Journal of Pure and Applied Algebra 65 (1990) 293-309
North-Holland

ON NEWTON'S RULE AND SYLVESTER'S THEOREMS

Jean-Claude YAKOUBSOHN
Laboratoire d'Analyse Numérique, 118 route de Narbonne, Université Paul Sabatier, 31062 Toulouse Cédex, France

Communicated by M.F. Coste-Roy
Received 4 April 1989

"This is to Newton's what Fourier's is to Descartes."
Sylvester

We study a rule given by Newton and proved by Sylvester, on an upper bound for the number of real roots of a polynomial. The notion of variation-permanence permits us to ameliorate Descartes' rule. We explain the link between a lemma given by Cauchy and Newton's rule and we give some applications.

0. Introduction

The Budan–Fourier theorem of which Descartes' rule is a special case, gives an upper bound for the number of roots that a polynomial has in a given interval. The rule stated by Newton is not well-known at present: it was generalized and proved by Sylvester in 1865. However this rule is better than Descartes'. The purpose of this article is to give a synthesis of Sylvester's work, but with a modern and different approach. Marchand had a similar goal when he wrote his thesis under the direction of Hurwitz [5]. We will follow a different method based on two original lemmas (Lemmas 4.6 and 4.8). Lemma 4.6 establishes a relation between the roots of a polynomial and its derivatives. This lemma uses a result given by Cauchy (Lemma 4.5). Lemma 4.8 is another way to write Lemma 4.6 using the theory developed by Karlin [4].

We explain the origin of the even number in the Budan–Fourier theorem. This approach is contained in Sections 1–4. In Section 5, we show how Newton's rule permits us to establish inequalities satisfied by the coefficients of a polynomial with real roots. Furthermore, we compute the average number of roots of a polynomial when the coefficients are equal to -1 or 1. In conclusion, we study the complexity of computation of Newton's rule.

1. Preliminaries and notations

Let

$$f(x) = \sum_{k=0}^{n} a_k x^{n-k}$$

be a polynomial of degree n. The derivatives of f are denoted by $f_0, f_1, \dots, f_k, \dots, f_n$ with $f_0 = f$. Rather than writing $g(x)$, it is more convenient to write g. We consider the functions F_k given by:

$$F_0 = f^2,$$
$$F_k = r_k f_k^2 - r_{k-1} f_{k-1} f_{k+1}, \quad k = 1, \dots, n-1,$$
$$F_n = f_n^2.$$

The constants r_0 and r_1 are given, and the r_k are defined by the two relations:

$$r_k > 0, \qquad k = 0, \dots, n-1, \tag{1}$$
$$r_{k+1} = 2r_k - r_{k-1}, \quad k = 1, \dots, n. \tag{2}$$

The second relation is introduced to simplify the calculation of the derivative of the functions F_k. These two relations are equivalent to:

$$nr_1 - (n-1)r_0 \geq 0, \tag{1'}$$
$$r_k = kr_1 - (k-1)r_0, \quad k = 2, \dots, n. \tag{2'}$$

We denote $r = (r_0, r_1)$.

Now let for $k = 1, \dots, n$,

$$S_k = \begin{pmatrix} f_{k-1} & f_k \\ F_{k-1} & F_k \end{pmatrix}$$

be the term of order k of the double sequence (S):

$$\begin{matrix} f_0 & f_1 & \cdots & f_n \\ F_0 & F_1 & \cdots & F_n. \end{matrix} \tag{S}$$

We investigate the signs of polynomials comprising S_k. If any of the f's and F's in S_k is zero in x, only the following cases occur:

Case 1. permanence-permanence denoted by pP

$$\begin{matrix} ++ & -- & ++ & -- \\ ++ & ++ & -- & -- \end{matrix}$$

Case 2. permanence-variation denoted by pV

$$\begin{matrix} ++ & -- & ++ & -- \\ +- & +- & -+ & -+ \end{matrix}$$

Case 3. variation–permanence denoted by vP

$+-$	$+-$	$-+$	$-+$
$++$	$--$	$++$	$--$

Case 4. variation–variation denoted by vV

$+-$	$-+$	$+-$	$-+$
$+-$	$+-$	$-+$	$-+$

The number of permanence–permanences of (S) at x for a value of r is denoted by $\mathrm{pP}_r(x,f)$. The notations $\mathrm{p}(x,f)$, $\mathrm{v}(x,f)$, $\mathrm{V}(x,f)$, $\mathrm{P}(x,f)$, $\mathrm{pV}_r(x,f)$, $\mathrm{vP}_r(x,f)$ and $\mathrm{vV}_r(x,f)$, are defined likewise.

Let a and b be two real numbers with $a<b$. We denote by $\mathrm{ZR}(a,b,f)$ the number of real roots of polynomial f in the interval $[a,b]$, by $\mathrm{ZR}_+(f)$ (resp. $\mathrm{ZR}_-(f)$) the number of positive (resp. negative) roots of f. Every root is counted with its order of multiplicity.

When there is no ambiguity, we abbreviate the previous notation into: $\mathrm{p}(x)$, $\mathrm{v}(x), \dots, \mathrm{pV}_r(x), \dots, \mathrm{vV}_r(x)$, $\mathrm{ZR}(a,b)$, ZR_+, ZR_-.

We denote by $\mathrm{vP}(x)$ (resp. $\mathrm{pP}(x)$), the minimum of $\mathrm{vP}_r(x)$'s (resp. $\mathrm{pP}_r(x)$'s).

2. Statements of Sylvester's theorem and Newton's rule

We suppose that the real numbers a and b are roots of neither f_k nor F_k.

Theorem 2.1. *The number of real roots of a polynomial f with real coefficients is given by*:

$$\mathrm{ZR}(a,b) = \mathrm{vP}(a) - \mathrm{vP}(b) - 2\alpha, \quad \textit{where } \alpha \in \mathbb{N}.$$

Theorem 2.2. *The number of real roots of a polynomial f with real coefficients is given by*:

$$\mathrm{ZR}(a,b) = \mathrm{pP}(b) - \mathrm{pP}(a) - 2\beta, \quad \textit{where } \beta \in \mathbb{N}.$$

Newton's rule. (1) *The number of real positive roots of a polynomial with real coefficients is given by*:

$$\mathrm{ZR}_+ = \mathrm{vP}(0) - 2\alpha, \quad \textit{where } \alpha \in \mathbb{N}.$$

(2) *The number of real negative roots of a polynomial with real coefficients is given by*:

$$\mathrm{ZR}_- = \mathrm{pP}(0) - 2\beta, \quad \textit{where } \beta \in \mathbb{N}.$$

3. Remarks

3.1. If either of the real numbers a, b is a root of function f_k or F_k $(k \geq 1)$ then vP(a) or vP(b) is not defined. In the case where none of the F_k's is identically zero, we consider that:

$$\mathrm{vP}(a)-\mathrm{vP}(b)=\mathrm{vP}(a+h)-\mathrm{vP}(b-h),$$

where h is a positive infinitesimal. We proceed identically with the difference pP(b) − pP(a).

3.2. Marchand states a necessary and sufficient condition for one of the F_k's to be identically zero ($F_k \equiv 0$):

Proposition 3.3. *A function F_k is identically zero iff the following two conditions hold*:

$$nr_1-(n-1)r_0=0, \tag{3}$$

$$f_{k-1}=c(x-d)^{n-k+1}, \quad c \text{ and } d \text{ being real numbers.} \tag{4}$$

Proof. For the proof see Marchand [5]. □

A corollary is that if $f_k \equiv 0$, then $F_{k+1} \equiv \cdots \equiv F_{n-1} \equiv 0$.

Starting with the sequence (S), we define a sequence (S′) giving sign-conventions for the F_k which are identically zero.

Convention 3.4. If at x,

$$F_{k-1} \neq 0, \quad F_k \equiv \cdots \equiv F_{n-1} \equiv 0, \qquad k \geq 2,$$

holds, in the sequence (S) we substitute F_{n-i} by $(-1)^i$ for $i=1,\dots,n-k$.

Convention 3.5. If at x

$$f_{k-1} \neq 0, \quad f_k = \cdots = f_{n-1} = 0, \quad f_n \neq 0,$$
$$F_{k-1} \neq 0, \quad F_k \equiv \cdots \equiv F_{n-1} \equiv 0, \quad F_n \neq 0,$$

holds, two cases may occur:

(1) $f_n f_{k-1} > 0$. We proceed as in Convention 3.4 for the f_{n-i} by Sgn(f_n) for $i=1,\dots,n-k$.

(2) $f_n f_{k-1} < 0$ in the sequence (S). We substitute the F_{n-i} and f_{n-i} respectively by $(-1)^i$ and Sgn(f_n) for $i=1,\dots,n-k+1$; and F_k and f_k respectively by Sgn(F_{k-1}) and Sgn(f_{k-1}).

Convention 3.6. If at x

$$f_0 \neq 0, \quad f_1 \neq 0, \dots, f_{n-1} \neq 0, \quad f_n \neq 0,$$
$$F_0 \neq 0, \quad F_1 \equiv \cdots \equiv F_{n-1} \equiv 0, \quad F_n \neq 0,$$

holds, we substitute in the sequence (S) F_{n-i} by Sgn(F_n) for $i=1,\dots,n-1$.

3.3. The method of proof for Theorems 2.1 and 2.2 (as used by Sylvester and Marchand) is the same as Fourier's: we investigate the changes of signs in the sequence (S) which hold only when one of f_k's or F_k's is zero at a point in $[a,b]$. The technical background is Taylor's formula. We give two results that state the signs of f_k's and F_k's in the neighbourhood of a point x when $f_k(x)=0$ or $F_k(x)=0$.

Proposition 3.7. *If at x,*

$$f_k=f_{k+1}=\cdots=f_{k+m-1}=0,\quad f_{k+m}\neq 0,\quad 0\le k\le n-1,\quad 1\le m\le n-k-1,$$

holds, then

$$F_k(x+h)=-\frac{h^{m-1}}{(m-1)!}r_{k-1}f_{k-1}f_{k+m}+o(h^{m-1}),\tag{5}$$

$$\begin{aligned}F_{k+i}(x+h)&=\frac{h^{2m-2i}}{[(m-i)!]^2}\,\frac{1}{(m-i+1)}\,r_{k+m}f_{k+m}^2\\&\quad+\frac{h^{2m-2i+1}}{(m-i)!\,(m-i+1)!}\,\frac{2}{(m-i)(m-i+1)(m-i+2)}\\&\quad\times r_{k+m+1}f_{k+m-1}f_{k+m+1}+o(h^{2m-2i+1}).\end{aligned}$$

for $i=1,\dots,m-1$. (6)

Proposition 3.8. *If at x*

$$F_k=F_{k+1}=\cdots=F_{k+m-1}=0,\quad F_{k+m}\neq 0,$$
$$f_{k+i}\neq 0,\quad \textit{for } i=0,\dots,m,$$
$$0\le k\le n-1,\quad 1\le m\le n-k-1,$$

holds, then

$$F_k(x+h)=\frac{h^m}{m!\,f_{k+m}}f_kF_{k+m}+o(h^m).$$

These two propositions are proved in the Appendix.

3.4. *Role of constants* r_0 *and* r_1. The introduction of the constants r_k is due to Sylvester and is also used by Marchand. But these authors do not explain how to calculate r_0 and r_1 in order to minimize the quantities $\mathrm{vP}_r(a)-\mathrm{vP}_r(b)$ or $\mathrm{pP}_r(b)-\mathrm{pP}_r(a)$. We give a method for this. The calculation of F_k is for $k=1,\dots,n$:

$$F_k=[kf_k^2-(k-1)f_{k-1}f_{k+1}]\,r_i-[(k-1)f_k^2-(k-2)f_{k-1}f_{k+1}]\,r_0.$$

Let

$$p_k=\frac{(k-1)f_k^2-(k-2)f_{k-1}f_{k+1}}{kf_k^2-(k-1)f_{k-1}f_{k+1}},$$

$$m_k=\inf(p_k,p_{k+1}),\quad M_k=\sup(p_k,p_{k-1}).$$

We consider the following notation:

$$D = \left\{(r_0, r_1):\ r_0 > 0,\ r_1 = \frac{n-1}{n} r_0\right\},$$

$$S = \left\{(r_0, r_1):\ r_0 > 0,\ r_1 > \frac{n-1}{n} r_0\right\},$$

$$D_k(x) = \{(r_0, r_1):\ r_0 > 0,\ r_1 = M_k r_0\},$$

$$d_k(x) = \{(r_0, r_1):\ r_0 > 0,\ r_1 = m_k r_0\},$$

$$S_k(x) = \{(r_0, r_1):\ r_1 > m_k r_0,\ r_1 < M_k r_0\}.$$

3.4.1. Minimization of $\mathrm{VP}_r(x)$. There are three cases:

(a) $\bar{S}_k \cap \bar{S} - (0,0) = \emptyset$ for any k for which $f_k f_{k+1} < 0$. Then the number of variation–permanences is the same for any (r_0, r_1) in $\bar{S}$.

(b) There exists $k_1, \ldots, k_l$ for which:

$$f_{k_i} f_{k_j+1} < 0, \quad \bar{S}_{k_i} \cap \bar{S} \neq \emptyset, \qquad \text{for } i = 1, \ldots, l, \quad \bigcap_{i=1}^{l} \bar{S}_{k_i} = \emptyset.$$

Then the number of variation–permanences is minimum for any (r_0, r_1) in one of $S_{k_i} \cap S$, D_{k_i} or d_{k_i}.

(c) There exist $k_1, \ldots, k_l$ such that:

$$I = \bar{S} \cap \left(\bigcap_{i=1}^{l} \bar{S}_{k_i}\right) \neq \emptyset.$$

Then the number of variation–permanences is minimum for any (r_0, r_1) in I.

3.4.2. Maximization of $\mathrm{VP}_r(x)$. The number of variation–permanences in x is maximum for any (r_0, r_1) in one of the $S - (\bar{S} \cap \bar{S}_{k_i})$, if cases (b) or (c) hold. In case (a), any (r_0, r_1) in $\bar{S}$ is admissible.

3.4.3. Minimization of $\mathrm{vP}_r(a) - \mathrm{vP}_r(b)$. The two previous descriptions are used here. We calculate respectively:

(a) the $S_{k_i}(a) \cap S$, $D_{k_i}(a)$ and $d_{k_i}(a)$ for $i = 1, \ldots, l$ when $f_{k_i}(a) f_{k_i}(a) < 0$;

(b) the $S - (\bar{S}_{k_j}(b) \cap \bar{S})$ for $j = 1, \ldots, n$ when $f_{k_j}(b) f_{k_j+1}(b) < 0$.

Thus we obtain q subsets Q_k of S. We order the Q_k via the following relation:

$$\text{if } Q_k = \left\{(r_0, r_1):\ r_0 > 0,\ r_1 \geq a_k r_0,\ r_1 \leq b_k r_0,\ \frac{n-1}{n} \leq a_k \leq b_k\right\},$$

$$Q_k < Q_i \quad \text{iff} \quad b_k < a_i,$$

$$Q_k > Q_i \quad \text{iff} \quad a_k > b_i,$$

$$Q_k \subset Q_i \quad \text{iff} \quad a_i \leq a_k \leq b_k \leq b_i.$$

Then we investigate the difference $\mathrm{vP}(a) - \mathrm{vP}(b)$ in each set of the partition of S

induced by the Q_k thus ordered. Then there exists a subset Q_k of S where the latter is minimal.

3.5. Examples.

Example 1. We want to know the number of positive roots of

$$f(x) = \sum_{k=0}^{7} (-1)^k x^{7-k}.$$

By Descartes' rule, we obtain $ZR_+ = 7 - 2\alpha$.

The calculation of F_k gives: $F_k = k!\,(k-1)!\,(r_1 - 2r_0)$ at $x=0$. For $r_0 = 1$ and $r_1 > 2$, F_k is positive and we conclude that $vP(0) = 7$. For $r_0 = 1$ at $\frac{6}{7} < r_1 < 2$, F_k is negative and we conclude that $vP(0) = 5$. So $ZR_+ = 5 - 2\alpha$.

For $r_0 = 1$ and $r_1 = 2$, all the F_k's $(1 \le k \le 6)$ are zeros. We use Proposition 3.8 in order to know the sign of F_k in an infinitesimal h. We obtain

$$F_k(h) = \frac{h^6}{6!\,f_7(0)} f_k(0)\,F_7(0).$$

The sign of $F_k(h)$ is that of $f_k(0)$. So $vP(0) = 1$. Finally $ZR_+ = 1$.

Sylvester compares Theorems 2.1 and 2.2 to "un fusil à deux coups, si l'un des canons rate l'autre peut atteindre le but". We give an example illustrating this.

Example 2. We want to know the number of roots in the interval $[0, 1]$ of the polynomial

$$f(x) = x^3 - x^2 - \tfrac{1}{2}x + 2.$$

For $x = 0$,

$$S \equiv \begin{pmatrix} 2 & -\frac{1}{2} & -2 & 6 \\ 4 & \frac{1}{4}r_1 + 4r_0 & 11r_1 - 4r_0 & 36 \end{pmatrix}.$$

For $x = 1$,

$$S \equiv \begin{pmatrix} \frac{3}{2} & \frac{1}{2} & 4 & 6 \\ \frac{9}{2} & \frac{1}{4}r_1 - 6r_0 & 29r_1 - 16r_0 & 36 \end{pmatrix}.$$

Then $v(0) - v(1) = vP(0) - vP(1) = 2$, and $pP(1) - pP(0) = 1 - 1 = 0$.

We now give an example that shows the limitations of Sylvester's theorem.

Example 3. The polynomial

$$f(x) = 5x^3 - 8x^2 + 4x - \tfrac{7}{12}$$

possesses one root in the interval $[0, 1]$. For $x = 0$,

$$S = \begin{pmatrix} -\frac{7}{12} & 4 & -16 & 30 \\ \frac{49}{144} & 16r_1 - \frac{28}{3}r_0 & 392r_1 - 256r_0 & 900 \end{pmatrix}.$$

For $x=1$,

$$S=\begin{pmatrix} \frac{5}{12} & 3 & 14 & 30 \\ \frac{25}{144} & 9r_1-\frac{35}{6} & 302r_1-196r_0 & 900 \end{pmatrix}.$$

We find: $\mathrm{v}(0)-\mathrm{v}(1)=\mathrm{vP}(0)-\mathrm{vP}(1)=\mathrm{pP}(1)-\mathrm{pP}(0)=3$.

4. Proof of Theorem 2.1

We use an inductive method in order to prove Theorem 2.1.

Definition 4.1. We define:

(1) *Extremum of the first kind*: a point with coordinates $(x,f(x))$ such that there exists a p for any infinitesimal h and for $k=1,\dots,2p$,

$$f_k(x)=0 \quad \text{and} \quad f_{2p}(x)f(x+h)<0;$$

(2) *Extremum of the second kind*: a point with coordinates $(x,f(x))$ such that there exists a p for any infinitesimal h and for $k=1,\dots,2p-1$,

$$f_k(x)=0 \quad \text{and} \quad f_{2p}(x)f(x+h)>0.$$

We denote by $E_1(a,b,f)$ and $E_2(a,b,f)$ respectively the number of extrema of polynomials of the first and second kind.

Definition 4.2. Let a and b such that: $f(a)f_1(a)f_2(a)\neq 0$ and $f(b)f_1(b)f_2(b)\neq 0$. We define $\sigma(a,b,f)$ by:

(1) $\sigma(a,b,f)=0$ if $f(a)f_1(a)f(b)f_1(b)>0$;

(2) $\sigma(a,b,f)=1$ if $f(a)f_1(a)f(b)f_1(b)<0$ and if $f(a)f(a)f_1(a)<0$ and $f(b)f_1(b)>0$;

(3) $\sigma(a,b,f)=-1$ if $f(a)f_1(a)f(b)f_1(b)<0$ and if $f(a)f_1(a)>0$ and $f(b)f_1(b)<0$.

Remark 4.3. If $f(a)f_1(a)f_2(a)=0$, we define $\sigma(a,b,f)$ by $\sigma(a+h,b,f)$ where h is infinitesimal, so that

$$f(a+h)f_1(a+h)f_2(a+h)\neq 0.$$

In the same way, if $f(b)f_1(b)f_2(b)=0$, we define $\sigma(a,b,f)$ by $\sigma(a,b-h,f)$.

Remark 4.4. $\sigma(a,b,f)$ indicates the behaviour of the curve that represents the polynomial f.

If $\sigma(a,b,f)=0$, the curve goes away from the x-axis in a neighbourhood of one of the bounds of the interval $[a,b]$, while it draws near to the x-axis in the neighbourhood of the other bound, while staying in $[a,b]$.

If $\sigma(a,b,f)=1$, the curve goes away from the x-axis in the neighbourhood of the bounds of the interval $[a,b]$, while staying in $[a,b]$.

If $\sigma(a,b,f)=-1$, the curve draws near to the x-axis in the neighbourhood of the bounds of the interval $[a,b]$, while staying in $[a,b]$.

We now give a relation between the number of roots of polynomial f, the number of the extrema of f in $[a,b]$ and $\sigma(a,b,f)$.

Lemma 4.5. (Cauchy). *Let* $\mathrm{NR}(a,b,f)$ *be the number of roots of the polynomial not counted with their order of multiplicity. Then*:

$$\mathrm{NR}(a,b,f) = E_1(a,b,f) - E_2(a,b,f) + \sigma(a,b,f).$$

Proof. We give a method of proof different from that of Cauchy's [1].

(1) In the first step, we note that in an interval $[a,b]$ such that $f(a)=f(b)$ and ($\forall x \in]a,b[$: $f(x)<f(a)$ or $f(x)>f(a)$),

$$E_1(a,b,f) - E_2(a,b,f) = 1.$$

An elementary use of the intermediate value theorem, of Rolle's theorem and Taylor's formula gives this result.

(2) In an interval where there is no root of the polynomial f we have:

$$E_1(a,b,f) - E_2(a,b,f) = 1 \quad \text{if} \quad f(a)f_1(a)f(b)f_1(b) < 0,$$
$$E_1(a,b,f) - E_2(a,b,f) = 0 \quad \text{if} \quad f(a)f_1(a)f(b)f_1(b) > 0.$$

(3) In an interval where the polynomial has a unique root, we write

$$E_1(a,b,f) - E_1(a,b,f)$$
$$= E_1(a,a_1,f) - E_2(a,a_1,f) + E_1(b_1,b,f) - E_2(b,b_1,f),$$

where the root belongs to the interval $[a_1,b_1]$ and f is strictly monotone. Then we use (2) for the different values of $\sigma(a,b,f)$ and we conclude as in Lemma 4.5. □

Lemma 4.6.

$$\mathrm{ZR}(a,b,f) = \mathrm{ZR}(a,b,f_1) + \sigma(a,b,f) - 2\alpha, \quad \textit{where } \alpha \in \mathbb{N}.$$

Proof. We introduce the two following notations: $I_1(a,b,f)$ is the number of roots of the polynomial f with horizontal tangent, $I_2(a,b,f)$ is the number of points on the curve that represent the polynomial f such that $f(x)\neq 0$ and $f_1(x)=0$ and that are not extrema.

We start the proof with:

$$\mathrm{NR}(a,b,f_1) = E_1(a,b,f) + E_2(a,b,f) + I_1(a,b,f) + I_2(a,b,f).$$

Then, by Lemma 4.5:

$$\mathrm{NR}(a,b,f) - \mathrm{NR}(a,b,f_1) = \sigma(a,b,f) - 2E_2(a,b,f) - I_1(a,b,f) - I_2(a,b,f).$$

On the other hand we observe that:

$$\begin{aligned}&\mathrm{ZR}(a,b,f)-\mathrm{ZR}(a,b,f_1)\\&\quad=\mathrm{NR}(a,b,f)+\sum(\alpha_k-1)-\mathrm{NR}(a,b,f_1)-\sum(\beta_k-1).\end{aligned}$$

We explain $\sum(\alpha_k-1)$ and $\sum(\beta_k-1)$. Now we write E_1, E_2, I_1, I_2 instead $E_1(a,b,f),\dots,$ etc.

$$\sum(\alpha_k-1)=\sum_{k=1}^{I_1}(\alpha_k-1),$$

where the α_k's are the orders of multiplicity of the roots of f in $[a,b]$.

$$\begin{aligned}\sum(\beta_k-1)&=\sum_{k=1}^{I_1}(\beta_k-1)+\sum_{k=I_1+1}^{I_1+I_2}(\beta_k-1)\\&\quad+\sum_{k=I_1+I_2+1}^{I_1+I_2+E_1}(\beta_k-1)+\sum_{k=I_1+I_2+E_1+1}^{I_1+I_2+E_1+E_2}(\beta_k-1),\end{aligned}$$

with:

(1) For $k=1,\dots,I_1$, β_k is the order of multiplicity of the roots of f and f_1 that verify:

$$f=f_1=\cdots=f_{2p}=0.$$

Then $\beta_k=\alpha_k-1$, which implies

$$\sum_{k=1}^{I_1}(\alpha_k-1)-\sum_{k=1}^{I_1}(\beta_k-1)=I_1.$$

(2) For $k=I_1+1,\dots,I_1+I_2$, β_k is the order of multiplicity of the root of f_1 that verifies:

$$f_1=\cdots=f_{2p}=0 \quad\text{and}\quad f\neq 0.$$

Then β_k is an even number, as $\sum_{k=I_1+1}^{I_1+I_2}\beta_k$.

(3) For $k=I_1+I_2+1,\dots,I_1+I_2+E_1+E_2$, β_k is the order of multiplicity of the root of f that verifies:

$$f_1=\cdots=f_{2p-1}=0 \quad\text{and}\quad f\neq 0.$$

Then β_k is an odd number and $\sum_{k=I_1+I_2+1}^{I_1+I_2+E_2+E_1}(\beta_k-1)$ is an even number.

Finally,

$$\begin{aligned}&\mathrm{ZR}(a,b,f)-\mathrm{ZR}(a,b,f_1)\\&\quad=\sigma(a,b,f)-2E_2-I_1-I_2+I_1-\sum_{k=I_1+1}^{I_1+I_2}\beta_k+I_2-\sum_{k=I_1+I_2+1}^{I_1+I_2+E_1+E_2}(\beta_k-1)\\&\quad=\sigma(a,b,f)-2E_2-\sum_{k=I_1+1}^{I_1+I_2}\beta_k-\sum_{k=I_1+I_2+1}^{I_1+I_2+E_1+E_2}(\beta_k-1)\\&\quad=\sigma(a,b)-\text{even number}.\end{aligned}$$

We denote this even number by

$$2\alpha = 2E_2 - \sum_{k=I_1+1}^{I_1+I_2} \beta_k - \sum_{k=I_1+I_2+1}^{I_1+I_2+E_1+E_2} (\beta_k - 1).$$

Lemma 4.7. *Let s, t, u, v be non-zero real numbers. Then*:

$$\mathrm{vP}\begin{pmatrix} u & v \\ -s & t \end{pmatrix} = \mathrm{vP}\begin{pmatrix} u & v \\ -s & t \end{pmatrix} \quad \text{if } uv > 0,$$

$$\mathrm{vP}\begin{pmatrix} u & v \\ -s & t \end{pmatrix} = 1 - \mathrm{vP}\begin{pmatrix} u & v \\ s & t \end{pmatrix} \quad \text{if } uv < 0.$$

Proof. Proving Lemma 4.7 is easy. □

Lemma 4.8.

$$\mathrm{vP}(a, f_1) - \mathrm{vP}(b, f_1) + \sigma(a, b, f)$$
$$= \mathrm{vP}(a, f) - \mathrm{vP}(b, f) - 2\gamma, \quad \text{where } \gamma \in \{0, 1\}.$$

Proof. We denote by

$$\mathrm{vP}_x\begin{pmatrix} g_1 \cdots g_n \\ h_1 \cdots h_n \end{pmatrix}$$

the number of variation–permanences in x of the sequence with polynomials g_i and h_i.

We investigate the different values of $\sigma(a, b, f)$.

Case 1. $\sigma(a, b, f) = 0$. Then $f(a) f_1(a) f(b) f_1(b) > 0$.
1.1. $f(a) f_1(a) > 0$ and $f(b) f_1(b) > 0$.

$$\mathrm{vP}(a, f) - \mathrm{vP}(b, f) = \mathrm{vP}_a\begin{pmatrix} f_1 \cdots f_n \\ F_1 \cdots F_n \end{pmatrix} - \mathrm{vP}_b\begin{pmatrix} f_1 \cdots f_n \\ F_1 \cdots F_n \end{pmatrix}.$$

1.1.1. $F_1(a) > 0$ and $F_1(b) > 0$. We write the sign of F_1 is f_1^2's in a and b. Then:

$$\mathrm{vP}(a, f) - \mathrm{vP}(b, f) = \mathrm{vP}(a, f_1) - \mathrm{vP}(b, f_1).$$

1.1.2. $F_1(a) > 0$ and $F_1(b) < 0$. $F_1(b)$ has the same sign as $-f_1^2(b)$. Since $F_1(b) < 0$ and $f(b) f_1(b) > 0$, it follows that $f_1(b) f_2(b) > 0$. Then

$$\mathrm{vP}(a, f) - \mathrm{vP}(b, f) = \mathrm{vP}(a, f_1) - \mathrm{vP}_b\begin{pmatrix} f_1 & \cdots f_n \\ -f_1^2 & \cdots F_n \end{pmatrix}.$$

Now,

$$\mathrm{vP}_b\begin{pmatrix} f_1 & \cdots f_n \\ -f_1^2 & \cdots F_n \end{pmatrix} = \mathrm{vP}_b\begin{pmatrix} f_1 & f_2 \\ -f_1^2 & F_2 \end{pmatrix} + \mathrm{vP}_b\begin{pmatrix} f_2 \cdots f_n \\ F_2 \cdots F_n \end{pmatrix}.$$

Since $f_1(b) f_2(b) > 0$ it follows from Lemma 4.7 that

$$\mathrm{vP}_b\begin{pmatrix} f_1 & f_2 \\ -f_1^2 & F_2 \end{pmatrix} = \mathrm{vP}_b\begin{pmatrix} f_1 & f_2 \\ f_1^2 & F_2 \end{pmatrix}.$$

We conclude

$$\mathrm{vP}(a,f) - \mathrm{vP}(b,f) = \mathrm{vP}(a,f_1) - \mathrm{vP}(b,f_1).$$

We prove the same result by the same methods when

$$(F_1(a) < 0 \text{ and } F_1(b) > 0) \quad \text{and} \quad (F_1(a) < 0 \text{ and } F_1(b) < 0).$$

1.2. $f(a)f_1(a)<0$ and $f(b)f_1(b)<0$.
1.2.1. $F_1(a)>0$ and $F_1(b)>0$.

$$\mathrm{vP}(a,f) - \mathrm{vP}(b,f) = 1 + \mathrm{vP}_a\begin{pmatrix} f_1\cdots f_n \\ F_1\cdots F_n \end{pmatrix} - 1 - \mathrm{vP}_b\begin{pmatrix} f_1\cdots f_n \\ F_1\cdots F_n \end{pmatrix}$$
$$= \mathrm{vP}_a(a,f_1) - \mathrm{vP}(b,f_1).$$

1.2.2. $F_1(a)>0$ and $F_1(b)<0$. Here $f_1(b)f_2(b)<0$.

$$\mathrm{vP}(a,f) - \mathrm{vP}(b,f) = 1 + \mathrm{vP}(a,f_1) - \mathrm{vP}_b\begin{pmatrix} f_1 & f_2\cdots f_n \\ -f_1^2 & F_2\cdots F_n \end{pmatrix}$$
$$= 1 + \mathrm{vP}(a,f_1) - \mathrm{vP}_b\begin{pmatrix} f_1 & f_2 \\ -f_1^2 & F_2 \end{pmatrix} - \mathrm{vP}_b\begin{pmatrix} f_2\cdots f_n \\ F_2\cdots F_n \end{pmatrix}.$$

From Lemma 4.7 and $f_1(b)f_2(b)<0$, it follows that:

$$\mathrm{vP}_b\begin{pmatrix} f_1 & f_2 \\ -f_1^2 & F_2 \end{pmatrix} = 1 - \mathrm{vP}_b\begin{pmatrix} f_1 & f_2 \\ f_1^2 & F_2 \end{pmatrix}.$$

We conclude

$$\mathrm{vP}(a,f) - \mathrm{vP}(b,f) = \mathrm{vP}(a,f_1) - \mathrm{vP}(b,f_1) + 2\mathrm{vP}_b\begin{pmatrix} f_1 & f_2 \\ f_1^2 & F_2 \end{pmatrix}.$$

1.2.3. ($F_1(a)<0$ and $F_1(b)>0$) or ($F_1(a)<0$ and $F_1(b)<0$). It is possible to select (r_0, r_1) such that $F_1(a)>0$ so as in Case 1.2.1 or 1.2.2. Indeed if $F_1(a)<0$, it follows that

$$r_1 < \frac{f(a)f_2(a)}{f_1^2(a)} r_0 \quad \text{and} \quad \frac{f(a)f_2(a)}{f_1^2(a)} \geq \frac{n-1}{n}.$$

Thus we consider (r_1, r_0) such that

$$r_1 > \frac{f(a)f_2(a)}{f_1^2(a)} r_0.$$

Then $F_1(a)>0$.

We can measure the technical importance of the choice of r_0 and r_1 and it is not in Sylvester's or Marchand's proof.

Case 2. $\sigma(a,b,f)=1$. Then $f(a)f_1(a)<0$ and $f(b)f_1(b)>0$. As before, we choose

(r_0, r_1) such that $F_1(a)>0$. We indicate the results in each case:

2.1. $F_1(a)>0$ and $F_1(b)>0$.

$$\mathrm{vP}(a,f)-\mathrm{vP}(b,f) = 1+\mathrm{vP}(a,f_1)-\mathrm{vP}(b,f_1).$$

2.2. $F_1(a)>0$ and $F_1(b)<0$. Here $f_1(b)f_2(b)<0$.

$$\mathrm{vP}(a,f)-\mathrm{vP}(b,f) = 1+\mathrm{vP}(a,f_1)-\mathrm{vP}(b,f_1)+2\mathrm{vP}_b\begin{pmatrix} f_1 & f_2 \\ f_1^2 & F_2 \end{pmatrix}.$$

Case 3. $\sigma(a,b,f)=-1$. Then $f(a)f_1(a)>0$ and $f(b)f_1(b)>0$. We obtain the same results as in Case 2 for the difference $\mathrm{vP}(a,f)-\mathrm{vP}(b,f)$ by substituting 1 by -1 in Cases 2.1 and 2.2. □

4.9. *Proof of Theorem 2.1.* It is easy to verify Theorem 2.1 for a polynomial of degree one. Suppose Theorem 2.1 is true for any polynomial of degree $n-1$. Let f be a polynomial f of degree n. For f_1 it follows that

$$\mathrm{ZR}(a,b,f_1) = \mathrm{vP}(a,f_1)-\mathrm{vP}(b,f_1)-2\beta, \quad \text{where } \beta\in\mathbb{N}.$$

From Lemmas 4.6 and 4.8 we conclude

$$\mathrm{ZR}(a,b,f) = \mathrm{vP}(a,f)-\mathrm{vP}(b,f)-2\gamma-2\alpha-2\beta. \qquad \square$$

Theorem 2.2 is proved in the same way by replacing Lemma 4.8 by:

Lemma 4.8′.

$$\mathrm{pP}(b,f_1)-\mathrm{pP}(a,f_1)+\sigma(a,b,f) = \mathrm{pP}(b,f)-\mathrm{pP}(a,f)-2\delta,$$
where $\delta\in\{0,1\}$.

To conclude, we add that Lemmas 4.5 and 4.6 hold for functions of class $C^m[a,b]$. Besides, Lemma 2 and its proof give us information on the kind of even numbers intervening in Descartes' rule, and in the theorems of Sylvester and Budan and Fourier. In fact, we have:

Lemma 4.9.

$$\mathrm{v}(a,f)-\mathrm{v}(b,f) = \mathrm{v}(a,f_1)-\mathrm{v}(b,f_1)+\sigma(a,b,f). \qquad \square$$

We can prove by induction the Bundan–Fourier theorem using Lemmas 4.6 and 4.9. The even number in this theorem is a function of number D_2 and I_2 of the functions f_k $(0\le n\le n)$. We have a different result of Schumaker [7] who establishes by induction:

$$\mathrm{ZR}(a,b,f) \le \mathrm{v}(a,f)-\mathrm{v}(b,f).$$

The fact that the difference between the two terms of the latter inequality is even is not considered by this author.

5. Applications

Application 5.1. Newton's rule can be used to prove the following theorem:

Theorem 5.2. *Let r_0 and r_1 such that $r_0 \leq ((n-1)/n)r_0$. We suppose that the polynomial*

$$f(x) = \sum_{k=0}^{n} a_k x^{n-k} \quad (a_0 \neq 0 \text{ and } a_n \neq 0)$$

possesses only real roots. Then the inequalities

$$\frac{r_k}{r_{k-1}} \frac{n-k}{n-k+1} a_k^2 - a_{k-1} a_{k+1} \geq 0, \quad 1 \leq k \leq n-1,$$

hold.

Proof. The computation of F_{n-k}'s at $x=0$ gives

$$F_{n-k} = r_k((n-k)!)^2 (n-k)!\,(n-k+1)! \left[\frac{r_k}{r_{k-1}} \frac{n-k}{n-k+1} a_k^2 - a_{k-1} a_{k+1} \right].$$

It is clear that if the polynomial has all of its roots positive, then the sequence of a_k's has n variations and the sequence of F_k's has n permanences. Since F_0 and F_n are positive, the inequalities hold.

We give the method of proof in the case where the polynomial f possesses n_1 positive roots and $-f$ possesses n_2 negative roots with $n = n_1 + n_2$. Then $\text{vP}(0, f) = n_1$ and $\text{vP}(0, -f) = n_2$.

The coefficient a_k (resp. $(-1)^k a_k$) can be divided into sets such that (1) in each set the a_k (resp. $(-1)^k a_k$) have the same sign and are not all zero, (2) the a_k's of two consecutive sets have opposite signs. If $a_k a_{k+1} < 0$, we have $F_k F_{k+1} > 0$.

An elementary observation of the members of each set of coefficients a_k and the fact that F_0 and F_n are positive, permit us to conclude. □

Application 5.3. Let P_n the set of polynomials such that $f(x) = \sum_{k=0}^{n} a_k x^k$ with $a_0 = 1$ and $a_k \in \{-1, 1\}$ for $1 \leq k \leq n$.

We compute the average number of variation-permanences at zero of a polynomial belonging to P_n. If we denote this number by vP_n, and by vP_{nk} the number of polynomials belonging to P_n which have k variation-permanences, then

$$\text{vP}_n = \frac{1}{2^n} \sum_{k=0}^{n} \text{vP}_{nk}.$$

The computation of the average number of variations at zero for a polynomial belonging to P_n gives $\frac{1}{2}n$.

We recall that the average number of roots for polynomials belonging to P_n is equal to $(2/\pi) \log n$. See [3].

Theorem 5.4. *The average number* vP_n *of variation-permanences at zero for polynomials belonging to* P_n *is equal to* $\frac{1}{16}(n+1)-\frac{1}{32}$.

Preliminary to the proof: Let $x=0$. We denote g for $g(0)$. Since $f_k=k!\,a_k$, we obtain:

$$p_k=\begin{cases}2, & \text{if } a_{k-1}a_{k+1}=1,\\ \dfrac{2k^2-2k-2}{2k^2-1}, & \text{if } a_{k-1}a_{k+1}=-1.\end{cases}$$

It is easy to prove that $F_k<(n-1)/n$ if $a_{k-1}a_{k+1}=-1$. We consider $r_0=1$ and $r_1=2$. Furthermore,

$$F_k=\begin{cases}0, & \text{if } a_{k-1}a_{k+1}=1,\\ (k-1)!^2(2k^2+2k), & \text{if } a_{k-1}a_{k+1}=-1.\end{cases}$$

By Proposition 3.8, the sign of F_k is the same as that of $a_k a_{k+m}$ where m is the least integer such that $F_{k+m}\neq 0$. Next we observe that a polynomial belonging to P_n has a variation-permanence iff

for $k=1: a_0=a_1$ and $a_2=-1$,
for $3\le k\le n-2: a_{k-2}=a_{k-1}=-a_k=-a_{k+1}$,
for $k=n: a_{n-2}=a_{n-1}=-a_n$.

Proof. We do not compute the numbers VP_{nk}. By the preliminary, we write $\sum_{k=0}^{n} k\,\mathrm{VP}_{nk}$ as the sum of three quantities:

(1) The number of polynomials belonging to P_n such that $a_{n-2}=a_{n-1}=-a_n$. It is equal to 2^{n-3}.

(2) The number of polynomials belonging to P_n such that there exists a k, $3\le k\le n-2$, such that $a_{k-2}=a_{k-1}=-a_k=-a_{k+1}$ and $a_{n-1}=a_n$. We obtain $\sum_{k=3}^{n-3}2^2 2^{n-6}+2\cdot 2^{n-6}=2^{n-5}(2n-9)$.

(3) The number of polynomials belonging to P_n such that: there does not exist k, $3\le k\le n-2$, such that $a_{k-2}=a_{k-1}=-a_k=-a_{k+1}$ and that the property $a_{n-2}=a_{n-1}=a_n$ does not hold. This number is equal to: $6\cdot 2^{n-5}$.

Furthermore,

$$\mathrm{VP}_n=(1/2^n)(2^{n-2}+2^{n-5}(2n-9)+6(2n-9)+6(2^{n-4}))=\tfrac{1}{16}(n+1)-\tfrac{1}{32}.$$

6. Conclusion

The study of the complexity is based on the work of Coste-Roy and Szpirglas [2].

Let $N(f)=(\sum_{k=0}^{n}a_k^2)^{1/2}$. It is known that the complexity of Sturm's method is at $O(n^4\log^2 N(f))$.

The computation of $\mathrm{VP}(0,f)+\mathrm{vP}(0,-f)$ requires the computation of F_k's and the sorting of p_k's defined in 3.4. The arithmetical operations are $O(n)$ and the sorting of p_k's $O(n \log n)$. Taking this into consideration, we obtain a complexity to computation at $O(n \log N(f)(\log N(f)+\log n))$.

Appendix

A.1. Proof of Proposition 3.7. The derivative of order j of F_{k+i} in x $(0 \le i \le m-1)$, is equal to:

$$F_{k+i}^{(j)}=\sum_{l=0}^{J} r_{k+i}\binom{j}{l} f_{k+i+l} f_{k+i+j-l}-\sum_{i=0}^{J} r_{k-1+i}\binom{j}{l} f_{k+i-1+l} f_{k+i+1+j-l}.$$

For $i=0$, it follows from the hypothesis of the proposition that

$$F_k^{(j)}=0, \quad j=0, \ldots, m-2,$$

and

$$F_k^{(m-1)}=-r_{k-1} f_{k-1} f_{k+m}.$$

We conclude from Taylor's formula the relation for F_k.

On the other hand, for $i \neq 0$ and $j=0, \ldots, 2m-2i-1$,

$$F_{k+i}^{(j)}=0.$$

For $j=2m-2i$, we obtain:

$$F_{k+i}^{(2m-2i)}=\frac{(2m-2i)!}{(m-i)!^2}\left(r_{k+i}-\frac{m-i}{m-i+1} r_{k+i-1}\right) f_{k+m}^2.$$

Since

$$r_{k+i}-\frac{m-i}{m-i+1} r_{k+i-1}=\frac{1}{m-i+1} r_{k+m},$$

if follows that

$$F_{k+i}^{(2m-i)}=\frac{(2m-2i)!}{(m-i)!^2(m-i+1)} f_{k+m} f_{k+m}^2.$$

The same calculation gives for $j=2m-2i+1$,

$$F_{k+i}^{(2m-2i+1)}=\frac{(2m-2i)!}{(m-i)!^2(m-i)(m-i+1)^2(m-i+2)} f_{k+m+1} f_{k+m-1} f_{k+m+1}.$$

Hence we have the second relation by applying Taylor's formula to the F_{k+i}'s. □

A.2. Proof of Proposition 3.8. We calculate F_k':

$$\begin{aligned} F_k' &= (2r_k-r_{k-1}) f_k f_{k+1}-r_{k-1} f_{k+1} f_{k-1} \\ &= r_{k+1} f_k f_{k+1}-r_{k-1} f_{k+1} f_{k-1}. \end{aligned}$$

Now we use the definition of F_k and obtain

$$\begin{aligned} f_{k+1}F_k' &= r_{k+1}f_kf_{k+1}^2 - r_kf_k^2f_{k+2} + F_kf_{k+2} \\ &= f_k F_{k+1} + f_{k+2} F_k. \end{aligned}$$

It follows inductively that

$$f_{j+k}F_k^{(j)} = \sum_{l=0}^{j-1} g_{kl}F_{k+l} + f_kF_{k+j}.$$

From the hypothesis of Proposition 3.8, we conclude

$$\begin{aligned} &F_k^{(j)} = 0, \quad j = 0, \dots, m-1, \\ &f_{k+m}F_k^{(m)} = f_kF_{k+m}. \end{aligned}$$

Hence Proposition 3.8 follows. □

Acknowledgment

The author would like to thank M.F. Coste-Roy and L. Gonzales for useful comments.

References

[1] A. Cauchy, Mémoire sur la détermination du nombre de racines réelles dans les équations algébriques (1813), in: Oeuvres II ème série, Tome 1 (Gauthiers-Villars, Paris, 1877).

[2] M.F. Coste-Roy and A. Szpirglas, Complexity of the computation on real algebraic numbers, J. Symbolic Computation, to appear.

[3] M. Kac, Probability and Related Topics in Physical Sciences (Interscience, London, 1959) 5–10.

[4] S. Karlin, Total Positivity (Stanford Univ. Press, Stanford, CA, 1968).

[5] E. Marchand, Sur les théorèmes de Sylvester et la règle de Newton, Thèse, Neuchâtel, 1913.

[6] A. Ostrowski, Bemerkungen zum beweise des Budan–Fourierchen und Newton–Sylvesterschen Satzes, in: Collected Mathematical Papers, Vol. 1 (Birkhäuser, Basel, 1983).

[7] L. Schumaker, Spline Functions. Basic Theory, Pure and Applied Mathematics (Wiley, Chichester, 1981).

[8] J.J. Sylvester, On an elementary proof and generalization of Sir Isaac Newton's hitherto undemonstrated rule for the discovery of imaginary roots (1865), in: Collected Mathematical Papers, Vol. 2 (Chelsea, New York, 1973).