Relationship Between High-Resolution Methods and Discrete Fourier Transform

S.MAYRARGUE and Th.BLU

CNET/PAB/RPE
38-40 rue du Général Leclerc
92131 Issy-les-Moulineaux (France)

ABSTRACT : In this paper, we establish a link between the Discrete Fourier Transform (DFT) and 2 high-resolution methods , MUSIC and the Tufts and Kumaresan's method (TK). Existence and location of the extraneous peaks of MUSIC, of the "noise" zeros of TK , are related to the minima of the DFT of the rectangular window filtering the data. Other properties of the "noise" zeros are given, in relation to polynomial theory.

I- INTRODUCTION

Time series harmonic analysis is a crucial problem in a number of practical situations. However, very often, data lengths are short, so that the separation between the frequencies to be retrieved may be shorter than the Fourier resolution limit. In this case, the so-called "high-resolution" methods are needed. Various such methods have been devised. In this paper, we will consider the MUSIC method [1], and Tufts and Kumaresan's method (TK) [2] [3]. In both methods, a "model order" is chosen.It corresponds to a rectangular windowing of the data.The sources are obtained as peaks in a spectrum (MUSIC) or zeros of a polynomial (TK).The number of such peaks (resp. zeros) is equal to the model order.The order is generally chosen in order to overestimate the number of frequencies, so that a number of extraneous peaks (resp. zeros) are introduced. These are generally attributed to the noise. We show that they originate in fact in the shape of the spectrum of the above-mentioned rectangular window.

We first show that MUSIC can be viewed as a special case of the periodogram, this latter being understood as the average spectrum of the data, taken over overlapping rectangular windows of width equal to the model order.The secondary peaks of the MUSIC spectrum are thus related to the minima of the Fourier transform of this rectangular window.

The TK method relies on the properties of the minimum-norm vector with first coefficient equal to 1 belonging to the so-called "noise subspace". It is well-known that the zeros of the associated polynomial can be divided into two subsets, namely the "source" zeros and the "noise" zeros.The "noise" zeros have been observed to be approximatively uniformly distributed on a circle centered in 0, and of radius <1.

This fact has never been explained, though [4] performs an explicit computation of the location of the roots in the similar case of the maximum-entropy method, but only in the single frequency case.

On another hand, [5] considers this problem from the point of view of measure theory. A probability measure is introduced, which assigns the same weight to each "noise" root. For the 2 frequencies case, [5] proves that for almost every value of the difference between these two frequencies, this measure converges weakly to the uniform measure on the unit circle when the model order tends to infinity.

In this paper, we present results on the structure of the "noise" polynomial, which leads to explain properties of their zeros. We then can relate the location of the zeros to the shape of the Fourier transform of the above-mentioned rectangular window.

We first show that the "noise" zeros are in fact roots of a lacunary polynomial. The properties of lacunary polynomials give an explanation to the distribution of the zeros.

We can then prove that the "noise" roots converge uniformly to the unit circle, except possibly for m-1 of them, m being the number of frequencies.The proof will not be given here, due to lack of space.

We study in detail the 1 frequency case, giving a direct proof of the convergence and the angular "equi"distribution of the roots.

For any number of frequencies, we give a first-order approximation of the lacunary polynomial, leading to prove the convergence of the roots to the unit circle.An explicit expression of this polynomial in the 2 frequencies case is given.

Finally, we give an heuristic explanation to the angular distribution of the zeros: since the "noise" zeros lie near the unit circle, the spectrum of the coefficients of the lacunary polynomial can be expected to have minima close to the location of these zeros. Using the above-mentioned first-order approximation , it is possible to prove that these minima are approximately angularly equispaced. The zeros are then likely to be also so. On another hand we have proven that the coefficients of the noise polynomial behaved more or less like the sum of the signal itself plus its derivative. The angular equispacing of the minima of its spectrum then corresponds to the minima of the spectrum of the rectangular window of width equal to the model order.

II- Problem Position

Throughout this paper, only noiseless data will be considered.

Let x be a signal composed of a sum of m complex sine waves, observed on N samples.

$$x_k = \sum_{i=1}^{m} d_i \exp[j\ \omega_i k]\ , \qquad \text{for } k = 0, \dots, N\text{-}1 \qquad (1)$$

where ω_i , d_i are resp. the pulsations and complex amplitudes of the i-th complex exponential.

III- Periodogram

In this paragraph, we recall the definition and properties of the periodogram.

The periodogram is commonly defined as the squared modulus of the discrete Fourier transform of the data. However, very often, overlapping windowing of the data is introduced and an averaging over the periodograms of the windowed data is used [6] .This can be written in matrix form as follows. Letting

CH2977-7/91/0000-3321 $1.00 © 1991 IEEE

$$X = \begin{bmatrix} x_0 & x_1 & & x_{N-L} \\ x_1 & x_2 & & x_{N-L+1} \\ & & & \\ x_{L-1} & & & x_{N-1} \end{bmatrix}$$

and $P(\omega)$ denote the averaged periodogram, we have $P(\omega) =$

$$1/(N-L+1)\,[1, e^{-j\omega}, \dots, e^{-j(L-1)\omega}]\; XX^+ [1, e^{-j\omega}, \dots, e^{-j(L-1)\omega}]^+$$

where the superscript + denotes conjugate transpose.

Let us investigate the shape of $P(\omega)$. Using (1) we obtain :

$$X = \begin{bmatrix} 1 & \dots & 1 \\ e^{j\omega_1} & & e^{j\omega_m} \\ & & \\ e^{j(K-1)\omega_1} & & e^{j(K-1)\omega_m} \end{bmatrix} A = S_L A$$

where A is the matrix containing the amplitudes of the windowed signals .

Thus $\quad XX^+ = S_L AA^+ S_L^+ \qquad (2)$

Note that A can be expressed as follows :

$$A = \begin{bmatrix} d_1 & & \\ & & d_m \end{bmatrix} \begin{bmatrix} 1 & e^{j\omega_1} & e^{j(N-L)\omega_1} \\ 1 & e^{j\omega_m} & e^{j(N-L)\omega_m} \end{bmatrix}$$

and thus that $1/(N-L+1)\, XX^+$ tends towards $S_L |D|^2 S_L^+$ when N tends to infinity, where $|D|^2 = \mathrm{diag}(|d_1|^2, \dots, |d_m|^2)$.

If m = 1, we clearly have $XX^+ =$

$(N-L+1)\, d_1^2\, [1, e^{-j\omega_1}, \dots, e^{-j(L-1)\omega_1}]^+ [1, e^{-j\omega_1}, \dots, e^{-j(L-1)\omega_1}]$

so that :

$P(\omega) = d_1^2 \left(\sin((\omega_1-\omega)L/2) / \sin(\omega_1-\omega)/2 \right)^2$

$P(\omega)$ has the well-known shape presenting a global maximum equal to d_1^2 and L-1 secondary maxima.

For well-separated sources, or for large N, we have

$P(\omega) \cong \Sigma\; d_i^2 \left(\sin((\omega_i-\omega)L/2) / \sin(\omega_i-\omega)/2 \right)^2$

It is then approximately true that this fonction admits L-m secondary lobes, L being the width of the rectangular windowing of the data.

IV- MUSIC

MUSIC is a high-resolution method, which we want to relate to the periodogram.

MUSIC relies on the eigendecomposition of XX^+. Indeed, let $XX^+ = U\Lambda U^+$, U (resp.Λ) being the matrix of the orthonormal eigenvectors (resp. eigenvalues) of XX^+. According to (2), XX^+ has rank m, so that only m eigenvalues are non-zero.Let Λ_s be the diagonal matrix containing these eigenvalues, and U_s be the matrix of the associated eigenvectors. We can write

$$XX^+ = U_s \Lambda_s U_s \qquad (3)$$

The columns of XX^+ span the same subspace as those of U_s, and also as those of S_L, according to (2) . This subspace is termed the "signal subspace".

The MUSIC method consists in seeking the values of ω which maximize the projection of the "steering" vector $[1, e^{j\omega}, \dots, e^{j(L-1)\omega}]$ on the "signal" subspace. Indeed, the steering vectors corresponding to the $\{\omega_i\}$ belong to the signal subspace.Since the projector on the signal subspace is

$\Pi_s = U_s U_s^+$, MUSIC in matrix form is given by

max $f(\omega)$, where

$f(\omega) = [1, e^{-j\omega}, \dots, e^{-j(L-1)\omega}]\, U_s U_s^+ [1, e^{-j\omega}, \dots, e^{-j(L-1)\omega}]^+$

Recall that $P(\omega)$ is proportional to :

$[1, e^{-j\omega}, \dots, e^{-j(L-1)\omega}]\, U_s \Lambda_s U_s^+ [1, e^{-j\omega}, \dots, e^{-j(L-1)\omega}]^+$

. For the special case of 1 frequency , $U_s = S_L/\sqrt{L}$ therefore XX^+ is proportional to $U_s U_s^+$. Thus, in this case, $f(\omega)$ and $P(\omega)$ are proportional.

. For any number of sources, we can write $U_s = S_L P$, where P is a non-singular square mxm matrix. This amounts only to express that U_s and S_L span the same subspace. Thus

$$P_s = U_s U_s^+ = S_L PP^+ S_L^+$$

and

$$f(\omega) = [1, \dots, e^{-j(L-1)\omega}]\, S_L (PP^+) S_L^+ [1, \dots, e^{-j(L-1)\omega}]^+$$

Recall that the periodogram was given by $P(\omega)$, with $P(\omega)$ proportional to $[1, \dots, e^{-j(L-1)\omega}]\, S_L (AA^+) S_L^+ [1, \dots, e^{-j(L-1)\omega}]^+$ we can thus consider $f(\omega)$ as the "periodogram" of a special signal whose amplitudes would be characterized by P instead of A. This explains why the MUSIC spectrum exhibits regularly spaced secondary peaks : they originate in the implicit windowing of the data, as it is the case for the periodogram.

Note that for well separated frequencies $AA^+ \cong |D|^2$ and $PP^+ \cong Id$. Then $f(\omega)$ tends to be proportional to $P(\omega)$.

V- Tufts and Kumaresan's method

In this paragraph, we first recall the principle of the TK method, and then relate the "noise" roots to the periodogram of the data.

V.1. Principle of the method

We defined in the above section the "signal" subspace as the one spanned by S_L. The subspace which is orthogonal to the "signal" subspace will be termed "noise" subspace. Any vector of the "noise" subspace is thus orthogonal to S_L. Consider a polynomial having as coefficients the elements of such a vector. The special structure of the columns of S_L implies that the $\{e^{j\omega_i}$ $i=1, \dots, m\}$ belong to the roots of this polynomial. There is an infinity of such polynomials. TK method singles out the minimum-norm vector belonging to the noise subspace, with first coefficient equal to 1 . The associated polynomial has thus L-m-1 roots besides the $\{e^{j\omega_i}\}$. These are termed the "noise" roots.

V.2 Problem position

It has been observed that the L-m-1 "noise" roots of the associated polynomial behave "regularly" in some sense : they seem to be approximately angularly equidistributed on a circle concentric to the unit circle, and of radius <1. This property facilitates the separation between these zeros and the $\{e^{j\omega_i}\}$.

However, while it has long been proven [3] that the noise roots are located inside the unit circle, their other properties have never been given an explanation, except by [4] for m =1 in the context of the maximum entropy method which leads to polynomials having similar properties.

In another respect, we have proven that the noise roots converge uniformly to the unit circle, except for at most m-1 of them (we in fact give an upper bound of the distance of those former roots to the unit circle). The proof will not be given here, due to lack of space.

We are going to study the properties of the "noise" polynomial coefficients (polynomial whose roots are the "noise" roots) . They originate in the following fact [3]. The coefficient vector of the noise polynomial is the AR prediction-error filter for a data sequence formed by the coefficients of the "signal" polynomial surrounded by zeros.

We establish that the "noise" roots are also roots of a lacunary polynomial. This latter property gives the key to the understanding of the location of these zeros.

We then put forth a complete study of the case $m = 1$.

For any number of sources, we give a first-order approximation of the lacunary polynomial.

For $m = 2$, we give an explicit expression for the lacunary polynomial.

Finally, we give an heuristic explanation to the almost equiangular distribution of the zeros, noting that since the zeros lie mostly near the unit circle, the spectrum of the noise polynomial coefficients can be expected to exhibit minima close to the location of these zeros. We have shown that the coefficients of this polynomial were the sum of two terms, one of them a linear sum of the same sine waves as the signal itself, the other one a linear sum of their derivatives. This hints that the spectrum of these coefficients has minima distributed as those of the spectrum of the rectangular window of width equal to the model order. This latter result can be proven using the first-order approximation of the lacunary polynomial.

V.3 The "noise" polynomial coefficients viewed as the AR solution to a special problem

Seeking the minimum-norm vector with first coefficient equal to 1 belonging to the "noise" subspace is equivalent to seeking a (L-m-1)-th degree monic polynomial with m given zeros and minimum-norm. The product of two polynomials can be written in matrix form as follows :

$$Ab = \begin{bmatrix} 1 & & 0 \\ \alpha_1 & 1 & \\ & \alpha_1 & \\ & & 1 \\ & & \alpha_1 \\ \alpha_m & & \\ & \alpha_m & \\ 0 & & \\ 0 & & \alpha_m \end{bmatrix} \begin{bmatrix} 1 = b_{L-m-1} \\ b_{L-m-2} \\ \\ \\ b_0 \end{bmatrix} = \begin{bmatrix} 1 = v_{L-1} \\ \\ \\ \\ v_1 \\ v_0 \end{bmatrix}$$

where $\alpha_1, \dots, \alpha_m$ (resp. $b_0, \dots, b_{L-m}$) are the coefficients of the "signal" (resp."noise") polynomial $A(z)$ (resp. $B_L(z)$). Assuming $\alpha_0 = 1$ involves no loss in generality.

The minimum-norm condition gives : b^+A^+Ab minimum under the constraint $b_0=1$. The solution to this problem is given by : A^+Ab proportional to $[1, 0, \dots, 0]^T$, i.e. b is the AR solution with respect to a signal formed by the $\{\alpha_i\ i=1, \dots, m\}$ surrounded by zeros [3].

V.4. A property of the coefficients of the "noise" polynomial

A^+A is a Toeplitz band matrix. Indeed

$$A^+A = \begin{bmatrix} r_0 & r_1 & & r_m & 0 & & 0 \\ r_1 & r_0 & r_1 & & & r_m & 0 \\ r_m & & & & & & r_m \\ & & & & & & r_1 \\ 0 & & & r_m & & & r_0 \end{bmatrix}$$

Therefore, since A^+Ab is proportional to $[1, 0, \dots, 0]^T$, the data sequence $\{\dots, 0, 0, 0, b_0, b_1, \dots, b_{L-m-1}, 0, \dots, 0, \dots\}$ follows a linear recursion relationship with coefficients $\{r_i\}$. It is clearly seen that since the zeros of the "signal" polynomial have modulus equal to 1, the $\{r_i\}$ are the coefficients of $(A(z))^2$.

Therefore, we can write

$$b_n = \sum_{k=1}^{m} (\lambda_k + n\,\mu_k)\, e^{j n \omega_k} \qquad (3)$$

The 2m values λ_k and μ_k are determined by writing that $b_{-1}, \dots, b_{-(m-1)}, b_{L-m}, \dots, b_{L-1}$ are zero and $b_{L-m-1} = 1$

V.5 A lacunary polynomial

A^+Ab does not represent the multiplication of the polynomial $\{r_i\}$ by the noise polynomial. This latter multiplication is written in matrix form using the augmented matrix $(A^+A)^{aug}$, so that

$$(A^+A)^{aug}\, b = \begin{bmatrix} r_m & & & & \\ r_1 & r_m & & & \\ r_0 & r_1 & r_m & 0 & 0 \\ r_1 & r_0 & r_1 & r_m & 0 \\ r_m & & & & r_m \\ & & & & r_1 \\ 0 & & r_m & & r_0 \\ & & & & r_m \end{bmatrix} \begin{bmatrix} 1 = b_{L-m-1} \\ \\ \\ b_1 \\ b_0 \end{bmatrix} = \begin{bmatrix} x \\ x \\ 1 \\ 0 \\ \vdots \\ 0 \\ x \\ x \\ x \end{bmatrix}$$

where the symbol x denotes a non-significant value.

This means that $(A(z))^2\, B_L(z)$ is a lacunary polynomial, i.e. of the following type $P_1(z)+z^{L-1}P_2(z)$ with $\deg(P_1)\le m-1$ and $\deg(P_2)\le m$, the polynomials depending implicitely on the model order L.

V.6 Case m = 1

We give here an exhaustive study of the location of the roots. Taking $\omega_1 = 0$ involves no loss in generality.

Using (3), we obtain $b_n = (\lambda + n\,\mu)$. We have $b_{-1} = 0$, thus $\lambda = \mu$. We also have $b_{L-m-1} = 1$. Thus $\lambda = 1/(L-m)$. Consequently

$$B_L(z) = 1/(L-m)\,[\Sigma(1+k)\, z^k]$$
$$= 1/(L-m)\ d/dz\,[(z^{L-m+1} - 1)/(z-1)]$$
$$= 1/(L-m)[(L-m)z^{L-m+1}-(L-m+1)z^{L-m}+1]/(z-1)^2$$

Following [4], [6], we can assert that the "noise" roots ξ_i are such that

$$(2(L-m))^{-1/(L-m)} \le |\xi_i| \le (L-m-1)/(L-m)$$

(See appendix for a proof).

Moreover, ([7] p.165)) implies that each root lies within the sector $(2k+1)\pi/(L-m+2) \pm \pi/(L-m+2)$ for $k=0..(L-m)$.

Therefore, we have proven that for m=1 the noise zeros converge to the unit circle when L tends to infinity, while being approximately angularly distributed.

V.7 General case

Let us write

$$A(z)^2 B_L(z) = 1/L\ Q_L(z) + z^{L-1}(1/L\ R_L(z) + A(z))$$

There, letting $\zeta_n = e^{j\omega_n}$, we write, using Lagrange's interpolation polynomials

$$\begin{cases} Q_L(z) = \displaystyle\sum_{n=1}^{m} \frac{Q_L(\zeta_n)}{A'(\zeta_n)} \frac{A(z)}{z-\zeta_n} \\ R_L(z) = \displaystyle\sum_{n=1}^{m} \frac{R_L(\zeta_n)}{A'(\zeta_n)} \frac{A(z)}{z-\zeta_n} \end{cases}$$

and thus, since ζ_n are 2-fold zeros of the lacunary polynomial :

$$\begin{cases} Q_L(\zeta_n)+\zeta_n^{L-1}R_L(\zeta_n)=0 \\ \sum_{k\neq n} R_L(\zeta_k)\frac{A'(\zeta_n)}{A'(\zeta_k)}\frac{\zeta_n^{L-1}-\zeta_k^{L-1}}{\zeta_n-\zeta_k} \\ +(L-1)\zeta_n^{L-2}R_L(\zeta_n)+L\zeta_n^{L-1}A'(\zeta_n)=0 \end{cases}$$

for n=1..m. The second equation implies that $R_L(\zeta_n)$ is bounded over L. Therefore, $R_L(z)$ and $Q_L(z)$ are two polynomials whose coefficients are bounded over L, wherefrom the Kumaresan polynomial $V_L(z)$ follows

$$V_L(z) = A(z)B_L(z)$$

$$= z^{L-1}+\frac{1}{L}\frac{Q_L(z)+z^{L-1}R_L(z)}{A(z)}$$

This is a useful property which shows in the limit L tends to infinity that, when its modulus is evaluated over the unit circle, the Kumaresan polynomial behaves as a constant except near the roots of A. More precisely,

$$\left|V_L(e^{i\theta})\right|^2 \underset{L\to\infty}{\cong} 1+\frac{2}{L}\Re\left(\frac{R_L(e^{i\theta})}{A(e^{i\theta})}\right)+\frac{2}{L}\Re\left(e^{-(L-1)i\theta}\frac{Q_L(e^{i\theta})}{A(e^{i\theta})}\right)$$

Thus, away from the $\{e^{j\omega_i}\}$, the first (constant) term dominates, while the second provides for a slowly varying 1/L correction and the third a fastly varying one, behaving like the Fourier transform of the rectangular window of width L-1.

V. 8 Case m = 2

Note first that chosing $\omega_1=-\omega_2=\theta$ involves no loss of genarality. We now solve the system of equations leading to the 4 coefficients of $R_L(z)$ and $Q_L(z)$ and obtain

$$\begin{cases} R_L(z)=\frac{1-S(\theta)\cos L\theta}{D(\theta)}+z\frac{S(\theta)\cos[(L-1)\theta]-\cos\theta}{D(\theta)} \\ Q_L(z)=\frac{S(\theta)\cos\theta-\cos[(L-1)\theta]}{D(\theta)}+z\frac{\cos L\theta-S(\theta)}{D(\theta)} \end{cases}$$

where $S(\theta)$ and $D(\theta)$ are defined as follows

$$\begin{cases} S(\theta)=\frac{\sin[(L-1)\theta]}{(L-1)\sin\theta} \\ D(\theta)=\frac{L-1}{2}(1-S(\theta)^2) \end{cases}$$

APPENDIX

1. We have

$B_L(z) = 1/(L-m)\, d/dz\, [(z^{L-m+1} - 1)/(z-1)]$

$= 1/(L-m)[\, 1+2z+ \ldots\ldots + (L-m)z^{L-m-1}\,]$

Using a theorem of [7] (p.137) about polynomials with positive coefficients, we find that $|\xi_i| \leq (L-m-1)/(L-m)$.

2. Let $N_L(z)=(L-m)z^{L-m+1}-(L-m+1)z^{L-m}+1$.

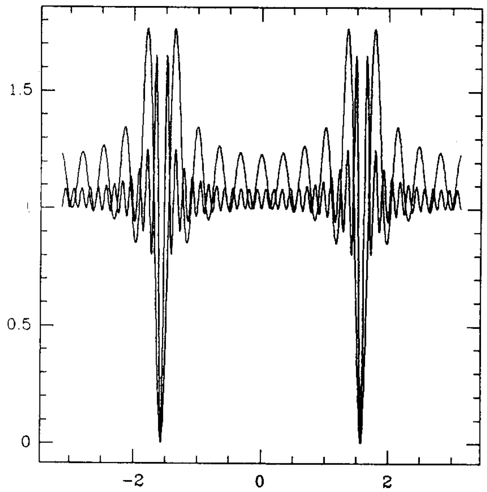
Consider now $z^{L-m+1}N_L(1/z)$, which we will denote by $M_L(z)$. $M_L(z) = (L-m) - z(L-m+1) + z^{L-m+1}$. The roots of $M_L(z)$ are $1/\xi_i$

[7] (Theorem 27.1) leads to seek the positive root of $U_L(z)$, with $U_L(z) = z^{L-m+1} - z(L-m+1)-(L-m)$.

Since $U_L(0)=-(L-m)$

and $U_L((2(L-m))^{1/(L-m)})$ can be checked to be positive, [7] asserts that

$1/|\xi_i| \leq 2(L-m)^{1/(L-m)}$



Spectrum $|V_L(e^{j\varphi})|^2$ of V_L (the minimum-norm TK vector) φ varying from $-\pi$ to $+\pi$, for 2 sources at $\pm e^{j\pi/2}$ and for L = 20, L=50 .

VI. Conclusion

We have established a link between the shape of the periodogram and, on the one hand, the location of the peaks in the MUSIC spectrum, on the other hand the location of the "noise" zeros in the TK method. In this latter case, the zeros are proven to root a lacunary polynomial. Further work will use this property to obtain a better understanding of the behavior of the roots. A future paper will present the proof of the convergence of the roots to the unit circle (except for a finite number of them) when the model order tends to infinity.

REFERENCES

[1] R.SCHMIDT Multiple Emitter Location and Signal Parameter Estimation. Proc. RADC Spectral Estimation Workshop (Rome, NY, 1979) pp. 243-256

[2] D.W.TUFTS, R.KUMARESAN Estimation of Frequencies of Multiple Sinusoids : Making Linear Prediction Perform Like Maximum Likelihood. Proc. of the IEEE, Vol.70, N°9, Sept.1982

[3] R.KUMARESAN Estimating the Parameters of Exponentially Damped or Undamped Sinusoidal Signals in Noise Ph.D. Thesis, University of Rhode Island, Kingston. 1982.

[4] J.L.LACOUME, CH. HANNA, J.L.NICOLAS Etalonnage de l'ANALYSE Specrtale par la Methode du Modèle Autoregressif Annales des Télécommunications,3 6, n°11-12, 1981.

[5] L.PAKULA Asymptotic Zero Distribution of Orthogonal Polynomials in Sinusoidal Frequency Estimation IEEE Trans. on Information Theory, Vol IT-33, No4, July 1987.

[6] S.KAY, L.MARPLE Spectrum Analysis- A modern Pespective. Proc. of the IEEE Vol.69, N°11, Nov.1981.

[7] M.MARDEN Geometry of Polynomials . Amercan Mathematical Society. Providence, Rhode Island 1966.