Equations algébriques et fonctions symétriques

JP Chabert

Table des matières

1 Introduction **1**

2 Equations du second degré **2**

3 Fonctions symétriques **5**

4 Matrices et déterminants de Vandermonde **15**

5 Equations algébriques de degré quelconque **22**

6 Formules de Newton-Girard et de Brioschi **29**

1 Introduction

En étudiant la suite de Fibonacci, on remarque son lien étroit avec l'équation : $x^2 = x+1$, dont les racines sont les deux nombres d'or Φ_1 et Φ_2, et avec la matrice : $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, dont les valeurs propres sont également les deux nombres d'or. Ces remarques sont riches de conséquences.

Nous allons chercher à savoir dans quelle mesure ces propriétés se généralisent ; pour cela, nous allons commencer par les équations du second degré, de la forme : $x^2 = ax+b$; ensuite nous ferons une incursion du côté des fonctions symétriques et des matrices de Vandermonde, avant de tenter une généralisation aux équations de degré supérieur à 2.

Par conséquent, on ne doit pas s'attendre à trouver ici un résumé des connaissances actuelles sur les équations algébriques, sur les méthodes de résolution formelles (résolution par radicaux) ou numériques (méthodes dichotomique, de fausse position, de Newton-Raphson, de Laguerre, de Ridder, de Brent, de Sturm ...), ni sur la théorie de Galois : on trouvera d'excellents sites sur ces sujets passionnants.

2 Equations du second degré

Considérons l'équation du second degré : $x^2 = a_1x + a_2$ (dans laquelle les coefficients a_1 et a_2 sont des nombres complexes) ; on a alors (formules de Viète) : $a_1 = \sigma_1 = \lambda_1 + \lambda_2$ et $a_2 = -\sigma_2 = -\lambda_1\lambda_2$ (où λ_1 et λ_2 sont les deux racines de l'équation, réelles ou complexes, distinctes ou non).

Posons : M=$\begin{pmatrix} a_1 & a_2 \\ 1 & 0 \end{pmatrix}$.

Cette matrice est appelée : "matrice compagnon" du polynôme $P(x) = x^2 - a_1x - a_2$.

D'autre part, définissons la suite (u_n) par la relation de récurrence : $\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix}$=M$\begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix}$, et par les valeurs initiales : $u_0 = 0$ et $u_1 = 1$.

On a alors : $\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix}$, donc :

(1) $u_{n+1} = a_1\ u_n + a_2\ u_{n-1}$, et :

(2) $\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = M^n \begin{pmatrix} u_1 \\ u_0 \end{pmatrix} = M^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Posons : $q_n = \dfrac{u_{n+1}}{u_n}$;

$q_{n+1} = \dfrac{u_{n+2}}{u_{n+1}} = \dfrac{a_1u_{n+1} + a_2u_n}{u_{n+1}} = \dfrac{a_1u_{n+1}}{u_{n+1}} + \dfrac{a_2u_n}{u_{n+1}} = a_1 + \dfrac{a_2}{\frac{u_{n+1}}{u_n}} = a_1 + \dfrac{a_2}{q_n}$.

Si on réitère cette opération, on obtient :

$q_{n+1} = a_1 + \frac{a_2}{q_n} = a_1 + \frac{a_2}{a_1 + \frac{a_2}{q_{n-1}}} = a_1 + \frac{a_2}{a_1 + \frac{a_2}{a_1 + \frac{a_2}{q_{n-2}}}} = ...$

On voit apparaître un développement en fraction continue.

De plus, si (q_n) converge vers une limite l, on doit avoir, d'après l'égalité $q_{n+1} = a_1 + \frac{a_2}{q_n}$:

$l = a_1 + \frac{a_2}{l}$, donc $l^2 = a_1l + a_2$; donc $l = \lambda_1\ ou\ \lambda_2$. Nous verrons plus loin, d'une part, les conditions de convergence, et d'autre part, laquelle des deux racines est la limite effective de la suite. (On verra que c'est la racine qui a le plus grand module.)

Passons à la diagonalisation de la matrice M ; recherchons ses valeurs propres :

$\begin{vmatrix} a_1 - \lambda & a_2 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - a_1\lambda - a_2$; ce déterminant s'annule pour $\lambda = \lambda_1\ ou\ \lambda_2$; les

valeurs propres de la matrice M sont donc les racines de l'équation $x^2 = a_1x+a_2$.

Recherchons les vecteurs propres :

$\begin{pmatrix} a_1 & a_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda_i \begin{pmatrix} x \\ y \end{pmatrix}$ $(i = 1\ ou\ 2)$; donc :

$$\begin{cases} a_1x + a_2y = \lambda_i x \\ \\ x = \lambda_i y \end{cases}$$

Ces deux équations équivalent à : $x = \lambda_i y$.

Donc les vecteurs propres de M sont proportionnels à $\overrightarrow{v_1} = \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix}$ et $\overrightarrow{v_2} = \begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix}$.

Remarquons que : $M\overrightarrow{v_1} = \lambda_1\overrightarrow{v_1} = \begin{pmatrix} \lambda_1^2 \\ \lambda_1 \end{pmatrix}$ et $M^n\overrightarrow{v_1} = \lambda_1^n\overrightarrow{v_1} = \begin{pmatrix} \lambda_1^{n+1} \\ \lambda_1^n \end{pmatrix}$;

de même : $M\overrightarrow{v_2} = \lambda_2\overrightarrow{v_2} = \begin{pmatrix} \lambda_2^2 \\ \lambda_2 \end{pmatrix}$ et $M^n\overrightarrow{v_2} = \lambda_2^n\overrightarrow{v_2} = \begin{pmatrix} \lambda_2^{n+1} \\ \lambda_2^n \end{pmatrix}$.

Pour passer de la base de vecteurs propres à la base canonique, on utilisera la matrice $\begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}$, et pour passer de la base canonique à la base de vecteurs propres, on utilisera la matrice inverse, qui est : $\frac{1}{\lambda_1-\lambda_2}\begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix}$.

Nous supposerons, pour le moment : $\lambda_1 \neq \lambda_2$, et même, plus précisément : $|\lambda_1| > |\lambda_2|$.

On a donc :

(3) $M = \frac{1}{\lambda_1-\lambda_2}\begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix}$, et :

(4) $M^n = \frac{1}{\lambda_1-\lambda_2}\begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix}\begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix} = \begin{pmatrix} \frac{\lambda_1^{n+1}-\lambda_2^{n+1}}{\lambda_1-\lambda_2} & -\lambda_1\lambda_2\frac{\lambda_1^n-\lambda_2^n}{\lambda_1-\lambda_2} \\ \\ \frac{\lambda_1^n-\lambda_2^n}{\lambda_1-\lambda_2} & -\lambda_1\lambda_2\frac{\lambda_1^{n-1}-\lambda_2^{n-1}}{\lambda_1-\lambda_2} \end{pmatrix}$.

Cette égalité est encore valable pour $n < 0$, à condition que M soit inversible, ce qui suppose : $detM = -a_2 \neq 0$, donc $\lambda_1 \neq 0$ et $\lambda_2 \neq 0$. Il est alors possible de définir u_n pour $n < 0$, grâce à la relation de récurrence : $u_{n-1} = \frac{u_{n+1}-a_1u_n}{a_2}$.

Voyons les conséquences de l'égalité (4) sur la convergence de la suite (q_n) :

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = M^n \begin{pmatrix} u_1 \\ u_0 \end{pmatrix} = \begin{pmatrix} \frac{\lambda_1^{n+1}-\lambda_2^{n+1}}{\lambda_1-\lambda_2} & -\lambda_1\lambda_2\frac{\lambda_1^n-\lambda_2^n}{\lambda_1-\lambda_2} \\ \frac{\lambda_1^n-\lambda_2^n}{\lambda_1-\lambda_2} & -\lambda_1\lambda_2\frac{\lambda_1^{n-1}-\lambda_2^{n-1}}{\lambda_1-\lambda_2} \end{pmatrix} \begin{pmatrix} u_1 \\ u_0 \end{pmatrix}, \text{ donc :}$$

$$u_n = \frac{\lambda_1^n-\lambda_2^n}{\lambda_1-\lambda_2}u_1 - \lambda_1\lambda_2\frac{\lambda_1^{n-1}-\lambda_2^{n-1}}{\lambda_1-\lambda_2}u_0 = \lambda_1^n\frac{u_1-\lambda_2u_0}{\lambda_1-\lambda_2} - \lambda_2^n\frac{u_1-\lambda_1u_0}{\lambda_1-\lambda_2} = \lambda_1^n\left[\frac{u_1-\lambda_2u_0}{\lambda_1-\lambda_2} - \left(\frac{\lambda_2}{\lambda_1}\right)^n\frac{u_1-\lambda_1u_0}{\lambda_1-\lambda_2}\right].$$

Si $n \to \infty$, $(\frac{\lambda_2}{\lambda_1})^n \to 0$, donc $u_n \sim \lambda_1^n\frac{u_1-\lambda_2u_0}{\lambda_1-\lambda_2}$, donc $\frac{u_{n+1}}{u_n} \to \lambda_1$ (à condition que : $u_1 \neq \lambda_2u_0$).

De même :

$$u_n = \lambda_2^n\left[\left(\frac{\lambda_1}{\lambda_2}\right)^n\frac{u_1-\lambda_2u_0}{\lambda_1-\lambda_2} - \frac{u_1-\lambda_1u_0}{\lambda_1-\lambda_2}\right].$$

Si $n \to -\infty$, $(\frac{\lambda_1}{\lambda_2})^n \to 0$, donc $u_n \sim \lambda_2^n\frac{-u_1+\lambda_1u_0}{\lambda_1-\lambda_2}$, donc $\frac{u_{n+1}}{u_n} \to \lambda_2$ (à condition que : $u_1 \neq \lambda_1u_0$).

On voit donc que la convergence est indépendante du choix des valeurs initiales u_0 et u_1, pourvu que le vecteur $\begin{pmatrix} u_1 \\ u_0 \end{pmatrix}$ ne soit pas un vecteur propre.

Par conséquent, à toute équation du second degré de la forme : $x^2 = a_1x + a_2$, on peut associer une suite (u_n) telle que $u_{n+1} = a_1u_n + a_2u_{n-1}$; le "ratio" q_n converge, dans le cas général, vers la racine de l'équation qui a le plus grand module (si $n \to +\infty$) ou vers celle qui a le plus petit module (si $n \to -\infty$). Le cas des équations ayant deux racines de même module doit être examiné séparément.

Dans le cas particulier de l'équation $x^2 = x + 1$, les racines λ_1 et λ_2 sont les nombres d'or Φ_1 et Φ_2, et la suite (u_n) est la suite de Fibonacci F(n).

Voyons maintenant quelles sont les conditions pour que λ_1 soit un PV-number. Nous avons vu que :

$$u_n = \lambda_1^n\frac{u_1-\lambda_2u_0}{\lambda_1-\lambda_2} - \lambda_2^n\frac{u_1-\lambda_1u_0}{\lambda_1-\lambda_2}.$$

Supposons a_1 et a_2 entiers. Le choix de u_0 et u_1 est libre ; prenons : $u_0 = a_1 = \lambda_1 + \lambda_2$ et $u_1 = 2$; comme a_1, a_2, u_0 et u_1 sont entiers, on peut conclure (par une récurrence évidente : voir la définition de u_n) que u_n est entier pour tout n.

$$u_n = \lambda_1^n\frac{\lambda_1+\lambda_2-2\lambda_2}{\lambda_1-\lambda_2} - \lambda_2^n\frac{\lambda_1+\lambda_2-2\lambda_1}{\lambda_1-\lambda_2} = \lambda_1^n + \lambda_2^n.$$

Supposons : $|\lambda_2| < 1$ et $|\lambda_1| > 1$. Quand n tend vers l'infini, λ_2^n tend vers 0, donc $|u_n - \lambda_1^n|$ tend vers 0. Comme u_n est entier, λ_1 est soit un entier, soit un PV-number. En résumé, nous avons dégagé des conditions suffisantes (mais

pas nécessaires) pour que λ_1 soit un PV-number : a_1 entier, a_2 entier, $|\lambda_2| < 1$, $|\lambda_1| > 1$.

3 Fonctions symétriques

Une fonction de plusieurs variables est symétrique si sa valeur ne change pas quand on permute les variables.

Considérons une équation de degré n : $(x-\lambda_1)(x-\lambda_2)(x-\lambda_3)...(x-\lambda_n) = 0$, à n racines réelles ou complexes λ_1, λ_2,...,λ_n ; si nous développons le membre de gauche, nous obtenons :

$$x^n - \sigma_1 x^{n-1} + \sigma_2 x^{n-2} - \sigma_3 x^{n-3} + \sigma_4 x^{n-4} - \sigma_5 x^{n-5} + ... + (-1)^n \sigma_n = 0,$$

où σ_1, σ_2,..., σ_n sont des polynômes homogènes et symétriques en λ_1, λ_2,...,λ_n. Pour être plus rigoureux, on pourra les noter $\sigma_i(\lambda_1, \lambda_2, ..., \lambda_n)$, ou $\sigma_i^{(n)}$ (si on veut seulement préciser le nombre de racines), ou encore $\sigma_i(f)$, f étant le polynôme : $(X-\lambda_1)(X-\lambda_2)...(X-\lambda_n)$.

Ces polynômes σ_i sont appelés : fonctions symétriques élémentaires des racines.

Pour une équation de degré 2 (n=2 ; racines :λ_1 et λ_2), on a :

$$\begin{cases} \sigma_0 = 1 \ ; \\ \sigma_1 = \lambda_1 + \lambda_2 \ ; \\ \sigma_2 = \lambda_1 \lambda_2. \end{cases}$$

Pour une équation de degré 3 (n=3 ; racines :λ_1, λ_2 et λ_3), on a :

$$\begin{cases} \sigma_0 = 1 \ ; \\ \sigma_1 = \lambda_1 + \lambda_2 + \lambda_3 \ ; \\ \sigma_2 = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 \ ; \\ \sigma_3 = \lambda_1\lambda_2\lambda_3. \end{cases}$$

Pour une équation de degré 4 (n=4 ; racines :λ_1, λ_2, λ_3 et λ_4), on a :

$$\begin{cases} \sigma_0 = 1 \ ; \\ \sigma_1 = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \ ; \\ \sigma_2 = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_1\lambda_4 + \lambda_2\lambda_3 + \lambda_2\lambda_4 + \lambda_3\lambda_4 \ ; \\ \sigma_3 = \lambda_1\lambda_2\lambda_3 + \lambda_1\lambda_2\lambda_4 + \lambda_1\lambda_3\lambda_4 + \lambda_2\lambda_3\lambda_4 \ ; \\ \sigma_4 = \lambda_1\lambda_2\lambda_3\lambda_4. \end{cases}$$

La formule générale est : $\sigma_m(\lambda_1, \lambda_2, ..., \lambda_n) = \sum_{i_1+i_2+...+i_n=m} \lambda_1^{i_1}\lambda_2^{i_2}...\lambda_n^{i_n}$, avec $i_1, i_2, ..., i_n = 0$ *ou* 1.

$\sigma_m^{(n)}$ est donc la somme de tous les produits distincts qu'on peut former en prenant au plus une fois chaque racine ; c'est un polynôme formé de C_n^m monômes de degré n. ($C_n^m = \frac{n!}{m!(n-m)!}$)

$\sigma_m^{(n)}$ s'annule pour m>n.

Le théorème fondamental des polynômes symétriques dit que tout polynôme homogène et symétrique en λ_1, λ_2,..., peut s'écrire de manière unique sous forme de polynôme en $\sigma_1(\lambda_1, \lambda_2, ...), \sigma_2(\lambda_1, \lambda_2, ...), ...$

On définit également les fonctions symétriques complètes des racines, de la façon suivante : $\tau_m(\lambda_1, \lambda_2, ..., \lambda_n) = \sum_{i_1+i_2+...+i_n=m} \lambda_1^{i_1}\lambda_2^{i_2}...\lambda_n^{i_n}$, avec $i_1, i_2, ..., i_n \geq 0$.

$\tau_m^{(n)}$ est donc la somme de tous les produits distincts qu'on peut former en prenant chaque racine autant de fois qu'on veut ; c'est un polynôme formé de C_{m+n-1}^n monômes de degré n. ($C_{m+n-1}^n = \frac{(m+n-1)!}{n!(m-1)!}$)

Contrairement à $\sigma_m^{(n)}$, $\tau_m^{(n)}$ ne s'annule pas pour m>n.

Pour une équation de degré 2 (n=2 ; racines :λ_1 et λ_2), on a :

$$\begin{cases} \tau_0 = 1 \ ; \\ \tau_1 = \lambda_1 + \lambda_2 \ ; \\ \tau_2 = \lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2; \\ \tau_3 = \lambda_1^3 + \lambda_2^3 + \lambda_1^2\lambda_2 + \lambda_2^2\lambda_1; \\ etc. \end{cases}$$

Pour une équation de degré 3 (n=3 ; racines :λ_1, λ_2 et λ_3), on a :

$$\begin{cases} \tau_0 = 1 \ ; \\ \tau_1 = \lambda_1 + \lambda_2 + \lambda_3 \ ; \\ \tau_2 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 \ ; \\ \tau_3 = \lambda_1^3 + \lambda_2^3 + \lambda_3^3 + \lambda_1^2\lambda_2 + \lambda_1^2\lambda_3 + \lambda_2^2\lambda_1 + \lambda_2^2\lambda_3 + \lambda_3^2\lambda_1 + \lambda_3^2\lambda_2 + \lambda_1\lambda_2\lambda_3 \ ; \\ etc. \end{cases}$$

Pour une équation de degré 4 (n=4 ; racines :λ_1, λ_2, λ_3 et λ_4), on a :

$$\begin{cases} \tau_0 = 1 \ ; \\ \tau_1 = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \ ; \\ \tau_2 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_4^2 + \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_1\lambda_4 + \lambda_2\lambda_3 + \lambda_2\lambda_4 + \lambda_3\lambda_4 \ ; \\ etc. \end{cases}$$

On peut démontrer facilement les formules suivantes :

(5) $\sigma_m^{(n+1)} = \lambda_{n+1}\sigma_{m-1}^{(n)} + \sigma_m^{(n)}$

(6) $\sigma_m^{(n)} = \lambda_n\sigma_{m-1}^{(n-1)} + \lambda_{n-1}\sigma_{m-1}^{(n-2)} + \lambda_{n-2}\sigma_{m-1}^{(n-3)} + ... + \lambda_{n-i}\sigma_{m-1}^{(n-i+1)} + ... + \lambda_m\sigma_{m-1}^{(m-1)}$

(7) $\tau_m^{(n+1)} = \lambda_{n+1}\tau_{m-1}^{(n+1)} + \tau_m^{(n)}$

(8) $\tau_m^{(n+1)} = \lambda_{n+1}^m + \lambda_{n+1}^{m-1}\tau_1^{(n)} + \lambda_{n+1}^{m-2}\tau_2^{(n)} + \lambda_{n+1}^{m-3}\tau_3^{(n)} + ... + \lambda_{n+1}\tau_{m-1}^{(n)} + \tau_m^{(n)}$

(9) $\tau_m^{(n)} = \lambda_n\tau_{m-1}^{(n)} + \lambda_{n-1}\tau_{m-1}^{(n-1)} + \lambda_{n-2}\tau_{m-1}^{(n-2)} + \lambda_{n-3}\tau_{m-1}^{(n-3)} + ... + \lambda_1\tau_{m-1}^{(1)}$

La formule qui suit est moins évidente ; nous en donnons une démonstration.

(10) Pour tout entier m non nul : $\sum_{i=0}^{m}\left[(-1)^i\sigma_i^{(n)}\tau_{m-i}^{(n)}\right] = 0,$

ou, si on préfère : $\sum_{i+j=m}\left[(-1)^i\sigma_i^{(n)}\tau_j^{(n)}\right] = 0 \ (i,j \in N).$

(On peut écrire aussi : $\sum_{i+j=m}\left[(-1)^i\sigma_i(f)\tau_j(f)\right] = 0$, f étant un polynôme quelconque.)

Voici une démonstration élémentaire, basée sur les dénombrements. (Nous donnerons plus loin une autre démonstration, plus astucieuse, basée sur la convolution.)

$\sigma_i^{(n)}$ est la somme de tous les monômes de la forme $\lambda_1^{e_1}\lambda_2^{e_2}...\lambda_n^{e_n}$, dans lesquels la suite des exposants $(e_1, e_2, ..., e_n)$ comprend i fois le nombre 1 et n-i fois le nombre 0 (donc chaque monôme est de degré total i) ; $\tau_{m-i}^{(n)}$ est la somme de tous les monômes de la forme $\lambda_1^{e'_1}\lambda_2^{e'_2}...\lambda_n^{e'_n}$, dans lesquels la suite des exposants $(e'_1, e'_2, ..., e'_n)$ a pour somme m-i (donc chaque monôme est de degré total m-i, et chaque exposant e'_j est compris entre 0 et m-i).

Développons $\sigma_i^{(n)}\tau_{m-i}^{(n)}$ (i et m fixés), en multipliant chaque monôme de $\tau_{m-i}^{(n)}$ par chaque monôme de $\sigma_i^{(n)}$. Nous obtenons des monômes ayant tous le même

degré : i+(m-i)=m, donc appartenant à $\tau_m^{(n)}$, mais, par ce procédé :
- on n'obtient pas tous les monômes de $\tau_m^{(n)}$, mais seulement ceux qui contiennent au minimum i exposants non nuls (propriété de $\sigma_i^{(n)}$) ;
- d'autres monômes, au contraire, vont apparaître plusieurs fois. Précisons :

Prenons les monômes de $\tau_m^{(n)}$ (les "antécédents") et multiplions-les par les monômes de $\sigma_i^{(n)}$, de manière à obtenir des monômes de $\tau_m^{(n)}$ (les "images"). Remarquons que, quand on multiplie un antécédent par un monôme appartenant à $\sigma_i^{(n)}$, le nombre d'exposants modifiés est égal à i : ces i exposants augmentent d'une unité, et les autres ne changent pas.

Combien chaque image a-t-elle d'antécédents ?

Pour répondre, il suffit de compter le nombre d'exposants non nuls contenus dans ce monôme image ; soit k ce nombre ($1 \leq k \leq n$). Pour retrouver un antécédent, on doit réduire d'une unité i exposants parmi les k exposants non nuls ; il y a C_k^i possibilités. Donc un monôme de $\tau_m^{(n)}$ ayant k exposants non nuls est représenté C_k^i fois dans le développement de $\sigma_i^{(n)}\tau_{m-i}^{(n)}$.

Dans $\sum_{i=0}^{n}\left[(-1)^i\sigma_i^{(n)}\tau_{m-i}^{(n)}\right]$, ce même monôme est représenté $\sum_{i=0}^{n}(-1)^iC_k^i$ fois, ou si on préfère $\sum_{i=0}^{k}(-1)^iC_k^i$ (on arrête la sommation à i=k, car $C_k^i = 0$ pour $i > k$; rappelons que : $1 \leq k \leq n$).

Or on sait que, pour $k > 0$: $\sum_{i=0}^{k}(-1)^iC_k^i = (1-1)^k = 0$, d'après la formule bien connue du binôme de Newton. (Pour k=0 : on obtient 1 et non 0.)

On en déduit que $\sum_{i=0}^{n}\left[(-1)^i\sigma_i^{(n)}\tau_{m-i}^{(n)}\right] = 0$: c'est la formule annoncée.

Conséquence : $(-1)^0\sigma_0^{(n)}\tau_{m-0}^{(n)}+\sum_{i=1}^{n}\left[(-1)^i\sigma_i^{(n)}\tau_{m-i}^{(n)}\right] = 0$, donc : $\tau_m^{(n)} = \sum_{i=1}^{n}\left[(-1)^{i+1}\sigma_i^{(n)}\tau_{m-i}^{(n)}\right]$.

Cas particulier : si m=0, on a : $\sum_{i=0}^{n}\left[(-1)^i\sigma_i^{(n)}\tau_{m-i}^{(n)}\right] = \sigma_0^{(n)}\tau_0^{(n)} = 1$ (car tous les autres termes s'annulent).

(11) $$\tau_p(\lambda_1, \lambda_2, ..., \lambda_n) = \sum_{i_1+2i_2+3i_3+...+ni_n=p} \frac{(i_1+i_2+...+i_n)!}{i_1!i_2!i_3!...i_n!}\alpha_1^{i_1}\alpha_2^{i_2}\alpha_3^{i_3}...\alpha_n^{i_n}$$
(avec : $\alpha_i = (-1)^i\sigma_i$ et $i_1, i_2, ..., i_n \in N$).

Démonstration :

Cette formule est vraie pour p=0 ; faisons un raisonnement par récurrence : supposons-la vraie pour tout indice inférieur à un nombre p donné, et montrons qu'elle est vraie pour p. Nous allons partir de la formule (10).

$$\tau_p^{(n)} = \sum_{j=1}^{n} \left[(-1)^j \sigma_j^{(n)} \tau_{p-j}^{(n)} \right] = \sum_{j=1}^{n} \left(\alpha_j \tau_{p-j}^{(n)} \right)$$

$$\tau_p^{(n)} = \sum_{j=1}^{n} \left[\alpha_j \sum_{i_1+2i_2+3i_3+...+ni_n=p-j} \frac{(i_1+i_2+...+i_n)!}{i_1!i_2!i_3!...i_n!} \alpha_1^{i_1} \alpha_2^{i_2} \alpha_3^{i_3} ... \alpha_n^{i_n} \right]$$

Dans le crochet, la multiplication par α_j a pour effet d'augmenter d'une unité l'exposant i_j ; le produit ji_j augmente de j unités.

$$\tau_p^{(n)} = \sum_{j=1}^{n} \left[\sum_{i_1+...+j(i_j+1)+...+ni_n=p} \frac{(i_1+i_2+...+i_n)!}{i_1!i_2!i_3!...i_n!} \alpha_1^{i_1} \alpha_2^{i_2} ... \alpha_j^{(i_j+1)} ... \alpha_n^{i_n} \right]$$

Réécrivons cette égalité en remplaçant $i_j + 1$ par i_j (nouvelle valeur de l'exposant) :

$$\tau_p^{(n)} = \sum_{j=1}^{n} \left[\sum_{i_1+2i_2+...+ni_n=p,\ i_j \geq 1} \frac{(i_1+...+(i_j-1)+...+i_n)!}{i_1!...(i_j-1)!...i_n!} \alpha_1^{i_1} \alpha_2^{i_2} \alpha_3^{i_3} ... \alpha_n^{i_n} \right]$$

$$\tau_p^{(n)} = \sum_{i_1+2i_2+...+ni_n=p} \left[\sum_{i_j \geq 1} \frac{(i_1+...+(i_j-1)+...+i_n)!}{i_1!...(i_j-1)!...i_n!} \alpha_1^{i_1} \alpha_2^{i_2} \alpha_3^{i_3} ... \alpha_n^{i_n} \right]$$

Il reste donc à prouver que $\dfrac{(i_1+i_2+...+i_n)!}{i_1!i_2!i_3!...i_n!} = \displaystyle\sum_{i_j \geq 1} \frac{(i_1+...+(i_j-1)+...+i_n)!}{i_1!...(i_j-1)!...i_n!}$, ce qui est un pur problème de dénombrement.

Remarquons que :

$$\frac{(i_1+i_2+...+i_n)!}{i_1!i_2!i_3!...i_n!} = \frac{(i_1+i_2+...+i_n)!}{(i_1+i_2+...+i_{n-1})!i_n!} \frac{(i_1+i_2+...+i_{n-1})!}{(i_1+i_2+...+i_{n-2})!i_{n-1}!} ... \frac{(i_1+i_2)!}{i_1!i_2!} \frac{i_1!}{i_1!}$$

$$\frac{(i_1+i_2+...+i_n)!}{i_1!i_2!i_3!...i_n!} = C_{i_1+i_2+...+i_n}^{i_n} C_{i_1+i_2+...+i_{n-1}}^{i_{n-1}} C_{i_1+i_2+...+i_{n-2}}^{i_{n-2}} ... C_{i_1+i_2}^{i_2} C_{i_1}^{i_1}$$

C'est le nombre de n-uplets qu'on peut construire avec les éléments α_1, α_2,...,α_n, chaque élément α_k étant représenté i_k fois.

Mais si on fixe le dernier élément du n-uplet (un élément α_j choisi, tel que $i_j \geq 1$), il reste $\dfrac{(i_1+...+(i_j-1)+...+i_n)!}{i_1!i_2!...(i_j-1)!...i_n!}$ façons de construire le n-uplet ; d'où l'égalité annoncée.

Pour étudier les propriétés des fonctions σ et τ, il est commode d'utiliser la notion de convolution.

Si $(a_i)_{i\in N}$ et $(b_j)_{j\in N}$ sont deux suites appartenant à C^N (suites de nombres complexes indexées sur N), leur produit de convolution est la suite $(c_k)_{k\in N}$ définie par : $c_k = \sum_{i+j=k} a_i b_j$.

Si $(a_i)_{i\in N}$ est la suite des coefficients d'un polynôme f et $(b_j)_{j\in N}$ la suite des coefficients d'un polynôme g, alors leur produit de convolution est la suite $(c_k)_{k\in N}$ des coefficients du polynôme fg (produit des deux polynômes).

On sait que C^N, muni de l'addition et de la convolution, est un anneau intégre et unitaire ; la convolution est associative, commutative, distributive par rapport à l'addition ; son élément neutre est (1,0,0,...). Les éléments inversibles de cet anneau sont les suites dont le premier terme est non nul.

Considérons l'ensemble P des polynômes dont le monôme de plus fort degré a pour coefficient 1 (ce qui revient à considérer les polynômes à un coefficient multiplicatif non nul près : deux polynômes ayant les mêmes racines, donc les mêmes fonctions symétriques σ et τ, sont équivalents).

A tout polynôme f de P, nous allons associer un polynôme $\tilde{f}$ obtenu de la façon suivante : si $f(X) = X^n - \sigma_1 X^{n-1} + \sigma_2 X^{n-2} - \sigma_3 X^{n-3} + ... + (-1)^n \sigma_n$, alors :

$\tilde{f}(X) = 1 - \sigma_1 X + \sigma_2 X^2 - \sigma_3 X^3 + ... + (-1)^n \sigma_n X^n$. (Remarquons que $\widetilde{fg} = \tilde{f}\ \tilde{g}$)

Associons également à f la suite $(a_i(f))_{i\in N}$ définie par : $a_i(f) = (-1)^i \sigma_i(f)$.

Notons que cette suite $(a_i(f))$ n'est pas la suite des coefficients de f, mais de $\tilde{f}$.

Etendons notre définition au corps des fractions de P : si $r = \frac{f}{g}$ $(f, g \in P)$, $(a_i(r))_{i\in N}$ sera le développement en série de $\frac{\tilde{f}}{\tilde{g}}$. Nous noterons : $\sigma_i(r) = (-1)^i a_i(r)$.

Considérons le polynôme $f(X) = (X - \lambda_1)(X - \lambda_2)$; on a alors : $\tilde{f}(X) = (1 - \lambda_1 X)(1 - \lambda_2 X)$.

$(a_i(\frac{1}{\tilde{f}}))_{i\in N}$ est le développement en série de $\frac{1}{\tilde{f}}$.

$\frac{1}{\tilde{f}(X)} = \frac{1}{(1-\lambda_1 X)(1-\lambda_2 X)}$

$\frac{1}{\tilde{f}(X)} = (1 + \lambda_1 X + \lambda_1^2 X^2 + \lambda_1^3 X^3 + ...)(1 + \lambda_2 X + \lambda_2^2 X^2 + \lambda_3^3 X^3 + ...)$

$\frac{1}{f(X)} = 1 + (\lambda_1 + \lambda_2)X + (\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2)X^2 + (\lambda_1^3 + \lambda_1^2\lambda_2 + \lambda_1\lambda_2^2 + \lambda_2^3)X^3 + ...$

On reconnait les termes de la suite $\tau(f)$. On peut démonter (par récurrence sur le degré) que, pour tout polynôme f de P, on a : $a_i(\frac{1}{f}) = \tau_i(f)$, donc :

(12) $\sigma_i(\frac{1}{f}) = (-1)^i\tau_i(f)$ et $\tau_i(\frac{1}{f}) = (-1)^i\sigma_i(f)$.

Calculons les fonctions symétriques d'un produit fg $(f, g \in P)$:

$a_k(fg) = \sum_{i+j=k} a_i(f)a_i(g)$, donc :

$(-1)^k\sigma_k(fg) = \sum_{i+j=k} (-1)^i\sigma_i(f)(-1)^j\sigma_i(g)$, soit :

(13) $\sigma_k(fg) = \sum_{i+j=k} \sigma_i(f)\sigma_j(g)$, et, en généralisant à un nombre quelconque p de polynômes :

$$\sigma_k(f_1f_2...f_p) = \sum_{i_1+i_2+...+i_p=k} \sigma_{i_1}(f_1)\sigma_{i_2}(f_2)...\sigma_{i_p}(f_p) \ (i_1, i_2, ..., i_p \in N)$$

De même pour τ :

$$\tau_k(fg) = a_k(\tfrac{1}{fg}) = \sum_{i+j=k} a_i(\frac{1}{f})a_j(\frac{1}{g}) = \sum_{i+j=k} \tau_i(f)\tau_j(g)$$

(14) $\tau_k(fg) = \sum_{i+j=k} \tau_i(f)\tau_j(g)$, et :

$$\tau_k(f_1f_2...f_p) = \sum_{i_1+i_2+...+i_p=k} \tau_{i_1}(f_1)\tau_{i_2}(f_2)...\tau_{i_p}(f_p) \ (i_1, i_2, ..., i_p \in N)$$

Calculons aussi les fonctions symétriques d'un quotient :

$$\sigma_k(\tfrac{f}{g}) = \sum_{i+j=k} \sigma_i(f)\sigma_j(\frac{1}{g}) = \sum_{i+j=k} \sigma_i(f)(-1)^j\tau_j(g)$$

(15) $\sigma_k(\frac{f}{g}) = \sum_{i+j=k} (-1)^j\sigma_i(f)\tau_j(g)$, et :

$$\sigma_k(\tfrac{f_1f_2...f_p}{g_1g_2...g_q}) = \sum_{i_1+...+i_p+j_1+...+j_q=k} (-1)^{j_1+...+j_q}\sigma_{i_1}(f_1)\sigma_{i_2}(f_2)...\sigma_{i_p}(f_p)\tau_{j_1}(g_1)\tau_{j_2}(g_2)...\tau_{j_q}(g_q)$$

$(i_1, i_2, ..., i_p, j_1, j_2, ..., j_q \in N)$

Remarquons que $\sigma_k(id) = \sigma_k(\frac{f}{f}) = \sum_{i+j=k} (-1)^j\sigma_i(f)\tau_j(f) = \delta_{k,0}$: cette somme est égale à 0 pour tout k différent de 0 (formule 10) ; pour k=0, on obtient 1, car $(-1)^0\sigma_0(f)\tau_0(f) = 1$.

De même pour τ :

$\tau_k(\frac{f}{g}) = (-1)^k \sigma_k(\frac{g}{f}) = \sum_{i+j=k} (-1)^j \sigma_i(g)\tau_j(f)$, ou :

(16) $\tau_k(\frac{f}{g}) = \sum_{i+j=k} (-1)^j \tau_i(f)\sigma_j(g)$ $(i, j \in N)$, et :

$\tau_k(\frac{f_1 f_2 ... f_p}{g_1 g_2 ... g_q}) = \sum_{i_1+...+i_p+j_1+...+j_q=k} (-1)^{j_1+...+j_q} \tau_{i_1}(f_1)\tau_{i_2}(f_2)...\tau_{i_p}(f_p)\sigma_{j_1}(g_1)\sigma_{j_2}(g_2)...\sigma_{j_q}(g_q)$

$(i_1, i_2, ..., i_p, j_1, j_2, ..., j_q \in N)$

Nous avions annoncé une seconde démonstration de la formule (10), basée sur la convolution ; la voici :

Définissons deux fonctions f et g, à variable complexe z, de la façon suivante :

$f(z) = \prod_{i=1}^{n}(1 - \lambda_i z)$ et $g(z) = \prod_{i=1}^{n} \frac{1}{1 - \lambda_i z}$.

D'après ce qui précède :

$f(z) = \sum_{i=0}^{\infty} \sigma_i(-z)^i = \sum_{i=0}^{\infty}(-1)^i \sigma_i z^i$ et $g(z) = \sum_{j=0}^{\infty} \tau_j z^j$; ces fonctions étant inverses :

$\left(\sum_{i=0}^{\infty}(-1)^i \sigma_i z^i\right)\left(\sum_{j=0}^{\infty} \tau_j z^j\right) = 1$, donc, en utilisant la convolution (et en posant $i + j = m$) :

$\sum_{m=0}^{\infty} \sum_{i+j=m} (-1)^i \sigma_i \tau_j z^{i+j} = 1$, donc :

$\sum_{m=0}^{\infty} \left(\sum_{i=0}^{m}(-1)^i \sigma_i \tau_{m-i}\right) z^m = 1.$

Donc tous les coefficients de la série sont nuls, sauf le premier :

$\sum_{i=0}^{m}(-1)^i \sigma_i \tau_{m-i} = \delta_{m,0}$ (où δ est le symbole de Kronecker).

Remarquons que cette formule permet de calculer, par récurrence, les nombres de la famille (τ_i) en utilisant la famille (σ_i), et réciproquement.

Voici maintenant (sans rapport avec la convolution) une formule qui va jouer un rôle essentiel, et que nous démontrerons plus loin :

(17) m et n étant deux entiers strictement positifs, et les racines λ_1, λ_2,..., λ_n étant toutes distinctes, on a l'égalité : $\tau_m^{(n)} = \sum_{i=1}^{n} \frac{\lambda_i^{m+n-1}}{\prod_{j\in\{1,...,n\}-\{i\}} (\lambda_i - \lambda_j)}$, où le produit est étendu à tous les entiers j différents de i, compris entre 1 et n.

Une autre formule qui nous sera utile est la suivante :

(18) $\sigma_p(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n) = \sigma_p(\lambda_1, \lambda_2, ..., \lambda_n) - \lambda_i \sigma_{p-1}(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n)$.

Cette formule se démontre facilement. En l'appliquant de manière répétée, on obtient :

$\sigma_p(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n) = \sigma_p(\lambda_1, \lambda_2, ..., \lambda_n) - \lambda_i(\sigma_{p-1}(\lambda_1, \lambda_2, ..., \lambda_n) - \lambda_i(\sigma_{p-2}(\lambda_1, \lambda_2, ..., \lambda_n) - \lambda_i(\sigma_{p-3}(\lambda_1, \lambda_2, ..., \lambda_n) - \lambda_i(...))))$, soit :

$\sigma_p(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n) = \sigma_p(\lambda_1, \lambda_2, ..., \lambda_n) - \lambda_i \sigma_{p-1}(\lambda_1, \lambda_2, ..., \lambda_n) + \lambda_i^2 \sigma_{p-2}(\lambda_1, \lambda_2, ..., \lambda_n) - \lambda_i^3 \sigma_{p-3}(\lambda_1, \lambda_2, ..., \lambda_n) + \lambda_i^4 \sigma_{p-4}(\lambda_1, \lambda_2, ..., \lambda_n) - ...$, ou encore :

(19) $\sigma_p(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n) = \sum_{k=0}^{p} (-1)^k \lambda_i^k \sigma_{p-k}(\lambda_1, ..., \lambda_n)$.

Etudions maintenant les dérivées partielles de $\sigma_i(\lambda_1, \lambda_2, \lambda_n)$ par rapport à λ_i. Prenons comme exemple le cas n=3.

$\sigma_0 = 1$
$\sigma_1 = \lambda_1 + \lambda_2 + ... + \lambda_n$
$\sigma_2 = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3$
$\sigma_3 = \lambda_1\lambda_2\lambda_3$

$\frac{\partial \sigma_1}{\partial \lambda_1} = 1 = \sigma_0(\lambda_2, \lambda_3)$

$\frac{\partial \sigma_2}{\partial \lambda_1} = \lambda_2 + \lambda_3 = \sigma_1(\lambda_2, \lambda_3)$

$\frac{\partial \sigma_3}{\partial \lambda_1} = \lambda_2\lambda_3 = \sigma_2(\lambda_2, \lambda_3)$

(De même pour les autres variables.)

$$\begin{pmatrix} \partial\sigma_1 \\ \partial\sigma_2 \\ \partial\sigma_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ \sigma_1(\lambda_2, \lambda_3) & \sigma_1(\lambda_1, \lambda_3) & \sigma_1(\lambda_1, \lambda_2) \\ \sigma_2(\lambda_2, \lambda_3) & \sigma_2(\lambda_1, \lambda_3) & \sigma_2(\lambda_1, \lambda_2) \end{pmatrix} \begin{pmatrix} \partial\lambda_1 \\ \partial\lambda_2 \\ \partial\lambda_3 \end{pmatrix} = T \begin{pmatrix} \partial\lambda_1 \\ \partial\lambda_2 \\ \partial\lambda_3 \end{pmatrix}$$

La matrice $T = \begin{pmatrix} 1 & 1 & 1 \\ \sigma_1(\lambda_2,\lambda_3) & \sigma_1(\lambda_1,\lambda_3) & \sigma_1(\lambda_1,\lambda_2) \\ \sigma_2(\lambda_2,\lambda_3) & \sigma_2(\lambda_1,\lambda_3) & \sigma_2(\lambda_1,\lambda_2) \end{pmatrix}$ a pour déterminant un polynôme Δ symétrique en $\lambda_1,\ \lambda_2,\ \lambda_3$; son degré est 3, car c'est une somme de produits de 3 termes appartenant respectivement à la première ligne (degré 0), à la seconde (degré 1) et à la troisième ligne (degré 2) ; et $0+1+2=3$.

Ce déterminant Δ s'annule pour $\lambda_1 = \lambda_2$ (dans ce cas, les 2 premières colonnes sont identiques), pour $\lambda_1 = \lambda_3$ et pour $\lambda_2 = \lambda_3$ (raisons similaires) ; donc il est divisible par $(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)(\lambda_2-\lambda_3)$, donc (puisque son degré est 3) il est égal, à un coefficient multiplicatif non nul près, au produit $(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)(\lambda_2-\lambda_3)$.

Il n'est pas très difficile de trouver l'inverse de cette matrice T. On obtient :

$$T^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ \sigma_1(\lambda_2,\lambda_3) & \sigma_1(\lambda_1,\lambda_3) & \sigma_1(\lambda_1,\lambda_2) \\ \sigma_2(\lambda_2,\lambda_3) & \sigma_2(\lambda_1,\lambda_3) & \sigma_2(\lambda_1,\lambda_2) \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ \lambda_2+\lambda_3 & \lambda_1+\lambda_3 & \lambda_1+\lambda_2 \\ \lambda_2\lambda_3 & \lambda_1\lambda_3 & \lambda_1\lambda_2 \end{pmatrix}^{-1}$$

$$T^{-1} = \begin{pmatrix} \frac{\lambda_1^2}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} & \frac{-\lambda_1}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} & \frac{1}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} \\ \\ \frac{\lambda_2^2}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} & \frac{-\lambda_2}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} & \frac{1}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} \\ \\ \frac{\lambda_3^2}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} & \frac{-\lambda_3}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} & \frac{1}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} \end{pmatrix}.$$

Pour une équation de degré n, la matrice T aura pour terme général :

(20) $\gamma_{i,j} = \sigma_{i-1}(\lambda_1, \lambda_2, ..., \lambda_{j-1}, \lambda_{j+1}, ..., \lambda_n)$;

et son inverse T^{-1} aura pour terme général :

(21) $\delta_{i,j} = (-1)^{j+1} \dfrac{\lambda_i^{n-j}}{\prod\limits_{k\in\{1,...,n\}-\{i\}} (\lambda_i - \lambda_k)}$.

Vérifions-le en multipliant la $i^{ème}$ ligne de T^{-1} avec la $j^{ème}$ colonne de T :

$$\sum_{l=1}^{n} \delta_{i,l}\gamma_{l,j} = \sum_{l=1}^{n} (-1)^{l+1} \frac{\lambda_i^{n-l}\sigma_{l-1}(\lambda_1, \lambda_2, ..., \lambda_{j-1}, \lambda_{j+1}, ..., \lambda_n)}{\prod\limits_{k\in\{1,...,n\}-\{i\}} (\lambda_i - \lambda_k)}$$

$$\sum_{l=1}^{n} \delta_{i,l}\gamma_{l,j} = \frac{\sum_{l=1}^{n}(-1)^{l+1}\lambda_i^{n-l}\sigma_{l-1}(\lambda_1,\lambda_2,...,\lambda_{j-1},\lambda_{j+1},...,\lambda_n)}{\prod_{k\in\{1,...,n\}-\{i\}}(\lambda_i-\lambda_k)}$$

Au numérateur, on reconnaît le développement de $(X-\lambda_1)(X-\lambda_2)...(X-\lambda_{j-1})(X-\lambda_{j+1})...(X-\lambda_n)$, où X a été remplacé par λ_i. Si $i=j$, ce produit est identique à celui qui se trouve au dénominateur, donc le quotient est égal à 1 ; sinon, l'un des facteurs s'annule, donc le produit est nul.

4 Matrices et déterminants de Vandermonde

Une matrice de Vandermonde est une matrice de la forme :

$$\begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & ... & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & ... & \lambda_2^{n-1} \\ 1 & ... & ... & ... & ... \\ 1 & \lambda_n & \lambda_n^2 & ... & \lambda_n^{n-1} \end{pmatrix}$$

On sait que son déterminant (déterminant de Vandermonde) est :

(22) $D(\lambda_1,\lambda_2,...,\lambda_n) = \prod_{1\leq j<i\leq n}(\lambda_i-\lambda_j)$ (produit de $\frac{n(n-1)}{2}$ facteurs).

Une démonstration consiste à remarquer que le déterminant s'annule si deux lignes sont identiques, ce qui se produit pour $\lambda_i=\lambda_j$; donc il est divisible par $\prod_{1\leq j<i\leq n}(\lambda_i-\lambda_j)$. De plus, ce déterminant, considéré comme un polynôme à n indéterminées $(\lambda_1,\lambda_2,...,\lambda_n)$, a pour degré total : $(n-1)+(n-2)+...+1=\frac{n(n-1)}{2}$; le polynôme $\prod_{1\leq j<i\leq n}(\lambda_i-\lambda_j)$ a précisément le même degré; donc $D(\lambda_1,\lambda_2,...,\lambda_n)$ est égal à $\prod_{1\leq j<i\leq n}(\lambda_i-\lambda_j)$, à un coefficient multiplicatif près. Il reste à prouver que ce coefficient est 1.

Faisons un raisonnement par récurrence : supposons que $D(\lambda_1,\lambda_2,...,\lambda_p) = \prod_{1\leq j<i\leq p}(\lambda_i-\lambda_j)$ pour $p<n$ (c'est vrai pour p=2) et que $D(\lambda_1,\lambda_2,...,\lambda_n) = k\prod_{1\leq j<i\leq n}(\lambda_i-\lambda_j)$, et montrons que k=1.

Si on développe le déterminant $\begin{vmatrix} 1 & \lambda_1 & \lambda_1^2 & ... & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & ... & \lambda_2^{n-1} \\ 1 & ... & ... & ... & ... \\ 1 & \lambda_n & \lambda_n^2 & ... & \lambda_n^{n-1} \end{vmatrix}$ en commençant par

λ_n^{n-1} (situé en bas à droite), on voit que son coefficient est :

$$\begin{vmatrix} 1 & \lambda_1 & \lambda_1^2 & ... & \lambda_1^{n-2} \\ 1 & \lambda_2 & \lambda_2^2 & ... & \lambda_2^{n-2} \\ 1 & ... & ... & ... & ... \\ 1 & \lambda_{n-1} & \lambda_{n-1}^2 & ... & \lambda_{n-1}^{n-2} \end{vmatrix} = D(\lambda_1, \lambda_2, ..., \lambda_{n-1}).$$

D'un autre côté, si on part de l'égalité $D(\lambda_1, \lambda_2, ..., \lambda_n) = k \prod_{1 \leq j < i \leq n} (\lambda_i - \lambda_j)$, on obtient (en séparant les facteurs qui contiennent λ_n de ceux qui ne le contiennent pas) :

$$D(\lambda_1, \lambda_2, ..., \lambda_n) = k \prod_{1 \leq j < n} (\lambda_n - \lambda_j) \prod_{1 \leq j < i < n} (\lambda_i - \lambda_j) = k \prod_{1 \leq j < n} (\lambda_n - \lambda_j) D(\lambda_1, \lambda_2, ..., \lambda_{n-1}).$$

Ceci permet de voir que le coefficient de λ_n^{n-1} est $kD(\lambda_1, \lambda_2, ..., \lambda_{n-1})$; d'où la conclusion : k=1.

Dans la suite, nous allons rencontrer des matrices de la forme :

$$\begin{pmatrix} \lambda_1^{n-1} & \lambda_2^{n-1} & ... & \lambda_n^{n-1} \\ ... & ... & ... & ... \\ \lambda_1^2 & \lambda_2^2 & ... & \lambda_n^2 \\ \lambda_1 & \lambda_2 & ... & \lambda_n \\ 1 & 1 & ... & 1 \end{pmatrix}$$

Une telle matrice s'obtient, à partir d'une matrice de Vandermonde, par :

- la "transposition" de la matrice (les lignes deviennent colonnes, et réciproquement) ; ceci ne modifie pas le déterminant ;

- l'inversion complète de l'ordre des lignes (cette permutation fait intervenir $\frac{n(n-1)}{2}$ interversions de lignes consécutives ou "transpositions", donc le déterminant est multiplié par $(-1)^{\frac{n(n-1)}{2}}$, ou $(-1)^{\left[\frac{n}{2}\right]}$, où le crochet désigne la partie entière de $\frac{n}{2}$:

$(-1)^{\left[\frac{n}{2}\right]}$ est égal à +1 si n=4p ou 4p+1, et à -1 si n=4p+2 ou 4p+3.

Démontrons maintenant la formule (17), déjà annoncée au paragraphe précédent :

m et n étant deux entiers strictement positifs, et les racines λ_1, λ_2,..., λ_n étant

toutes distinctes, on a l'égalité : $\tau_m^{(n)} = \sum_{i=1}^{n} \frac{\lambda_i^{m+n-1}}{\prod_{j\in\{1,...,n\}-\{i\}} (\lambda_i - \lambda_j)}$, où le produit est étendu à tous les entiers j différents de i, compris entre 1 et n.

- Pour n=1, l'égalité se réduit à : $\tau_m^{(1)} = \lambda_1^m$, ce qui est manifestement vrai.

- Pour m=1, l'égalité à démontrer est : $\tau_1^{(n)} = \sum_{i=1}^{n} \frac{\lambda_i^n}{\prod_{j\in\{1,...,n\}-\{i\}} (\lambda_i - \lambda_j)}$.

Au dénominateur, remplaçons $(\lambda_i - \lambda_j)$ par $(\lambda_j - \lambda_i)$ chaque fois que j est supérieur à i (donc n-i fois) ; ceci revient à multiplier le dénominateur par $(-1)^{n-i}$. Ensuite, multiplions-le par $D(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n)$ (où λ_i a été exclu) ; nous obtenons alors $D(\lambda_1, \lambda_2, ..., \lambda_n)$. Faisons de même au numérateur. Il vient :

$$\sum_{i=1}^{n} \frac{\lambda_i^n}{\prod_{j\in\{1,...,n\}-\{i\}} (\lambda_i - \lambda_j)} = \frac{\sum_{i=1}^{n} (-1)^{n-i} \lambda_i^n D(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n)}{D(\lambda_1, \lambda_2, ..., \lambda_n)}$$

Etudions le numérateur de cette fraction :

$N(\lambda_1, \lambda_2, ..., \lambda_n) = \sum_{i=1}^{n} (-1)^{n-i} \lambda_i^n D(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n)$; montrons d'abord qu'il est multiplié par (-1) quand on permute deux racines.

On voit que $\sum_{i=1}^{n} \frac{\lambda_i^n}{\prod_{j\in\{1,...,n\}-\{i\}} (\lambda_i - \lambda_j)}$ est indépendant de l'ordre des racines ; si on permute deux racines quelconques, le dénominateur $D(\lambda_1, \lambda_2, ..., \lambda_n)$ change de signe, donc il en est de même pour le numérateur $N(\lambda_1, \lambda_2, ..., \lambda_n)$.

Montrons ensuite que $N(\lambda_1, \lambda_2, ..., \lambda_n)$ s'annule si deux racines λ_k et λ_l $(k \neq l)$ sont égales. Nous pouvons supposer que k=1 et l=2, puisqu'une permutation des racines ne peut pas modifier la valeur absolue de $N(\lambda_1, \lambda_2, ..., \lambda_n)$.

Si $\lambda_1 = \lambda_2$, tous les déterminants de Vandermonde contenant à la fois λ_1 et λ_2 s'annulent, et $N(\lambda_1, \lambda_2, ..., \lambda_n)$ se réduit donc à une somme de deux termes :

$N(\lambda_1, \lambda_2, ..., \lambda_n) = (-1)^{n-1} \lambda_1^n D(\lambda_2, \lambda_3, ..., \lambda_n) + (-1)^{n-2} \lambda_2^n D(\lambda_1, \lambda_3, ..., \lambda_n) = 0$.

On en déduit que $N(\lambda_1, \lambda_2, ..., \lambda_n)$ est divisible par $(\lambda_k - \lambda_l)$ $(k \neq l)$, donc par $D(\lambda_1, \lambda_2, ..., \lambda_n)$. Notons $Q(\lambda_1, \lambda_2, ..., \lambda_n)$ le polynôme quotient. Ce polynôme

est symétrique (quand on inverse deux racines, $N(\lambda_1, \lambda_2, ..., \lambda_n)$ et $D(\lambda_1, \lambda_2, ..., \lambda_n)$ changent de signe, donc leur quotient n'est pas modifié) et homogène.

Le degré total de $N(\lambda_1, \lambda_2, ..., \lambda_n)$ est $n+\frac{(n-1)(n-2)}{2} = \frac{n^2-n+2}{2}$, celui de $D(\lambda_1, \lambda_2, ..., \lambda_n)$ est $\frac{n(n-1)}{2} = \frac{n^2-n}{2}$, donc le degré total de $Q(\lambda_1, \lambda_2, ..., \lambda_n)$ est : $\frac{n^2-n+2}{2} - \frac{n^2-n}{2} = 1$.

Le degré partiel de $N(\lambda_1, \lambda_2, ..., \lambda_n)$ (par rapport à une indéterminée quelconque λ_i) est n, celui de $D(\lambda_1, \lambda_2, ..., \lambda_n)$ est n-1, donc le degré partiel de $Q(\lambda_1, \lambda_2, ..., \lambda_n)$ est : $n - (n-1) = 1$.

D'après le théorème fondamental des polynômes symétriques, $Q(\lambda_1, \lambda_2, ..., \lambda_n)$ peut s'exprimer comme polynôme en $\sigma_i(\lambda_1, \lambda_2, ..., \lambda_n)$; mais le seul σ_i qui soit de degré 1 en λ_i est $\sigma_1(\lambda_1, \lambda_2, ..., \lambda_n) = \lambda_1 + \lambda_2 + ... + \lambda_n = \tau_1(\lambda_1, \lambda_2, ..., \lambda_n)$. Donc $Q(\lambda_1, \lambda_2, ..., \lambda_n)$ est égal à $\lambda_1 + \lambda_2 + ... + \lambda_n$, à un coefficient multiplicatif non nul près. Il reste à expliquer pourquoi ce coefficient est 1.

Dans $(-1)^{n-i}\lambda_i^n D(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n)$, le degré partiel de λ_i est n, et le degré partiel des autres indéterminées est n-2.

Dans $N(\lambda_1, \lambda_2, ..., \lambda_n)$, considéré comme un polynôme à une seule indéterminée λ_i, cherchons le monôme en λ_i ayant le plus haut degré :

c'est $(-1)^{n-i}\lambda_i^n D(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n)$; son degré est n, et son coefficient est $(-1)^{n-i}D(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n)$.

D'autre part, dans le polynôme :

$$D(\lambda_1, \lambda_2, ..., \lambda_n) = (-1)^{n-i}D(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n) \prod_{j\in\{1,...,n\}-\{i\}} (\lambda_i - \lambda_j),$$

le monôme de plus haut degré en λ_i est $(-1)^{n-i}D(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n)\lambda_i^{n-1}$: son degré est n-1, son coefficient est encore $(-1)^{n-i}D(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n)$.

En divisant selon les puissances décroissantes, on voit que le coefficient de λ_i dans $Q(\lambda_1, \lambda_2, ..., \lambda_n)$ ne peut être que 1.

- pour n>1 et m>1 : faisons un raisonnement par récurrence. Nous supposons que la formule est vérifiée pour τ_{m-1}^n et pour τ_m^{n-1}, et nous la démontrons pour τ_m^n.

Partons de la formule (7) : $\tau_m^{(n)} = \lambda_n \tau_{m-1}^{(n)} + \tau_m^{(n-1)}$.

$$\tau_m^{(n)} = \lambda_n \sum_{i=1}^{n} \frac{\lambda_i^{m+n-2}}{\prod_p rod_{j\in\{1,...,n\}-\{i\}}(\lambda_i - \lambda_j)} + \sum_{i=1}^{n-1} \frac{\lambda_i^{m+n-2}}{\prod_{j\in\{1,...,n-1\}-\{i\}} (\lambda_i - \lambda_j)}$$

$$\tau_m^{(n)} = \sum_{i=1}^{n} \frac{\lambda_n \lambda_i^{m+n-2}}{\prod\limits_{j\in[1,n]-\{i\}} (\lambda_i - \lambda_j)} + \sum_{i=1}^{n-1} \frac{\lambda_i^{m+n-2}(\lambda_i - \lambda_n)}{\prod\limits_{j\in\{1,\ldots,n\}-\{i\}} (\lambda_i - \lambda_j)}$$

(On peut sommer jusqu'à i=n, car le dernier terme est nul.)

$$\tau_m^{(n)} = \sum_{i=1}^{n} \frac{\lambda_i^{m+n-2}(\lambda_n + (\lambda_i - \lambda_n))}{\prod\limits_{j\in\{1,\ldots,n\}-\{i\}} (\lambda_i - \lambda_j)} = \sum_{i=1}^{n} \frac{\lambda_i^{m+n-1}}{\prod\limits_{j\in\{1,\ldots,n\}-\{i\}} (\lambda_i - \lambda_j)}.$$

La formule est donc vérifiée pour τ_m^n.

Attention : cette récurrence portant sur deux indices à la fois doit être bien comprise. Sachant que la formule est vraie pour n=1 (m quelconque) et pour n=1 (m quelconque), on peut déduire qu'elle est vraie pour n=2 et m=2, 3, 4, ..., puis pour n=3 et m=2, 3, 4, ..., puis pour n=4 et m=2, 3, 4, ..., etc.

Inverse d'une matrice de Vandermonde

Considérons la matrice : $\begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \ldots & \lambda_2^{n-1} \\ 1 & \ldots & \ldots & \ldots & \ldots \\ 1 & \lambda_n & \lambda_n^2 & \ldots & \lambda_n^{n-1} \end{pmatrix}$.

Pour qu'elle soit inversible, il faut et il suffit que les λ_i soient tous distincts ; c'est ce que nous supposerons dans les calculs qui suivent.

Nous voulons calculer l'inverse de cette matrice. Pour cela, il faut calculer le cofacteur de chaque coefficient de la matrice, remplacer chaque coefficient par son cofacteur, puis transposer la matrice ainsi obtenue, et la diviser par le déterminant de la matrice d'origine.

Pour calculer le cofacteur d'un coefficient donné, par exemple λ_i^{j-1} (qui est situé sur la $i^{ème}$ ligne et la $j^{ème}$ colonne), on supprime la $i^{ème}$ ligne et la $j^{ème}$ colonne, on calcule le déterminant de la matrice ainsi obtenue, et on le multiplie par $(-1)^{i+j}$.

Après avoir supprimé la ligne et la colonne de λ_i^{j-1}, nous obtenons une matrice de rang n-1 ; son déterminant est un polynôme de degré n-1, qui s'annule si $\lambda_k = \lambda_l$ $(k \neq i,\ l \neq i,\ k \neq l)$ (car dans ce cas, il y a deux lignes identiques). Il est donc divisible par $D(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n)$. Si on effectue la division, on obtient un polymôme symétrique et homogène $Q_{i,j}$.

Le déterminant : $D_1 = \begin{vmatrix} 1 & \lambda_1 & \lambda_1^2 & ... & \lambda_1^{j-2} & \lambda_1^j & ... & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & ... & \lambda_2^{j-2} & \lambda_2^j & ... & \lambda_2^{n-1} \\ 1 & ... & ... & ... & ... & ... & ... & ... \\ 1 & \lambda_{i-1} & \lambda_{i-1}^2 & ... & \lambda_{i-1}^{j-2} & \lambda_{i-1}^j & ... & \lambda_{i-1}^{n-1} \\ 1 & \lambda_{i+1} & \lambda_{i+1}^2 & ... & \lambda_{i+1}^{j-2} & \lambda_{i+1}^j & ... & \lambda_{i+1}^{n-1} \\ 1 & ... & ... & ... & ... & ... & ... & ... \\ 1 & \lambda_n & \lambda_n^2 & ... & \lambda_n^{j-2} & \lambda_n^j & ... & \lambda_n^{n-1} \end{vmatrix}$

...a pour degré partiel n-1 (ou n-2, si la colonne supprimée est la dernière) ; son degré total est la somme des exposants d'une même ligne, soit : $0+1+2+...+(n-1)-(j-1) = \frac{n(n-1)}{2} - j + 1$.

Le déterminant $D_2 = D(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n)$ a pour degré partiel n-2 (il y a n-2 facteurs contenant λ_n, par exemple) ; son degré total est le nombre total de facteurs, soit : $\frac{(n-1)(n-2)}{2}$.

Le degré partiel et le degré total du polynôme quotient $Q_{i,j}$ s'obtiennent par soustraction :

- degré partiel : $(n-1)-(n-2) = 1$;

- degré total : $\frac{n(n-1)}{2} - (j-1) - \frac{(n-1)(n-2)}{2} = n - j$.

Le seul polynôme symétrique et homogène de degré partiel 1 et de degré total n-j qui convienne est $\sigma_{n-j}(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n)$.

Donc $Q_{i,j} = \sigma_{n-j}(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n)$, et :

$$D_1 = Q_{i,j}.D_2 = \sigma_{n-j}(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n) D(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n).$$

Notons $\alpha_{i,j}$ le terme de la matrice inverse de la matrice de Vandermonde situé sur la $i^{ème}$ ligne et la $j^{ème}$ colonne ; il s'obtient, à partir de $Q_{i,j}$, par une transposition (inversion des rôles de i et j), une multiplication par $(-1)^{i+j}$, et une division par le déterminant $D(\lambda_1, \lambda_2, ..., \lambda_n)$.

$$\alpha_{i,j} = (-1)^{i+j} \frac{\sigma_{n-i}(\lambda_1, \lambda_2, ..., \lambda_{j-1}, \lambda_{j+1}, ..., \lambda_n) D(\lambda_1, \lambda_2, ..., \lambda_{j-1}, \lambda_{j+1}, ..., \lambda_n}{D(\lambda_1, \lambda_2, ..., \lambda_n)}.$$

$$\alpha_{i,j} = (-1)^{i+j} \frac{\sigma_{n-i}(\lambda_1, \lambda_2, ..., \lambda_{j-1}, \lambda_{j+1}, ..., \lambda_n)}{(-1)^{n-j} \prod\limits_{k \in \{1,...,n\}-\{j\}} (\lambda_j - \lambda_k)}$$

$$(23)\ \alpha_{i,j} = (-1)^{n-i} \frac{\sigma_{n-i}(\lambda_1, \lambda_2, ..., \lambda_{j-1}, \lambda_{j+1}, ..., \lambda_n)}{\prod\limits_{k \in \{1,...,n\}-\{j\}} (\lambda_j - \lambda_k)}.$$

Par exemple, pour n=3, on obtient :

$$\begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 \\ 1 & \lambda_2 & \lambda_2^2 \\ 1 & \lambda_3 & \lambda_3^2 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{(\lambda_2\lambda_3)}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} & \frac{(\lambda_1\lambda_3)}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} & \frac{(\lambda_1\lambda_2)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} \\ \frac{-(\lambda_2+\lambda_3)}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} & \frac{-(\lambda_1+\lambda_3)}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} & \frac{-(\lambda_1+\lambda_2)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} \\ \frac{1}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} & \frac{1}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} & \frac{1}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} \end{pmatrix}$$

...et pour n=4 :

$$\begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \lambda_1^3 \\ 1 & \lambda_2 & \lambda_2^2 & \lambda_2^3 \\ 1 & \lambda_3 & \lambda_3^2 & \lambda_3^3 \\ 1 & \lambda_4 & \lambda_4^2 & \lambda_4^3 \end{pmatrix}^{-1} =$$

$$\begin{pmatrix} \frac{-(\lambda_2\lambda_3\lambda_4)}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)(\lambda_1-\lambda_4)} & \frac{-(\lambda_1\lambda_3\lambda_4)}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)(\lambda_2-\lambda_4)} & \frac{-(\lambda_1\lambda_2\lambda_4)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)(\lambda_3-\lambda_4)} & \frac{-(\lambda_1\lambda_2\lambda_3)}{(\lambda_4-\lambda_1)(\lambda_4-\lambda_2)(\lambda_4-\lambda_3)} \\ \frac{(\lambda_2\lambda_3+\lambda_2\lambda_4+\lambda_3\lambda_4)}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)(\lambda_1-\lambda_4)} & \frac{(\lambda_1\lambda_3+\lambda_1\lambda_4+\lambda_3\lambda_4)}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)(\lambda_2-\lambda_4)} & \frac{(\lambda_1\lambda_2+\lambda_1\lambda_4+\lambda_2\lambda_4)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)(\lambda_3-\lambda_4)} & \frac{(\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_2\lambda_3)}{(\lambda_4-\lambda_1)(\lambda_4-\lambda_2)(\lambda_4-\lambda_3)} \\ \frac{-(\lambda_2+\lambda_3+\lambda_4)}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)(\lambda_1-\lambda_4)} & \frac{-(\lambda_1+\lambda_3+\lambda_4)}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)(\lambda_2-\lambda_4)} & \frac{-(\lambda_1+\lambda_2+\lambda_4)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)(\lambda_3-\lambda_4)} & \frac{-(\lambda_1+\lambda_2+\lambda_3)}{(\lambda_4-\lambda_1)(\lambda_4-\lambda_2)(\lambda_4-\lambda_3)} \\ \frac{1}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)(\lambda_1-\lambda_4)} & \frac{1}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)(\lambda_2-\lambda_4)} & \frac{1}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)(\lambda_3-\lambda_4)} & \frac{1}{(\lambda_4-\lambda_1)(\lambda_4-\lambda_2)(\lambda_4-\lambda_3)} \end{pmatrix}$$

Considérons maintenant la matrice $\begin{pmatrix} \lambda_1^{n-1} & \lambda_2^{n-1} & ... & \lambda_n^{n-1} \\ ... & ... & ... & ... \\ \lambda_1^2 & \lambda_2^2 & ... & \lambda_n^2 \\ \lambda_1 & \lambda_2 & ... & \lambda_n \\ 1 & 1 & ... & 1 \end{pmatrix}$, obtenue à partir d'une matrice de Vandermonde par une transposition, suivie d'une inversion de l'ordre des lignes (ou par une inversion de l'ordre des colonnes, suivie d'une transposition) ; son inverse s'obtient, à partir de l'inverse de la matrice de Vandermonde, par une inversion de l'ordre des lignes, suivie d'une transposition (ou par une transposition, suivie d'une inversion de l'ordre des colonnes). Soit $\beta_{i,j}$ le coefficient situé sur la $i^{ème}$ ligne et la $j^{ème}$ colonne ; on a :

(24) $$\beta_{i,j} = \alpha_{n-j+1,i} = (-1)^{j-1} \frac{\sigma_{j-1}(\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_n)}{\prod_{k\in\{1,...,n\}-\{i\}} (\lambda_i - \lambda_k)}.$$

Par exemple, pour n=3, on obtient :

$$\begin{pmatrix} \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ 1 & 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} & \frac{-(\lambda_2+\lambda_3)}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} & \frac{(\lambda_2\lambda_3)}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} \\ \frac{1}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} & \frac{-(\lambda_1+\lambda_3)}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} & \frac{(\lambda_1\lambda_3)}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} \\ \frac{1}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} & \frac{-(\lambda_1+\lambda_2)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} & \frac{(\lambda_1\lambda_2)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} \end{pmatrix}.$$

5 Equations algébriques de degré quelconque

Considérons l'équation de degré n suivante :

$x^n = a_1x^{n-1} + a_2x^{n-2} + ... + a_{n-1}x + a_n$

(dans laquelle les coefficients a_i sont des nombres complexes) ; on a alors (formules de Viète) : $a_i = (-1)^{i+1}\sigma_i(\lambda_1, ..., \lambda_n)$ (où λ_1,..., λ_n sont les racines de l'équation, réelles ou complexes, distinctes ou non).

Posons : M= $\begin{pmatrix} a_1 & a_2 & a_3 & ... & a_{n-1} & a_n \\ 1 & 0 & 0 & ... & 0 & 0 \\ 0 & 1 & 0 & ... & 0 & 0 \\ 0 & 0 & 1 & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 & 0 \end{pmatrix} = \begin{pmatrix} \sigma_1 & -\sigma_2 & \sigma_3 & ... & (-1)^{n-1}\sigma_n \\ 1 & 0 & 0 & ... & 0 \\ 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 0 \end{pmatrix}.$

M est la "matrice compagnon" du polynôme $P(x) = x^n - a_1x^{n-1} - a_2x^{n-2} - ... - a_{n-1}x - a_n$.

On peut vérifier facilement que le déterminant de cette matrice est égal à $(-1)^{n-1}a_n$, donc : $det\ M = \sigma_n = \lambda_1\lambda_2...\lambda_n$.

La matrice M est inversible si toutes les racines sont non nulles. Dans ce cas, on a :

$$M^{-1} = \frac{1}{a_n}\begin{pmatrix} 0 & a_n & 0 & ... & 0 & 0 \\ 0 & 0 & a_n & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & a_n & 0 \\ 0 & 0 & 0 & ... & 0 & a_n \\ 1 & -a_1 & -a_2 & ... & -a_{n-2} & -a_{n-1} \end{pmatrix}$$

$$M^{-1} = \begin{pmatrix} 0 & 1 & 0 & ... & 0 & 0 \\ 0 & 0 & 1 & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 & 0 \\ 0 & 0 & 0 & ... & 0 & 1 \\ \frac{1}{\sigma_n} & -\frac{\sigma_1}{\sigma_n} & \frac{\sigma_2}{\sigma_n} & ... & (-1)^{n-2}\frac{\sigma_{n-2}}{\sigma_n} & (-1)^{n-1}\frac{\sigma_{n-1}}{\sigma_n} \end{pmatrix}.$$

On peut remarquer que $\dfrac{\sigma_i(\lambda_1,...,\lambda_n)}{\sigma_n(\lambda_1,...,\lambda_n)} = \dfrac{\sigma_i(\lambda_1,...,\lambda_n)}{\lambda_1\lambda_2\lambda_3...\lambda_n} = \sigma_{n-i}\left(\dfrac{1}{\lambda_1},\dfrac{1}{\lambda_2},\dfrac{1}{\lambda_3},...,\dfrac{1}{\lambda_n}\right)$.

D'autre part, définissons la suite (u_i) par la relation de récurrence :

$$\begin{pmatrix} u_i \\ u_{i-1} \\ u_{i-2} \\ ... \\ u_{i-n+1} \end{pmatrix} = M \begin{pmatrix} u_{i-1} \\ u_{i-2} \\ u_{i-3} \\ ... \\ u_{i-n} \end{pmatrix}$$, et par les valeurs initiales :

$u_0 = 1$, et $u_i = 0$ pour $i < 0$ (valeurs arbitraires, qui peuvent être éventuellement modifiées).

On a donc :

$$\begin{pmatrix} u_p \\ u_{p-1} \\ u_{p-2} \\ ... \\ u_{p-n+1} \end{pmatrix} = M^p \begin{pmatrix} u_0 \\ u_{-1} \\ u_{-2} \\ ... \\ u_{-n+1} \end{pmatrix} = M^p \begin{pmatrix} 1 \\ 0 \\ 0 \\ ... \\ 0 \end{pmatrix} \tag{25}$$

Nous allons diagonaliser la matrice M ; pour cela, cherchons ses valeurs propres ; ce sont les racines de l'équation :

$$\Delta = \begin{vmatrix} a_1-\lambda & a_2 & a_3 & ... & a_{n-1} & a_n \\ 1 & -\lambda & 0 & ... & 0 & 0 \\ 0 & 1 & -\lambda & ... & 0 & 0 \\ 0 & 0 & 1 & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 & -\lambda \end{vmatrix} = 0\,;$$

Montrons, par récurrence, que $\Delta = (-1)^n[\lambda^n - (a_1\lambda^{n-1} + a_2\lambda^{n-2} + a_3\lambda^{n-3} + ... + a_n)]$ (formule vraie pour n=1) ; développons le déterminant à partir de la dernière colonne, et supposons la formule vraie pour tout déterminant du même type, de degré inférieur à n :

$\Delta = -\lambda(-1)^{n-1}[\lambda^{n-1} - (a_1\lambda^{n-2} + a_2\lambda^{n-3} + a_3\lambda^{n-4} + ... + a_n)] + (-1)^{n-1}a_n$; ce qui équivaut bien à la formule annoncée.

Notre équation devient donc :

$(-1)^n[\lambda^n - (a_1\lambda^{n-1} + a_2\lambda^{n-2} + a_3\lambda^{n-3} + ... + a_n)] = 0$;

ses solutions sont λ_1,...,λ_n ; ce sont les valeurs propres de M. Nous suppose-

rons, pour le moment, qu'elles sont toutes distinctes. Dans ce cas, les vecteurs propres constituent une base.

Notons $x_0, x_1, ..., x_{n-1}$ les coordonnées d'un vecteur propre associé, par exemple, à la valeur propre λ_1. On doit avoir :

$$\begin{pmatrix} a_1 & a_2 & a_3 & ... & a_{n-1} & a_n \\ 1 & 0 & 0 & ... & 0 & 0 \\ 0 & 1 & 0 & ... & 0 & 0 \\ 0 & 0 & 1 & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 & 0 \end{pmatrix} \begin{pmatrix} x_{n-1} \\ x_{n-2} \\ x_{n-3} \\ x_{n-4} \\ ... \\ x_0 \end{pmatrix} = \lambda_1 \begin{pmatrix} x_{n-1} \\ x_{n-2} \\ x_{n-3} \\ x_{n-4} \\ ... \\ x_0 \end{pmatrix};$$

En commençant par la dernière ligne, nous déduisons :

$x_1 = \lambda_1 x_0$;
$x_2 = \lambda_1 x_1 = \lambda_1^2 x_0$;
$x_3 = \lambda_1 x_2 = \lambda_1^3 x_0$;
...
$x_{n-1} = \lambda_1 x_{n-2} = \lambda_1^{n-1} x_0$

La première ligne ne fait que confirmer :

$$a_1 x_{n-1} + a_2 x_{n-2} + ... + a_n x_0 = (a_1 \lambda_1^{n-1} + a_2 \lambda_1^{n-2} + ... + a_n) x_0 = \lambda_1^n x_0 = \lambda_1 x_{n-1}.$$

Donc les vecteurs propres associés à λ_1 sont proportionnels à $\overrightarrow{u_1} = \begin{pmatrix} \lambda_1^{n-1} \\ \lambda_1^{n-2} \\ ... \\ \lambda_1^3 \\ \lambda_1^2 \\ \lambda_1 \\ 1 \end{pmatrix}$.

Avant de continuer, faisons quelques remarques :

- le polynôme $P(x) = x^n - a_1 x^{n-1} - a_2 x^{n-2} - ... - a_{n-1} x - a_n$ est le polynôme caractéristique de M ;

- d'après le théorème de Cayley-Hamilton, on a l'égalité matricielle : $P(M) = M^n - a_1 M^{n-1} - a_2 M^{n-2} - ... - a_{n-1} M - a_n I_n = 0_n$, où I_n est la matrice identité d'ordre n, et 0_n la matrice nulle d'ordre n ;

- dans le cas présent, cette égalité est facile à comprendre : calculons par exemple $P(M)(\overrightarrow{u_1})$, image d'un vecteur propre par $P(M)$:
sachant que $M(\overrightarrow{u_1}) = \lambda_1 \overrightarrow{u_1}$, $M^2(\overrightarrow{u_1}) = \lambda_1^2 \overrightarrow{u_1}$, $M^3(\overrightarrow{u_1}) = \lambda_1^3 \overrightarrow{u_1}$, etc., nous pouvons écrire :

$P(M)(\overrightarrow{u_1}) = M^n(\overrightarrow{u_1}) - a_1 M^{n-1}(\overrightarrow{u_1}) - a_2 M^{n-2}(\overrightarrow{u_1}) - ... - a_{n-1} M(\overrightarrow{u_1}) - a_n \overrightarrow{u_1}$
$= \lambda_1^n \overrightarrow{u_1} - a_1 \lambda_1^{n-1} \overrightarrow{u_1} - a_2 \lambda_1^{n-2} \overrightarrow{u_1} - ... - a_{n-1} \lambda_1^{n-1} \overrightarrow{u_1} - a_n \overrightarrow{u_1}$
$= (\lambda_1^n - (a_1 \lambda_1^{n-1} + a_2 \lambda_1^{n-2} + a_3 \lambda_1^{n-3} + ... + a_n)) \overrightarrow{u_1} = 0 \overrightarrow{u_1} = \overrightarrow{0}$.
De même, les images de tous les autres vecteurs propres par $P(M)$ s'annulent ; si les vecteurs propres constituent une base, alors on peut en déduire que $P(M) = 0_n$.

Pour passer de la base de vecteurs propres à la base canonique, on utilise la matrice :

$$\begin{pmatrix} \lambda_1^{n-1} & \lambda_2^{n-1} & ... & \lambda_{n-1}^{n-1} & \lambda_n^{n-1} \\ \lambda_1^{n-2} & \lambda_2^{n-2} & ... & \lambda_{n-1}^{n-2} & \lambda_n^{n-2} \\ ... & ... & ... & ... & ... \\ \lambda_1^2 & \lambda_2^2 & ... & \lambda_{n-1}^2 & \lambda_n^2 \\ \lambda_1 & \lambda_2 & ... & \lambda_{n-1} & \lambda_n \\ 1 & 1 & ... & 1 & 1 \end{pmatrix}.$$

Si les racines λ_i sont toutes distinctes, comme nous l'avons supposé, cette matrice est inversible, et son inverse se calcule selon la technique vue à la fin du paragraphe précédent (formule 24).

Par exemple, pour $n = 3$, on obtient l'égalité :

$$M = \begin{pmatrix} \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} \frac{1}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} & \frac{-(\lambda_2+\lambda_3)}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} & \frac{(\lambda_2\lambda_3)}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} \\ \frac{1}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} & \frac{-(\lambda_1+\lambda_3)}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} & \frac{(\lambda_1\lambda_3)}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} \\ \frac{1}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} & \frac{-(\lambda_1+\lambda_2)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} & \frac{(\lambda_1\lambda_2)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} \end{pmatrix},$$

ce qui entraîne :

$$M^p = \begin{pmatrix} \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1^p & 0 & 0 \\ 0 & \lambda_2^p & 0 \\ 0 & 0 & \lambda_3^p \end{pmatrix} \begin{pmatrix} \frac{1}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} & \frac{-(\lambda_2+\lambda_3)}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} & \frac{(\lambda_2\lambda_3)}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} \\ \frac{1}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} & \frac{-(\lambda_1+\lambda_3)}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} & \frac{(\lambda_1\lambda_3)}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} \\ \frac{1}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} & \frac{-(\lambda_1+\lambda_2)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} & \frac{(\lambda_1\lambda_2)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} \end{pmatrix}.$$

Ce produit est facile à développer. Pour une équation de degré n, la matrice M^p a pour terme général :

$$(26)\ \gamma_{i,j}(p) = (-1)^{j-1} \sum_{k=1}^{n} \frac{\lambda_k^{p+n-i} \sigma_{j-1}(\lambda_1, \lambda_2, ..., \lambda_{k-1}, \lambda_{k+1}, ..., \lambda_n)}{\prod_{l \in \{1,...,n\}-\{k\}} (\lambda_k - \lambda_l)}.$$

Pour les termes de la première colonne (j=1), la formule se simplifie :

$$\gamma_{i,1}(p) = \sum_{k=1}^{n} \frac{\lambda_k^{p+n-i}}{\prod\limits_{l\in\{1,\dots,n\}-\{k\}} (\lambda_k - \lambda_l)} = \tau_{p+1-i}(\lambda_1, \lambda_2, ..., \lambda_n).$$

Signalons au passage une conséquence intéressante : si on pose : $u_0 = 1$ et $u_i = 0$ pour $i < 0$, alors on a $u_p = \tau_p$ pour tout $p \geq 0$.

Dans la formule (26), remplaçons $\sigma_{j-1}(\lambda_1, \lambda_2, ..., \lambda_{k-1}, \lambda_{k+1}, ..., \lambda_n)$ par sa valeur, tirée de (19), que nous écrivons sous la forme : $\sum\limits_{q=1}^{j} (-1)^{q-1} \lambda_k^{q-1} \sigma_{j-q}(\lambda_1, ..., \lambda_n)$; nous obtenons :

$$\gamma_{i,j}(p) = (-1)^{j-1} \sum_{k=1}^{n} \left[\frac{\lambda_k^{p+n-i}}{\prod\limits_{l\in\{1,\dots,n\}-\{k\}} (\lambda_k - \lambda_l)} \sum_{q=1}^{j} (-1)^{q-1} \lambda_k^{q-1} \sigma_{j-q}(\lambda_1, ..., \lambda_n) \right]$$

$$\gamma_{i,j}(p) = \sum_{q=1}^{j} \left[(-1)^{j-q} \sigma_{j-q}(\lambda_1, ..., \lambda_n) \sum_{k=1}^{n} \frac{\lambda_k^{p+n+q-i-1}}{\prod\limits_{l\in\{1,\dots,n\}-\{k\}} (\lambda_k - \lambda_l)} \right]$$

Nous remarquons que la seconde somme est égale à $\tau_{p+q-i}(\lambda_1, ..., \lambda_n)$, d'après la formule (17). D'où :

$$(27)\ \gamma_{i,j}(p) = \sum_{q=1}^{j} \left[(-1)^{j-q} \sigma_{j-q}^{(n)} \tau_{p+q-i}^{(n)} \right].$$

Pour la première colonne (j=1), on retrouve facilement la simplification signalée précédemment.

Pour n=3, par exemple, la formule (27) donne :

$$M^p = \begin{pmatrix} \tau_p^{(3)} & -\sigma_2^{(3)}\tau_{p-1}^{(3)} + \sigma_3^{(3)}\tau_{p-2}^{(3)} & \sigma_3^{(3)}\tau_{p-1}^{(3)} \\ \tau_{p-1}^{(3)} & -\sigma_2^{(3)}\tau_{p-2}^{(3)} + \sigma_3^{(3)}\tau_{p-3}^{(3)} & \sigma_3^{(3)}\tau_{p-2}^{(3)} \\ \tau_{p-2}^{(3)} & -\sigma_2^{(3)}\tau_{p-3}^{(3)} + \sigma_3^{(3)}\tau_{p-4}^{(3)} & \sigma_3^{(3)}\tau_{p-3}^{(3)} \end{pmatrix}$$

Pour n=4, on obtient :

$$M^p = \begin{pmatrix} \tau_p^{(4)} & -\sigma_2^{(4)}\tau_{p-1}^{(4)} + \sigma_3^{(4)}\tau_{p-2}^{(4)} - \sigma_4^{(4)}\tau_{p-3}^{(4)} & \sigma_3^{(4)}\tau_{p-1}^{(4)} - \sigma_4^{(4)}\tau_{p-2}^{(4)} & -\sigma_4^{(4)}\tau_{p-1}^{(4)} \\ \tau_{p-1}^{(4)} & -\sigma_2^{(4)}\tau_{p-2}^{(4)} + \sigma_3^{(4)}\tau_{p-3}^{(4)} - \sigma_4^{(4)}\tau_{p-4}^{(4)} & \sigma_3^{(4)}\tau_{p-2}^{(4)} - \sigma_4^{(4)}\tau_{p-3}^{(4)} & -\sigma_4^{(4)}\tau_{p-2}^{(4)} \\ \tau_{p-2}^{(4)} & -\sigma_2^{(4)}\tau_{p-3}^{(4)} + \sigma_3^{(4)}\tau_{p-4}^{(4)} - \sigma_4^{(4)}\tau_{p-5}^{(4)} & \sigma_3^{(4)}\tau_{p-3}^{(4)} - \sigma_4^{(4)}\tau_{p-4}^{(4)} & -\sigma_4^{(4)}\tau_{p-3}^{(4)} \\ \tau_{p-3}^{(4)} & -\sigma_2^{(4)}\tau_{p-4}^{(4)} + \sigma_3^{(4)}\tau_{p-5}^{(4)} - \sigma_4^{(4)}\tau_{p-6}^{(4)} & \sigma_3^{(4)}\tau_{p-4}^{(4)} - \sigma_4^{(4)}\tau_{p-5}^{(4)} & -\sigma_4^{(4)}\tau_{p-4}^{(4)} \end{pmatrix}$$

Nous aurions pu démontrer la formule (27) plus simplement, par récurrence : elle est évidente pour p=1 ; ensuite, on la suppose vraie pour une valeur quelconque de p, et on développe le produit $M.M^p$. En utilisant de manière répétée la conséquence de la formule (10) $\left(\text{c'est-à-dire : } \tau_m^{(n)} = \sum_{i=1}^{n} \left[(-1)^{i+1}\sigma_i^{(n)}\tau_{m-i}^{(n)}\right]\right)$, on montre qu'elle est vraie pour $p+1$.

Ceci prouve que la formule (27), contrairement à la formule (26), reste valable lorsque l'équation a des racines multiples.

Mais la formule (26) permet d'étudier facilement le comportement de M^p quand p tend vers l'infini ; de plus, si la matrice M est inversible (donc si les racines λ_i sont toutes non nulles), elle est valable pour tout $p \in Z$, alors que la formule (27) suppose $p > 0$.

Supposons que l'une des racines (λ_s) ait un module strictement supérieur aux autres ; il en résulte que $\left(\frac{\lambda_k}{\lambda_s}\right)^p$ tend vers 1 (pour $k = s$) ou vers 0 (pour $k \neq s$) quand p tend vers l'infini ; donc la matrice $\left(\frac{1}{\lambda_s}M\right)^p$ converge :

$$\lim_{p\to+\infty} \frac{\gamma_{i,j}(p)}{\lambda_s^p} = (-1)^{j-1} \frac{\lambda_s^{n-i}\sigma_{j-1}(\lambda_1, \lambda_2, ..., \lambda_{s-1}, \lambda_{s+1}, ..., \lambda_n)}{\prod_{l\in\{1,...,n\}-\{s\}} (\lambda_s - \lambda_l)}.$$

On remarque, en particulier, que :

$$\lim_{p\to+\infty} \frac{\gamma_{i,j}(p)}{\gamma_{i+1,j}(p)} = \lambda_s \,;$$

$$\lim_{p\to+\infty} \frac{\gamma_{i,j}(p)}{\gamma_{i,1}(p)} = (-1)^{j-1}\sigma_{j-1}(\lambda_1, \lambda_2, ..., \lambda_{s-1}, \lambda_{s+1}, ..., \lambda_n).$$

Si la matrice M est inversible, on peut faire tendre n vers $-\infty$. Supposons que l'une des racines (λ_t) ait un module strictement inférieur aux autres ; on a alors :

$$\lim_{p\to-\infty}\frac{\gamma_{i,j}(p)}{\lambda_t^p}=(-1)^{j-1}\frac{\lambda_t^{n-i}\sigma_{j-1}(\lambda_1,\lambda_2,...,\lambda_{t-1},\lambda_{t+1},...,\lambda_n)}{\prod\limits_{l\in\{1,...,n\}-\{t\}}(\lambda_s-\lambda_l)}.$$

$$\lim_{p\to-\infty}\frac{\gamma_{i,j}(p)}{\gamma_{i+1,j}(p)}=\lambda_t\,;$$

$$\lim_{p\to-\infty}\frac{\gamma_{i,j}(p)}{\gamma_{i,1}(p)}=(-1)^{j-1}\sigma_{j-1}(\lambda_1,\lambda_2,...,\lambda_{t-1},\lambda_{t+1},...,\lambda_n).$$

L'étude de M^p quand p tend vers $\pm\infty$ permet donc de calculer des valeurs approchées de λ_s (racine ayant le plus grand module) ou de λ_t (racine ayant le plus petit module), et de $\sigma_{j-1}(\lambda_1,\lambda_2,...,\lambda_{s-1},\lambda_{s+1},...,\lambda_n)$ ou $\sigma_{j-1}(\lambda_1,\lambda_2,...,\lambda_{t-1},\lambda_{t+1},...,\lambda_n)$.

On va pouvoir utiliser la famille $(\sigma_{j-1}(\lambda_1,\lambda_2,...,\lambda_{s-1},\lambda_{s+1},...,\lambda_n))_{1\le j\le n}$ ou $(\sigma_{j-1}(\lambda_1,\lambda_2,...,\lambda_{t-1},\lambda_{t+1},...,\lambda_n))_{1\le j\le n}$ pour former une nouvelle matrice, d'ordre n-1, qui va permettre de calculer la racine suivante, et ainsi de suite (à condition que toutes les racines aient des modules distincts).

Revenons à la formule (25). Le premier terme de M^p (en haut à gauche) est : $\gamma_{1,1}=\tau_p(\lambda_1,...,\lambda_n)$, donc : $u_p=\tau_p(\lambda_1,...,\lambda_n)$.

Ceci nous permet de calculer le "ratio" q_n :

$$q_n=\frac{u_{p+1}}{u_p}=\frac{\tau_{p+1}(\lambda_1,...,\lambda_n)}{\tau_p(\lambda_1,...,\lambda_n)}=\frac{\sum\limits_{i=1}^{n}\dfrac{\lambda_i^{p+n}}{\prod\limits_{j\in\{1,...,n\}-\{i\}}(\lambda_i-\lambda_j)}}{\sum\limits_{i=1}^{n}\dfrac{\lambda_i^{p+n-1}}{\prod\limits_{j\in\{1,...,n\}-\{i\}}(\lambda_i-\lambda_j)}}=\frac{\sum\limits_{i=1}^{n}\lambda_i^{p+n}}{\sum\limits_{i=1}^{n}\lambda_i^{p+n-1}}$$ (d'après la formule 17).

$$q_n=\frac{\lambda_s^{p+n}\sum\limits_{i=1}^{n}(\frac{\lambda_i}{\lambda_s})^{p+n}}{\lambda_s^{p+n-1}\sum\limits_{i=1}^{n}(\frac{\lambda_i}{\lambda_s})^{p+n-1}}$$

Quand p tend vers l'infini, $(\frac{\lambda_i}{\lambda_s})^{p+n}$ et $(\frac{\lambda_i}{\lambda_s})^{p+n-1}$ tendent vers 1 pour $i=s$, et vers 0 pour $i\neq s$; on en déduit que q_n tend vers λ_s quand p tend vers $+\infty$. On peut montrer de manière analogue que q_n tend vers λ_t quand p tend vers $-\infty$ (si M est inversible).

Jusqu'ici, nous avons supposé que les racines (λ_i) avaient des modules tous distincts ; imaginons maintenant que la racine ayant le plus grand module soit

double : $\lambda_s = \lambda_{s'}$. Dans ce cas, $(\frac{\lambda_i}{\lambda_s})^{p+n}$ et $(\frac{\lambda_i}{\lambda_s})^{p+n-1}$ tendent vers 1 pour $i = s$ et pour $i = s'$, et vers 0 sinon ; on en déduit que q_n tend encore vers λ_s. Il en est de même lorsque la racine de plus grand module est multiple, d'ordre quelconque.

Mais s'il y a plusieurs racines de plus grand module non égales (par exemple : $\lambda_s = \rho e^{i\theta_1}$ et $\lambda_{s'} = \rho e^{i\theta_2}$, avec $\theta_1 \neq \theta_2$ modulo π), alors la suite q_n diverge.

Cherchons maintenant une condition pour que l'une des solutions de l'équation soit un PV-number.

Introduisons ici une nouvelle famille de fonctions symétriques des racines : il s'agit de la famille (s_p), définie par :

$$s_p(\lambda_1, ..., \lambda_n) = \lambda_1^p + ... + \lambda_n^p = \sum_{i=1}^{n} \lambda_i^p.$$

Ces fonctions s_p s'expriment comme polynômes en σ_i à coefficients entiers ; les formules de Newton-Girard (voir paragraphe suivant) le montrent clairement. On pourrait d'ailleurs démontrer, très généralement, que, si un polynôme en λ_i est symétrique et à coefficients entiers, alors il s'exprime comme polynôme en σ_i à coefficients entiers (il suffit de reprendre la démonstration du théorème fondamental des fonctions symétriques, et de vérifier qu'à chaque étape de la transformation, les coefficients restent entiers).

Supposons qu'une équation de la forme $X^n = a_1 X^{n-1} + a_2 X^{n-2} + ... + a_n$ ait des coefficients (a_i) tous entiers ; il en résulte que $\sigma_i(\lambda_1, ..., \lambda_n)$ et $s_i(\lambda_1, ..., \lambda_n)$ sont des entiers, quel que soit i. Supposons de plus qu'une racine λ_s ait un module supérieur à 1, et que toutes les autres aient un module inférieur à 1. Dans ce cas, $s_p = \lambda_1^p + ... + \lambda_n^p$ est équivalent à λ_s^p quand p tend vers l'infini (car tous les autres termes tendent vers 0) ; donc λ_s est soit un entier, soit un PV-number.

6 Formules de Newton-Girard et de Brioschi

Nous aurions pu aborder ce formules dans le paragraphe II (fonctions symétriques) ; mais comme elles ont un rôle marginal dans notre travail, nous les traitons seulement en appendice.

Elles font intervenir la famille de fonctions symétriques des racines (s_p) :

$$s_p(\lambda_1, ..., \lambda_n) = \lambda_1^p + ... + \lambda_n^p = \sum_{i=1}^{n} \lambda_i^p.$$

Posons : $\overrightarrow{v_i} = \begin{pmatrix} \lambda_i^{n-1} \\ \lambda_i^{n-2} \\ ... \\ \lambda_i \\ 1 \end{pmatrix}$.

C'est un vecteur propre de la matrice M, et il vérifie :

$$M.\overrightarrow{v_i} = \lambda_i \overrightarrow{v_i} = \begin{pmatrix} \lambda_i^{n} \\ \lambda_i^{n-1} \\ ... \\ \lambda_i^{2} \\ \lambda_i \end{pmatrix}, \text{ et : } M^p.\overrightarrow{v_i} = \lambda_i^p \overrightarrow{v_i} = \begin{pmatrix} \lambda_i^{n-1+p} \\ \lambda_i^{n-2+p} \\ ... \\ \lambda_i^{1+p} \\ \lambda_i^{p} \end{pmatrix}.$$

Considérons maintenant le vecteur $\overrightarrow{w} = \overrightarrow{v_i} + ... + \overrightarrow{v_n}$ (qui n'est pas un vecteur propre, mais la somme de tous les vecteurs propres). On a :

$$\overrightarrow{w} = \begin{pmatrix} s_{n-1} \\ s_{n-2} \\ ... \\ s_1 \\ s_0 \end{pmatrix}, \text{ et :}$$

$$M^p.\overrightarrow{w} = M^p.\overrightarrow{v_i} + ... + M^p.\overrightarrow{v_n} = \begin{pmatrix} \lambda_1^{n-1+p} + ... + \lambda_n^{n-1+p} \\ \lambda_1^{n-2+p} + ... + \lambda_n^{n-2+p} \\ ... \\ \lambda_1^{1+p} + ... + \lambda_n^{1+p} \\ \lambda_i^{p} + ... + \lambda_i^{p} \end{pmatrix} = \begin{pmatrix} s_{n-1+p} \\ s_{n-2+p} \\ ... \\ s_{1+p} \\ s_p \end{pmatrix};$$

d'où on tire :

$$\begin{pmatrix} s_{n-1+p} \\ s_{n-2+p} \\ ... \\ s_{1+p} \\ s_p \end{pmatrix} = M^p. \begin{pmatrix} s_{n-1} \\ s_{n-2} \\ ... \\ s_1 \\ s_0 \end{pmatrix}.$$

En remplaçant M^p par sa valeur donnée par la formule (26) ou (27), on voit qu'il est possible de calculer tous les termes de la suite (s_i) en fonction des n premiers, à condition d'avoir calculé au préalable les puissances de M.

On peut aussi se contenter d'établir une formule de récurrence, en écrivant, par exemple (pour $p \geq n$) :

$$\begin{pmatrix} s_p \\ s_{p-1} \\ ... \\ s_{p+2-n} \\ s_{p+1-n} \end{pmatrix} = M. \begin{pmatrix} s_{p-1} \\ s_{p-2} \\ ... \\ s_{p+1-n} \\ s_{p-n} \end{pmatrix} = \begin{pmatrix} \sigma_1 & -\sigma_2 & \sigma_3 & ... & (-1)^{n-1}\sigma_n \\ 1 & 0 & 0 & ... & 0 \\ 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 0 \end{pmatrix} \begin{pmatrix} s_{p-1} \\ s_{p-2} \\ ... \\ s_{p+1-n} \\ s_{p-n} \end{pmatrix}.$$

Par conséquent :

$s_p = \sigma_1 s_{p-1} - \sigma_2 s_{p-2} + \sigma_3 s_{p-3} - \sigma_4 s_{p-4} + ... + (-1)^{n-1}\sigma_n s_{p-n}$

(28) $s_p(\lambda_1, ..., \lambda_n) = \sum_{i=1}^{n} (-1)^{i-1}\sigma_i(\lambda_1, ..., \lambda_n) s_{p-i}(\lambda_1, ..., \lambda_n)$ $(p \geq n)$.

Une autre méthode très simple pour démontrer cette formule consiste à partir de l'équation : $x^n = \sigma_1 x^{n-1} - \sigma_2 x^{n-2} + \sigma_3 x^{n-3} - ... + (-1)^{n-1}\sigma_n$; en remplaçant successivement x par chacune des racines $(\lambda_1, ..., \lambda_n)$, on obtient n égalités, qu'on ajoute membre à membre. Puis on multiplie les deux membres par x^{p-n}.

Le problème est que cette formule n'est valable que pour $p \geq n$; pour pouvoir l'utiliser, il faut calculer d'abord s_0, s_1,..., s_{p-1}. Pour cela, on utilise une formule légèrement différente (mais plus difficile à démontrer) :

(29) $s_p = \sum_{i=1}^{p-1} (-1)^{i-1}\sigma_i s_{p-i} + (-1)^{p-1} p\sigma_p$.

On part de : $s_0 = n$, $s_1 = \sigma_1$; ensuite, on peut calculer successivement : $s_2 = s_1\sigma_1 - 2\sigma_2$, $s_3 = s_2\sigma_1 - s_1\sigma_2 + 3\sigma_3$, etc.

Contrairement à la formule 28, la formule 29 (formule de Newton-Girard) est valable pour tout entier p supérieur ou égal à 2. (Pour $p > n$, on a : $\sigma_p = 0$, donc le dernier terme disparaît.)

Pour ceux que cela intéresse, nous en donnons ci-dessous une démonstration connue, plus analytique qu'algébrique :

L'idée est de partir des fonctions à variable complexe $f(z) = Log \prod_{i=1}^{n} \frac{1}{1-\lambda_i z}$ et $g(z) = \sum_{i=1}^{n} Log \frac{1}{1-\lambda_i z}$; ces deux fonctions sont évidemment égales (propriété du logarithme), mais on va les développer en séries de deux manières différentes, puis les dériver, et écrire que les dérivées sont égales.

$f(z) = -Log \prod_{i=1}^{n} (1-\lambda_i z) = -Log\ h(z)$, avec :

$h(z) = \prod_{i=1}^{n} (1-\lambda_i z) = \sum_{j=0}^{+\infty} (-1)^j \sigma_j z^j$;

$h'(z) = \sum_{j=1}^{+\infty} (-1)^j \sigma_j j z^{j-1}$;

$$f'(z) = -\frac{h'(z)}{h(z)} = \frac{\sum_{j=1}^{+\infty}(-1)^{j-1}\sigma_j j z^{j-1}}{\sum_{j=0}^{+\infty}(-1)^j \sigma_j z^j}.$$

Pour développer $g(z)$, on utilise le développement bien connu : $Log\frac{1}{1-u} = u + \frac{u^2}{2} + \frac{u^3}{3} + ...$:

$$g(z) = \sum_{i=1}^{n}\left(\sum_{j=1}^{+\infty}\frac{\lambda_i^j z^j}{j}\right) = \sum_{j=1}^{+\infty}\left(\frac{z^j}{j}\sum_{i=1}^{n}\lambda_i^j\right) = \sum_{j=1}^{+\infty}\frac{z^j s_j}{j}\ ;$$

$$g'(z) = \sum_{j=1}^{+\infty} s_j z^{j-1}.$$

$f'(z) = g'(z)$, donc :

$$\sum_{p=1}^{+\infty}(-1)^{p-1}p\sigma_p z^{p-1} = \left(\sum_{j=1}^{+\infty} s_j z^{j-1}\right)\left(\sum_{i=0}^{+\infty}(-1)^i\sigma_i z^i\right).$$

Le produit des deux sommes, dans le second membre, se développe selon le principe de la convolution ; on pose : i+j=p ; on obtient :

$$\sum_{p=1}^{+\infty}(-1)^{p-1}p\sigma_p z^{p-1} = \sum_{p=1}^{+\infty}\sum_{i+j=p, j\geq 1} s_j\sigma_i(-1)^i z^{p-1} = \sum_{p=1}^{+\infty}\left(\sum_{i=0}^{p-1}(-1)^i s_{p-i}\sigma_i\right) z^{p-1}$$

Quel que soit $p \geq 1$, les coefficients de z^{p-1} doivent être égaux :

$$(-1)^{p-1}p\sigma_p = \sum_{i=0}^{p-1}(-1)^i s_{p-i}\sigma_i = s_p + \sum_{i=1}^{p-1}(-1)^i s_{p-i}\sigma_i, \text{ autrement dit :}$$

$$s_p = \sum_{i=1}^{p-1}(-1)^{i-1}\sigma_i s_{p-i} + (-1)^{p-1}p\sigma_p.$$

La formule de Brioschi est voisine de celle de Newton-Girard, mais elle fait intervenir τ_i à la place de σ_i :

$$(30)\ s_p = p\tau_p - \sum_{i=1}^{p-1}\tau_i s_{p-i}.$$

Pour la démontrer, reprenons les mêmes fonctions f et g que ci-dessus, mais développons $f(z)$ en utilisant la famille (τ_i) au lieu de (σ_i) :

$$f(z) = Log \prod_{i=1}^{n} \frac{1}{1-\lambda_i z} = Log \sum_{i=0}^{\infty} \tau_i z^i \,;$$

$$f'(z) = \frac{\left(\sum_{i=0}^{\infty} \tau_i z^i\right)'}{\sum_{i=0}^{\infty} \tau_i z^i} = \frac{\sum_{p=1}^{\infty} \tau_p p z^{p-1}}{\sum_{i=0}^{\infty} \tau_i z^i} \,;$$

on a toujours : $g'(z) = \sum_{j=1}^{\infty} s_j z^{j-1}$ et $f'(z) = g'(z)$, donc :

$$\sum_{p=1}^{\infty} p\tau_p z^{p-1} = \left(\sum_{i=0}^{\infty} \tau_i z^i\right)\left(\sum_{j=1}^{\infty} s_j z^{j-1}\right),$$

donc (en utilisant la convolution, et en posant : i+j=p) :

$$\sum_{p=1}^{\infty} p\tau_p z^{p-1} = \sum_{p=1}^{\infty} \left(\sum_{i=0}^{p-1} \tau_i s_{p-i}\right) z^{p-1}.$$

On écrit que les coefficients des deux séries sont égaux, pour tout p :

$p\tau_p = \sum_{i=0}^{p-1} \tau_i s_{p-i}$, ce qui conduit bien à la formule annoncée.