tag{1}$$

Proof: We simply use the binomial expansion

$$(X+Y)^s = \sum_{t=0}^{s} \binom{s}{t} X^{s-t} Y^t$$

and linearity. □

Corollary 0.1.9 *Let K be a field of characteristic zero and let*

$$P(X) \;=\; a_0 + a_1 X + \ldots + a_n X^n \in K[X]\,.$$

If Y is another variable, then

$$P(X+Y) = P(X) + \frac{Y}{1!}P^{(1)}(X) + \frac{Y^2}{2!}P^{(2)}(X) + \ldots + \frac{Y^n}{n!}P^{(n)}(X)\,. \tag{2}$$

Remark: $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are examples of fields of characteristic zero. Note that a finite field has not characteristic zero.

Definition: The relation (1) is called *Taylor's generalized formula* for univariate polynomials, while the relation (2) is called *Taylor's formula* for univariate polynomials.

Corollary 0.1.10 *Let K be a field of characteristic zero and $P \in K[X]$. Then $a \in K$ is a root of order $k \geq 1$ of the polynomial P if and only if $P^{(i)}(a) = 0$ for all $i \in \{0, 1, \ldots, k-1\}$ and $P^{(k)}(a) \neq 0$.*

Proof: Apply (2) with $X = a$ and $Y = X - a$. □

[1] I.e. if $n \cdot 1 \neq 0$ in A for all $n \in \mathbb{N} \setminus \{0\}$.

0.1.2 Elimination theory. Resultants

Elimination theory originates in the study of the following problem: If f, g are two univariate polynomials with coefficients in an integral domain A, find necessary and sufficient conditions for f and g to have common roots in an extension of the domain A. There exists a computable function $\mathrm{Res}(f,g)$ – it is a function of all the coefficients of f and g – such that f and g have a common root if and only if $\mathrm{Res}(f,g)=0$. The function $\mathrm{Res}(f,g)$ is called the *resultant* of the polynomials f and g.

There are several possibilities to introduce the resultant of two polynomials. We will deal with the resultants considered by Bézout, Jacobi and Cauchy. They obtained the resultant as the determinant of a *resultant matrix* with entries the coefficients of the corresponding polynomials.

Problems of elimination theory, such as the study of coprimeness and the determination of the greatest common divisors of two or several polynomials or polynomial matrices, play important roles in many domains as ring theory, differential systems, linear algebra, algebraic geometry, commutative algebra.

Observe that $\gcd(f,g)\neq 0$ if and only if there exist polynomials u, v such that

$$uf+vg=0, \quad \text{with} \quad \deg(u)<\deg(g). \tag{3}$$

In his celebrated memoir on elimination theory [**?**] (1764), *Bézout* described several methods to construct the resultant as the determinant of a convenient *resultant matrix*. Similar, but less general attempts had also been considered by *Euler* [**?**] (1748). Bézout's results were reconsidered during 19th century, among others, by *Jacobi* [**?**] (1836), *Sylvester* [**?**] (1840) and *Cauchy* [**?**] (1840).

Let

$$f(X)=\sum_{i=0}^{m} a_i X^i, \qquad g(X)=\sum_{i=0}^{n} b_i X^i$$

be nonconstant polynomials in one variable with the coefficients in a field k. The usual resultant matrix associated to them is the *Sylvester matrix*

$$S(f,g)=\begin{pmatrix}
a_0 & a_1 & \dots & \dots & a_{m-1} & a_m & 0 & \dots & 0 & 0\\
0 & a_0 & a_1 & \dots & \dots & a_{m-1} & a_m & \dots & 0 & 0\\
0 & 0 & a_0 & a_1 & \dots & \dots & a_{m-1} & a_m & \dots & 0\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
0 & \dots & \dots & 0 & a_0 & a_1 & \dots & a_{m-1} & a_m & 0\\
0 & 0 & \dots & \dots & 0 & a_0 & a_1 & \dots & a_{m-1} & a_m\\
b_0 & b_1 & \dots & \dots & b_{n-1} & b_n & 0 & \dots & 0 & 0\\
0 & b_0 & b_1 & \dots & \dots & b_{n-1} & b_n & 0 & \dots & 0\\
0 & 0 & b_0 & b_1 & \dots & \dots & b_{n-1} & b_n & \dots & 0\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
0 & \dots & \dots & 0 & b_0 & b_1 & \dots & b_{n-1} & b_n & 0\\
0 & \dots & \dots & 0 & 0 & b_0 & b_1 & \dots & b_{n-1} & b_n
\end{pmatrix}.$$

which is a $(m+n)\times(m+n)$–matrix. Note that $\det S(f,g)$ is the determinant of the linear system associated to (3).

Suppose that $\deg(f)=m\geq n=\deg(g)$. Bézout associated to f and g another resultant matrix:

$$B(f,g)=\begin{pmatrix}
c_{00} & c_{01} & \dots & \dots & c_{0,n-1} & c_{0n} & \dots & \dots & c_{0,m-1}\\
c_{10} & c_{11} & \dots & \dots & c_{1,n-1} & c_{1n} & \dots & \dots & c_{1,m-1}\\
c_{20} & c_{21} & \dots & \dots & c_{2,n-1} & c_{2n} & \dots & \dots & c_{2,m-1}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
c_{n-1,0} & c_{n-1,1} & \dots & \dots & c_{n-1,n-1} & c_{n-1,n} & \dots & \dots & c_{n-1,m-1}\\
b_0 & b_1 & \dots & b_{n-1} & b_n & 0 & 0 & \dots & 0\\
0 & b_0 & b_1 & \dots & b_{n-1} & b_n & \dots & \dots & 0\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
0 & 0 & \dots & b_0 & b_1 & \dots & \dots & b_{n-1} & b_n
\end{pmatrix}.$$

This device is called the *abridged method*. It produces a resultant matrix $B(f,g)$ (called the *Bézoutian* of order $m = \max\big(\deg(f), \deg(g)\big)$, while Sylvester's matrix $S(f,g)$ is of order $m+n$.

0.1.3 The abridged method of Bézout

Bézout started the exposition of his abridged method by considering the case of two polynomial equations of second degree

$$\left\{ \begin{array}{rcl} f(x) & = & Ax^2 + Bx + C = 0 \\ \\ g(x) & = & A'x^2 + B'x + C' = 0\,. \end{array} \right. \tag{4}$$

He constructed two polynomial equations from (4) multiplied by the polynomials $u = Mx + N$, respectively $v = M'x + N'$. From the identification with 0 he obtained a homogeneous linear system in the indeterminates $M,\ N,\ M'$ and N' and deduced that if the polynomial equations (4) have a common root then their coefficients satisfy the condition

$$(AB' - A'B)(B'C - BC') + (AC' - A'C)^2 = 0 \tag{5}$$

The condition (5) is equivalent to

$$\left| \begin{array}{cc} BC' - B'C & AC' - A'C \\ \\ AC' - A'C & AB' - A'B \end{array} \right| = 0,$$

which means

$$\det B(f,g) = 0. \tag{6}$$

After the presentation of this example, the method is described for two polynomials of the same degree $m = n$. Bézout carefully worked out the cases $m = 2$, $m = 3$ and $m = 4$ ([**?**] pp. 535–540). In the general case his techniques lead to the following construction.

For every $i \in \{0, 1, 2, \ldots, n-1\}$ let

$$f_i(x) = a_{i+1} + a_{i+2}x + \ldots + a_n x^{n-i-1} = \frac{1}{x^{i+1}}\left(f(x) - \sum_{j=0}^{i} a_j x^j\right),$$

$$g_i(x) = b_{i+1} + b_{i+2}x + \ldots + b_n x^{n-i-1} = \frac{1}{x^{i+1}}\left(g(x) - \sum_{j=0}^{i} b_j x^j\right)$$

be associated to the polynomials f and g, and construct the polynomials

$$B_i = f_i g - g_i f\,.$$

Note that $\deg B_i \leq n-1$. Writing

$$B_i(x) = \sum_{j=0}^{n-1} c_{ij} x^j = c_{i0} + c_{i1}x + \ldots + c_{ij}x^j + \ldots + c_{i,n-1}x^{n-1},$$

and mastering the linear techniques developed by Cramer in his treatise on linear systems and determinants ([**?**], 1750), Bézout considered the system

$$B_i(x) \,=\, 0, \qquad (0 \leq i \leq n-1)\,, \tag{7}$$

which is linear in $x^0, x^1, x^2, \ldots, x^{n-1}$. The system (7) leads to the nullity of $\det B(f,g)$, where

$$B(f,g) = \begin{pmatrix} c_{00} & c_{01} & \ldots & c_{0i} & \ldots & c_{0,n-1} \\ c_{10} & c_{11} & \ldots & c_{1i} & \ldots & c_{1,n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ c_{n-1,0} & c_{n-1,1} & \ldots & c_{n-1,i} & \ldots & c_{n-1,n-1} \end{pmatrix}.$$

The above considerations of Bézout essentially prove the following

Theorem 0.1.11 *The polynomials f and g have a common root if and only if* $\det B(f,g) = 0$.

Notice that one implication in Theorem 0.1.11 is obvious: if the polynomials f and g have a common root, say x, then the homogeneous linear system (7) has a nontrivial solution. Thus $\det B(f,g) = 0$.

The next step in Bézout's exposition of the abridged method is the study of the case of unequal degrees. To eliminate x between

$$f(x) = a_0 + a_1 x + \ldots + a_m x^m = 0$$
$$g(x) = b_0 + b_1 x + \ldots + b_n x^n = 0 \qquad m \geq n,$$

he considered derived relations which involve differences between polynomials of degree n, $n+1$, ...:

$$b_n f(x) - a_m x^{m-n} g(x) = 0,$$

$$\big(b_n x + b_{n-1}\big) f(x) - \big(a_m x^{m-n+1} + a_{m-1} x^{m-n}\big) g(x) = 0\,,$$

and so on.

This method is practically the abridged method for equal degrees in the case of the polynomials $f(x)$ and $x^{m-n}g(x)$. Bézout remarked the possible existence of extraneous factors and observed that their removal *"needs enough tedious care"*.

0.1.4 Jacobi's version

The next step in the development of the *abridged method* was a memoir of *Jacobi* [**?**] (1836), one of the last important algebraic works written in *Latin*[2]. In this paper devoted to elimination theory Jacobi focused on the case of two polynomials of equal degrees. In the introduction he recorded the methods of Euler and Bézout. Jacobi declared to develop the abridged method of Bézout just from the first lines of his memoir. He succeeded in giving a clear exposition of the abridged method in the case of equal degrees and he added new results and applications.

Like Bézout, Jacobi considered two polynomials in one variable of degree n,

$$f(X) = \sum_{i=0}^{n} a_i X^i, \qquad g(X) = \sum_{i=0}^{n} b_i X^i$$

and he associated to them the polynomials

$$\begin{aligned} f_i(X) &= a_{i+1} + a_{i+2} X + \ldots + a_n X^{n-i-1}, \\ g_i(X) &= b_{i+1} + b_{i+2} X + \ldots + b_n X^{n-i-1}, \end{aligned}$$

for $i \in \{0, 1, 2, \ldots, n-1\}$. Afterwards he followed Bézout's device, considering the polynomials

$$\begin{aligned} B_i(X) &= f_i g - g_i f \\ &= c_{0i} X^0 + c_{1i} X^1 + c_{2i} X^2 + \ldots + c_{n-1,i} X^{n-1} \end{aligned}$$

and establishing that f and g have a common root if and only if $\det B(f,g) = 0$, where $B(f,g)$ is the same resultant matrix considered by *Bézout* (p. 102).

Remark: In the memoir of *Jacobi* there are used the notations m_i for B_i, respectively α_{ji} for c_{ij} as defined in the former section.

Jacobi remarked that

$$c_{ij} = c_{ji}, \quad \text{for all} \quad i, j \in \{0, 1, 2, \ldots, n-1\}$$

and obtained the following relation:

[2]The last serious attempt to revive the use of Latin in Science was G. Peano's *Latino sine Flexione* (1889).

Theorem 0.1.12

$$\left\{\begin{array}{rcl} c_{ij} & = & a_{i+1}b_j + a_{i+2}b_{j-1} + \ldots + a_{i+j+1}b_0 \\ & - & \left(b_{i+1}a_j + b_{i+2}a_{j-1} + \ldots + b_{i+j+1}a_0\right) \\ & = & a_{j+1}b_i + a_{j+2}b_{i-1} + \ldots + a_{i+j+1}b_0 \\ & - & \left(b_{j+1}a_i + b_{j+2}a_{i-1} + \ldots + b_{i+j+1}a_0\right). \end{array}\right. \tag{8}$$

0.1.5 Cauchy's contribution

As with many of his papers, the memoir of Cauchy [**?**] contains an exhaustive presentation of a particular topic, in this case of elimination theory. He discussed the works of his predecessors Bézout [**?**] and Euler [**?**], but also a paper of his contemporary Sylvester [**?**]. However, he completely ignored the memoir of Jacobi [**?**], published four years before the mentioned paper of Sylvester.

After the exposition of the methods that led Euler, Bézout and Sylvester to the Sylvester matrix, he focused on the abridged method of Bézout.

Remark: The paper of Sylvester mentioned by Cauchy contains only some hints about an elimination method. Cauchy noticed that the resultant matrix suggested by Sylvester is of order $m+n$. It seems that the Sylvester matrix got his name because of the subsequent quotations of the paper of Sylvester, without reference to the earlier work of Bézout and Euler.

Cauchy derived the entries of the Bézoutian by a method different from those of Bézout and Jacobi. As we shall see, his techniques improve the algorithm of the abridged method. He states that it is sufficient to consider the case of two polynomials of the same degree, because it may be supposed that some of the coefficients involved are zero in the case of different degrees.

Let

$$f(X) = \sum_{i=0}^{n} a_i X^{n-i}, \qquad g(X) = \sum_{i=0}^{n} b_i X^{n-i}.$$

be such polynomials.

For $j \in \{0, 1, 2, \ldots, n-1\}$ Cauchy considered the polynomials

$$f_j = \sum_{i=0}^{j} a_i X^{j-i}, \qquad \widetilde{f_j} = \sum_{i=j+1}^{n} a_i X^{n-i}$$

associated to f and the corresponding polynomials g_j, $\tilde{g}_j$ associated to g. He defined A_{kj} as the coefficient of X^{n-k-1} in the polynomial

$$\begin{array}{rcl} C_j & = & \dfrac{1}{X^{n-j}}(f_j\widetilde{g}_j - g_j\widetilde{f}_j) \\ & = & (a_0X^j + a_1X^{j-1} + \cdots + a_j)\cdot(b_{j+1}X^{n-j-1} + \cdots + b_{n-1}X + b_n) \\ & - & (b_0X^j + b_1X^{j-1} + \cdots + b_j)\cdot(a_{j+1}X^{n-j-1} + \cdots + a_{n-1}X + a_n). \end{array}$$

If f and g have a common root, then

$$C_j(x) = \sum_{i=0}^{n-1} A_{ij}x^{n-i-1} = 0, \quad \text{for all} \quad j \in \{0, 1, 2, \ldots, n-1\}, \tag{9}$$

But (9) can be considered as a homogeneous linear system in the indeterminates x^0, x^1, x^2, ..., x^{n-1} having a nontrivial solution[3]. Therefore

$$\det\Big(A_{ij}\Big)_{0\leq i,j\leq n-1} = 0\,. \tag{10}$$

Proposition 0.1.13 *Let f, g be nonconstant polynomials of equal degrees. Then the resultant matrices of Bézout and Cauchy are equal.*

Proof: Suppose

$$f(X) = \sum_{i=0}^{n} a_i X^{n-i}, \qquad g(X) = \sum_{i=0}^{n} b_i X^{n-i}\,.$$

With the notations of Bézout and Cauchy we have

$$\begin{aligned} C_i &= f_i\widetilde{g}_i - g_i\widetilde{f}_i \\ &= f_i(g-g_i) - g_i(f-f_i) \\ &= f_i g - g_i f \\ &= B_i\,. \end{aligned}$$

Therefore the resultant matrices formed with the coefficients of the polynomials $B_0, \ldots, B_{n-1}$, respectively $C_0, \ldots, C_{n-1}$, are equal. □

Remark: Cauchy uses a smaller number of computations, because he uses only cancellations of f and g. Therefore the cost of his method is smaller than the cost of the method of Bézout.

Definition: Let $f \in A[X]$, $n = \deg(f)$. The product

$$D(f) = \mathrm{lc}(f)^{2n-2} \prod_{1\leq j<i\leq n} (\alpha_i - \alpha_j)^2$$

is called the *discriminant* of the polynomial f.

0.1.6 The companion matrix

Consider the polynomial

$$F(X) = X^n + a_{n-1}X^{n-1} + \ldots + a_0 \in K[X],$$

where K is a field. To F there is associated a companion matrix C_F.

Definition:

$$C_F = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \ldots & -a_{n-2} & -a_{n-1} \end{pmatrix}.$$

is called the *companion matrix* of the polynomial F.

Note that C_F is a $n \times n$–matrix.

Remark: The eigenvalues of the companion matrix C_F are exactly the roots of f, counted with the same multiplicities. Therefore the characteristic polynomial of the matrix C_F is F.

[3]We take have $x^0 = 1 \neq 0$.

Proposition 0.1.14 *If $F \in K[X]$ is nonconstant and $A = C_F$ is the companion matrix of F, then $F(A) = 0$.*

Proof: We consider the quotient ring $R = K[X]/(F)$. Let $T : R \longrightarrow R$ be the linear mapping defined by $T(P) = XP \bmod F$. Then C_F is the matrix associated to T with respect to the basis $\mathcal{B} = \{1, X, X^2, \ldots, X^{n-1}\}$ of A. It follows that F is the characteristic polynomial of the matrix $A = C_F$, thus $F(A) = 0$. □

Exercises

1. Let $A = \mathbb{Z}_4$, the ring of the remainders modulo 4. Prove that:

i. $\mathbb{Z}_4$ has zero divisors.

ii. If $a \in A$ is a zero divisor, then the polynomial $P(X) = aX$ has at least two distinct roots.

Hint: We have $\widehat{2} \cdot \widehat{2} = \widehat{0}$ in $\mathbb{Z}_4$.

2. Let $A = \mathbb{Z}_4$ and $P(X) = 2X^2 - 2X \in \mathbb{Z}_4[X]$. Use Proposition 0.1.1 for proving that P has not a unique representation as a product of polynomials of degree one.

Hint: Observe that from $P(0) = 0$ we have $P = X \cdot (2X-2)$, but from $P(2) = 0$ it follows that $P = (X-2) \cdot (2X+2)$.

3. Let $A = \mathbb{Z} \times \mathbb{Z}$. For $a = (b, c)$, $a' = (b', c') \in A$ define $a + a' = (b + b', c + c')$ and $a \cdot a' = (bb', cc')$. Prove that:

i. $\mathbb{Z} \times \mathbb{Z}$ is not an integral ring.

ii. There exist nonzero polynomials over $\mathbb{Z} \times \mathbb{Z}$ which have an infinite number of roots.

Hint: For ii consider, for example, $P(X) = (1, 0)X \in A[X]$.

4. Let $P \in \mathbb{Q}[X]$ be such that $P(X^2 + 1)$ is the null polynomial. Prove that $P(X)$ is the null polynomial.

Hint: Suppose $P \neq 0$ and let $d = \deg(P)$. Then notice that $P(1) = P(2) = \ldots = P((d-1)^2 + 1) = P(d^2 + 1)$.

5. Let A be an integral domain and let $f, g, h \in A[X]$. Prove that:

i. $\operatorname{Res}(f, 0) = 0$.

ii. $\operatorname{Res}(f, g) = (-1)^{\deg(f)\deg(g)} \operatorname{Res}(g, f)$.

iii. $\operatorname{Res}(f, gh) = \operatorname{Res}(f, g) \operatorname{Res}(f, h)$.

6. Let $f(X) = a(X - \alpha_1) \ldots (X - \alpha_m)$ and $g(X) = b(X - \beta_1) \ldots (X - \beta_n)$. Prove that

$$\operatorname{Res}(f, g) = a^n \prod_{i=1}^{m} g(\alpha_i) = (-1)^{mn} b^m \prod_{j=1}^{n} f(\beta_j) = a^n b^m \prod_{\substack{1 \le i \le m \\ 1 \le j \le n}} (\alpha_i - \beta_j).$$

7. Let $f \in A[X]$, $n = \deg(f)$. Prove that

$$\operatorname{Res}(f, f') = (-1)^{\frac{n(n-1)}{2}} \operatorname{lc}(f) D(f).$$

8. Let A be an integral domain and let $f \in A[X]$. Prove that the polynomial f has multiple roots if and only if $D(f) = 0$.

9. Let $f(X) = X^3 + bX + c \in A[X]$. Prove that

$$D(f) = -4b^3 - 27c^2.$$

10. Let A be a domain and $f_1, \dots, f_m \in A[X]$.

i. Prove that

$$D_k\big(f_1 \cdots f_m\big) = \sum_{k_1 + \cdots + k_m = k} D_{k_1}(f_1) \cdots D_{k_m}(f_m), \quad \text{for all} \quad k \in \mathbb{N},$$

where D_k is the hyperderivative of order k.

ii. If $a \in A$, then

$$D_k(X-a)^m = \binom{m}{k} (X-a)^{m-k}, \quad \text{for all} \quad k, m \in \mathbb{N}.$$

11. Let f, g be nonconstant polynomials of degree $n \geq 1$ over a domain A and let $C_i, \widetilde{f}_i, \widetilde{g}_i$ be the polynomials associated to f, g in Cauchy's elimination method. Prove that

$$C_i = -\widetilde{f}_i g + \widetilde{g}_i f, \quad \text{for} \quad i = 0, 1, \dots, n-1.$$

Hint: Use the decomposition $f = f_i + \widetilde{f}_i$ and the relation $C_i = f_i g - g_i f$ obtained in Proposition 0.1.13.

12. Let $F(X)$ and $G(X)$ be coprime polynomials with integer coefficients and of degree at most d. Prove the following two results:

i. There exists a positive integer R, depending on F and G, such that for all rational numbers a/b,

$$\gcd\big(b^d F(a/b), b^d G(a/b)\big) \text{ divides } R.$$

ii. There are constants k_1 and k_2, depending on F and G, such that for all rational numbers a/b which are not roots of G,

$$d\, \mathrm{h}(a/b) - k_1 \leq \mathrm{h}\left(\frac{F(a/b)}{G(a/b)}\right) \leq d\, \mathrm{h}(a/b) + k_2,$$

where the *height* $\mathrm{h}(x)$ of a nonzero rational number x is defined by $\mathrm{h}(x) = \log \max\{|m|, |n|\}$ with $x = m/n$ and $\gcd(m, n) = 1$, whereas $\mathrm{h}(0) = 0$.

Hint: Prove i) using the resultant A of the polynomials F and G and the fact that there exist integer polynomials U and V such that

$$U(X)F(X) + V(X)G(X) = A.$$

The proof of the second inequality in ii) is similar to the proof of i). To prove the first one use again the previous resultant identity. For more details see:
J. H. Silverman and J. Tate [**?**], pp. 72–75.

13. Let F be a univariate polynomial of degree $n \geq 1$ over the integral domain A. Prove that the companion matrix C_F is the matrix associated to the linear mapping

$$T : K[X]/(F) \longrightarrow K[X]/(F),$$

defined by $T(P) = XP \bmod F$, with respect to the basis $\{1, X, \ldots, X^{n-1}\}$.

Hint: Note that $\widehat{P} \in K[X]/(F)$ is the set $\widehat{P} = \{P + QF; Q \in K[X]\}$.

14. Compute the rank of the companion matrix of the polynomial:

i. $F(X) = X^5 - 2X^2 + X - 4 \in \mathbb{Z}[X]$.

ii. $F(X) = X^6 + \frac{1}{2}X^5 - \frac{3}{7}X^3 + 2X + 4 \in \mathbb{Q}[X]$.