PANORAMAS ET SYNTHÈSES 6

ALGORITHME DE SCHUR, ESPACES À NOYAU REPRODUISANT ET THÉORIE DES SYSTÈMES

Daniel Alpay

Société Mathématique de France 1998

Daniel Alpay
Département de mathématiques, Université Ben–Gurion du Negev, POB 653, 84105 Beer-Sheva, Israel.
E-mail : dany@black.bgu.ac.il
Url : http://www.cs.bgu.ac.il/~dany

***Classification mathématique par sujets* (2000).** — 46E22, 93-02.
Mots clefs. — Théorie des systèmes, interpolation, noyaux reproduisants.

ALGORITHME DE SCHUR, ESPACES À NOYAU REPRODUISANT ET THÉORIE DES SYSTÈMES

Daniel Alpay

***Résumé.* —** Les mêmes fonctions positives (au sens des noyaux reproduisants) apparaissent de manière naturelle dans deux domaines différents, à savoir la modélisation des systèmes linéaires dissipatifs et invariants dans le temps et la théorie des opérateurs linéaires. Nous utilisons les espaces de Hilbert à noyau reproduisant associés à ces fonctions pour étudier les liens entre ces domaines. Le problème de diffusion inverse (inverse scattering problem) joue un rôle central dans les développements. L'approche des noyaux reproduisants permet également d'attaquer de manière naturelle des cas plus généraux, tels que les systèmes non stationnaires, le cas de métriques non positives et celui de paires d'opérateurs non autoadjoints qui commutent.

***Abstract* (The Schur Algorithm, Reproducing Kernel Spaces and System Theory)**

The same positive functions (in the sense of reproducing kernel spaces) appear in a natural way in two different domains, namely the modeling of time–invariant dissipative linear systems and the theory of linear operators. We use the associated reproducing kernel Hilbert spaces to study the relationships between these domains. The inverse scattering problem plays a key role in the exposition. The reproducing kernel approach allows to tackle in a natural way more general cases, such as nonstationary systems, the case of non positive metric and the case of pairs of commuting nonself–adjoint operators.

Pour ma fille Elinor Alpay et mon épouse Liora Alpay

TABLE DES MATIÈRES

CHAPITRE 1

INTRODUCTION

1.1. Introduction

L'objet principal que nous voulons étudier dans ce travail est l'ensemble des fonctions analytiques dont le module est borné par 1 dans le disque unité $\mathbb{D}$. Nous noterons cet ensemble par $\mathcal{S}$ et appellerons ses éléments (pour des raisons qui seront exposées dès les paragraphes suivants) *fonctions de Schur*. Soit donc $s \in \mathcal{S}$. On peut écrire son développement de Taylor à l'origine :

$$s(z) = a_0 + a_1 z + a_2 z^2 + \cdots$$

Bien évidemment la suite $a_0, a_1, \ldots$ définit de manière unique la fonction s. Cependant, ce n'est pas toujours la meilleure caractérisation possible de s ; pour illustrer ce point, considérons les questions connexes suivantes :

Problème 1

- *Comment caractériser les coefficients de Taylor des fonctions de Schur ?*
- *En d'autres termes, comment voir que l'opérateur de Toeplitz*

$$\begin{pmatrix} \cdot \\ \cdot \\ y_1 \\ y_0 \end{pmatrix} = \begin{pmatrix} \cdot & \cdot & \cdot & \\ \cdot & a_0 & a_1 & a_2 \\ \cdot & 0 & a_0 & a_1 \\ \cdot & 0 & 0 & a_0 \end{pmatrix} \begin{pmatrix} \cdot \\ \cdot \\ x_1 \\ x_0 \end{pmatrix},$$

  est une contraction de ℓ_2 dans lui-même ?
- *Comment caractériser les fonctions de Schur s telles que $s(\omega_0) = a_0$, où ω_0 et a_0 sont donnés ?*

Répondre à ces questions ne semble pas aisé en termes des coefficients de Taylor. Une autre suite de coefficients, appelée suite des coefficients de Schur de s et introduite par I. Schur, caractérise de manière unique la fonction s et permet d'aborder ces questions. Cela nous ramène au début du $XX^{ème}$ siècle et en particulier à deux travaux [**324**], [**325**] de I. Schur de 1917 et 1918. Motivé par les travaux de Carathéodory, Fejér [**117**], Herglotz [**213**] et Toeplitz [**341**] (et en particulier par le problème des moments trigonométriques), Schur associe à une fonction s analytique et dont le module est borné par 1 dans le disque unité la suite $s_0, s_1, \ldots$ de fonctions définies par

$$\begin{aligned} s_0(z) &= s(z) \\ s_{n+1}(z) &= \frac{s_n(z) - s_n(0)}{z(1 - s_n(z)s_n(0)^*)}, \qquad n \geq 0 \end{aligned} \tag{1.1.1}$$

où * désigne la conjugaison dans $\mathbb{C}$. Le fait que les fonctions s_n sont encore analytiques et bornées par 1 en module dans $\mathbb{D}$ provient du principe du maximum. La suite s'arrête au rang n si $|s_n(0)| = 1$; sinon, elle est infinie. La suite des nombres $\rho_n = s_n(0)$ s'appelle suite des coefficients de Schur, ou coefficients de réflexion de s ; elle caractérise de manière unique la fonction s et joue un rôle important dans des domaines aussi divers que l'interpolation des fonctions analytiques, le filtrage des processus stochastiques stationnaires et la recherche sismique. Cette suite est aussi utile pour l'étude topologique des fonctions intérieures ; voir [**90**], [**118**]. La construction de la suite $s_0, s_1, \ldots$ s'appelle *algorithme de Schur*.

Exemple 1.1.1. — Calculer la suite des coefficients de Schur pour la fonction

$$b_\omega(z) = \frac{z - \omega}{1 - z\omega^*}, \tag{1.1.2}$$

et plus généralement pour un produit fini

$$c \prod b_{\omega_i}, \tag{1.1.3}$$

où $c \in \mathbb{C}$ est de module 1 et où les ω_i appartiennent à $\mathbb{D}$.

I. Schur a démontré que la suite des coefficients est finie si et seulement si s est égal à un tel produit (1.1.3) (appelé produit de Blaschke fini).

Notons que l'algorithme de Schur donne une représentation de la fonction s en fraction continue

$$s(z) = \rho_0 + \cfrac{z(1 - |\rho_0|^2)}{\rho_0{}^* z - \cfrac{1}{\rho_1 + \cfrac{z(1 - |\rho_1|)^2}{\rho_1^* z - \cdots}}}.$$

Voir [**354**, theorem 77.1, p. 285]. Nous renvoyons au livre de Bultheel [**113**] pour plus de détails sur les liens entre les fractions continues, l'approximation et la théorie des systèmes.

Plus généralement, on peut se fixer une suite (finie ou infinie) $\omega_0, \omega_1, \ldots$ de points de $\mathbb{D}$ et définir de manière récursive une suite s_n de fonctions de Schur par $s_0 = s$ et (pourvu que $|s_n(\omega_n)| < 1$)

$$s_{n+1}(z) = \frac{s_n(z) - s_n(\omega_n)}{b_{\omega_n}(z)(1 - s_n(z)s_n(\omega_n)^*)}, \qquad n \geq 0 \tag{1.1.4}$$

où b_ω est donnée par (1.1.2), et caractériser la fonction de Schur s par la suite (finie ou infinie) des valeurs $s_n(\omega_n)$.

La propriété des fonctions de Schur qui nous intéresse le plus est la positivité, au sens des noyaux reproduisants, de la fonction

$$K_s(z,\omega) = \frac{1 - s(z)s(\omega)^*}{1 - z\omega^*} \tag{1.1.5}$$

dans le disque unité. La démonstration de cette propriété est très simple et revient à vérifier que l'opérateur de multiplication par s est une contraction de l'espace de Hardy $\mathbf{H}_2$ dans lui–même ; voir le chapitre 2. (Nous rappelons la définition de l'espace de Hardy $\mathbf{H}_2$ dans le chapitre suivant ; voir l'exemple 2.1.5 et la section 2.6.) Soit $x(z) = \sum_0^\infty x_n z^n \in \mathbf{H}_2$ et soit $s(z) = \sum_0^\infty a_n z^n$ le développement de s en série de Taylor dans un voisinage de l'origine. L'opérateur de multiplication par s se traduit sur les coefficients des puissances de z par

$$y_n = a_n x_0 + a_{n-1}x_1 + \cdots + a_0 x_n, \qquad n \geq 0 \tag{1.1.6}$$

que l'on peut réécrire comme

$$\begin{pmatrix} y_0 \\ y_1 \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} a_0 & 0 & 0 & \cdot \\ a_1 & a_0 & 0 & \cdot \\ a_2 & a_1 & a_0 & \\ \cdot & \cdot & \cdot & \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \cdot \\ \cdot \end{pmatrix}$$

ou comme

$$\begin{pmatrix} \cdot \\ \cdot \\ y_1 \\ y_0 \end{pmatrix} = \begin{pmatrix} \cdot & \cdot & \cdot & \\ \cdot & a_0 & a_1 & a_2 \\ \cdot & 0 & a_0 & a_1 \\ \cdot & 0 & 0 & a_0 \end{pmatrix} \begin{pmatrix} \cdot \\ \cdot \\ x_1 \\ x_0 \end{pmatrix}, \tag{1.1.7}$$

c'est–à–dire en terme d'un opérateur de Toeplitz (il y a une très abondante bibliographie sur ces opérateurs ; nous renvoyons à [**147**, Chapter 7] et à [**271**, Appendix 4, pp. 299–398]).

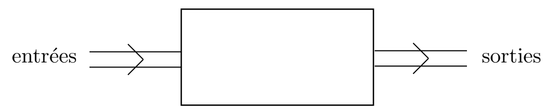
Ici intervient une des premières coïncidences heureuses qui lient la théorie des systèmes linéaires et la théorie des fonctions analytiques. Mais d'abord quelques mots sur ce terme. Dans ce travail, l'expression *système linéaire* désigne en fait une fonction analytique dans le disque unité (et souvent de plus rationnelle) à valeurs matricielles. Dans le cas rationnel, une telle fonction peut être abordée de plusieurs points de vue :

(1) en termes de polynômes matriciels premiers entre eux,
(2) en terme d'une matrice de Hankel,
(3) en termes d'une réalisation (*i.e.* d'un quadruplet de matrices ; voir la section 3.1).

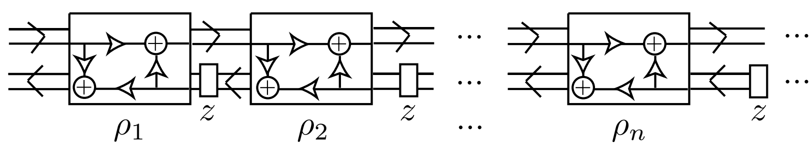
Voir [**173**, p. 1] pour le cas scalaire. Nous reviendrons par la suite sur certains de ces aspects.

Soit $(\mathbb{C}^p)^{\mathbb{N}}$ l'ensemble des suites finies indexées par les entiers naturels et à valeurs dans $\mathbb{C}^p$; pour l'instant, il suffit de voir un système causal à p entrées et q sorties comme un opérateur linéaire de $(\mathbb{C}^p)^{\mathbb{N}}$ dans $(\mathbb{C}^q)^{\mathbb{N}}$ de la forme (1.1.6). Interprétant les entiers naturels comme une échelle de temps discret, la causalité se traduit par le fait que la sortie au temps n ne dépend que des entrées jusqu'au temps n. Pour plus de détails sur la notion de causalité et les liens avec les fonctions analytiques, voir [**178**].

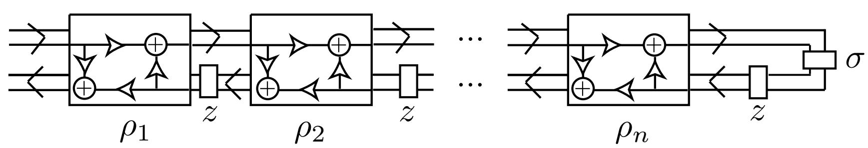
A ce stade, aucune notion de continuité n'intervient. On décrit schématiquement un système comme une boîte noire (*black box*)



Plutôt que de considérer la suite (x_n), il est souvent plus commode de travailler avec sa *transformée en* z, à savoir la série $\sum_0^\infty x_n z^n$. Cette série converge dans le disque unité, et en fait est dans l'espace de Hardy $\mathbf{H}_2$ lorsque le signal est d'énergie finie : on définit l'énergie du signal comme la norme ℓ_2 de la suite correspondante ; cette condition s'écrit donc $\sum_0^\infty |x_n|^2 < \infty$. Les systèmes dissipatifs sont particulièrement importants ; d'un point de vue mathématique le système est dissipatif si l'énergie du signal de sortie est toujours inférieure ou égale à l'énergie du signal d'entrée. La dissipativité du système se traduit par la positivité de la fonction (1.1.5). Dans cette optique, l'algorithme de Schur a une importance capitale ; il correspond à la décomposition du système en une cascade



ou, si l'algorithme se termine après un nombre fini d'étapes, en une cascade

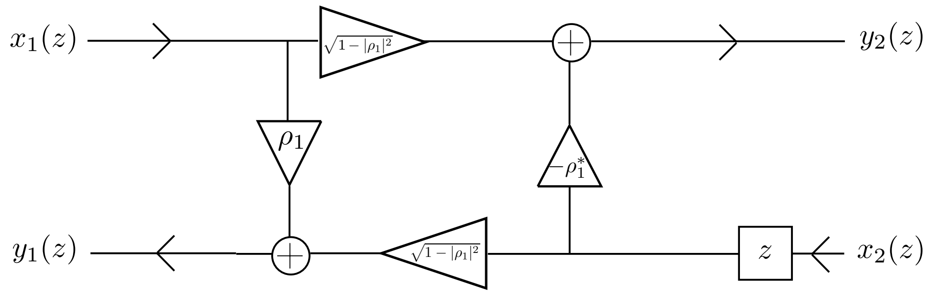


où σ est la fonction de transfert d'un filtre passif (la « charge ») et où chaque élément représente un filtre F_i sans pertes à deux entrées et deux sorties (caractérisé par le coefficient ρ_i) de fonction de transfert de chaîne (pour la définition, voir (3.5.2)),

$$\Theta_i(z) = \frac{1}{\sqrt{1-|\rho_i|^2}} \begin{pmatrix} 1 & \rho_i \\ \rho_i^* & 1 \end{pmatrix} \begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix}, \tag{1.1.8}$$

Les transformées en z des entrées $x_1(z)$ et $x_2(z)$ et des sorties $y_1(z)$ et $y_2(z)$ du filtre F_i sont reliées par les équations

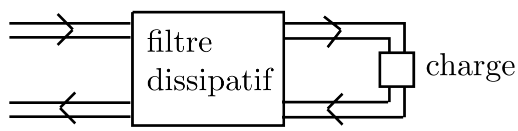
$$y_1(z) = \rho_i x_1(z) + z\sqrt{1-|\rho_i|^2}\, x_2(z) \tag{1.1.9}$$
$$y_2(z) = \sqrt{1-|\rho_i|^2}\, x_1(z) - z\rho_i^* x_2(z). \tag{1.1.10}$$



Les triangles représentent des opérateurs de multiplication, $\oplus$ représente l'addition.

Ces « filtres en échelle » (*ladder filters*) servent aussi à modéliser les milieux stratifiés (layered medium) et ont été considérés dans de nombreuses publications ; voir par exemple le livre [**191**] et les articles [**57**], [**179**], [**345**], [**346**], [**347**]. Nous reviendrons à ces filtres dans les chapitres 3 et 4.

D'une manière plus générale, le « *inverse scattering problem* » associé à s en théorie des systèmes (que nous traduisons par *problème de diffusion inverse*) consiste à trouver toutes les représentations de s de la forme



comme branchement d'un filtre dissipatif (ou mieux, sans pertes) à deux entrées et deux sorties de fonction de transfert de chaîne Θ, et d'une « charge » caractérisée par la fonction de transfert σ. La version projective de ce problème, appelée *problème de diffusion inverse partiel* joue aussi un rôle important ; voir la définition 4.3.2.

Le problème de diffusion inverse se traduit de la manière suivante en termes de fonctions analytiques : il s'agit d'exprimer une fonction de Schur s (la fonction de transfert du système dissipatif considéré) comme

$$s(z) = \frac{a(z)\sigma(z) + b(z)}{c(z)\sigma(z) + d(z)}. \tag{1.1.11}$$

où σ est encore une fonction de Schur et où la fonction matricielle $\Theta = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ est méromorphe dans $\mathbb{D}$ et telle que

$$\Theta(z)J_0\Theta(z)^* \leq J_0 \tag{1.1.12}$$

pour tout point d'analyticité de Θ. Dans cette expression,

$$J_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.1.13}$$

et l'inégalité signifie que la différence $J_0 - \Theta(z)J_0\Theta(z)^*$ est une matrice hermitienne positive. Cette condition peut aussi s'exprimer de la manière suivante : la matrice

$\Theta(z)$ est une contraction de $\mathbb{C}^2$ dans lui–même, lorsque cet espace est muni de la métrique indéfinie

$$\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right] = |x_1|^2 - |x_2|^2.$$

La dissipativité du filtre à deux entrées et deux sorties se traduit donc par la contractivité d'une fonction matricielle, mais par rapport à une métrique indéfinie. Les espaces munis de métrique indéfinie, et en particulier les espaces de Pontryagin, interviendront tout au long de cette étude.

La définition précédente du problème de diffusion inverse provient de la théorie des systèmes linéaires. Il n'en est que plus surprenant que ce problème corresponde exactement au problème de diffusion inverse associé aux paires d'opérateurs unitaires (voir [**109**], [**102**], [**107**]). Le problème de diffusion inverse peut sembler être un problème uniquement technique ; il intervient dans un nombre impressionnant de situations. De plus, lorsqu'on le réexprime dans le language des noyaux reproduisants (comme cela a été fait par L. de Branges et J. Rovnyak [**107**]), il suggère des généralisations à des situations nouvelles, traitées dans la troisième partie de ce travail.

La condition (1.1.12) implique que $d(z) \not\equiv 0$ (et en fait, d^{-1} est une fonction de Schur et est donc analytique dans $\mathbb{D}$; voir le chapitre 3 et la preuve de l'exemple 7.1.4). On peut donc réécrire

$$s(z) = \Sigma_{12}(z) + \Sigma_{11}(z)\sigma(z)(1 - \Sigma_{21}(z)\sigma(z))^{-1}\Sigma_{22}(z), \tag{1.1.14}$$

avec

$$\begin{aligned} \Sigma_{11}(z) &= a(z) - b(z)d(z)^{-1}c(z), \\ \Sigma_{12}(z) &= b(z)d(z)^{-1}, \\ \Sigma_{21}(z) &= -d^{-1}(z)c(z), \\ \Sigma_{22}(z) &= d^{-1}(z). \end{aligned}$$

La fonction $\Sigma = \left(\begin{smallmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{smallmatrix}\right)$ est contractive dans le disque unité. La transformation inverse $\Sigma \mapsto \Theta$ s'appelle la transformée de Potapov–Ginzburg . Les transformations homographiques du type (1.1.14) s'appellent aussi transformations de Redheffer.

Un problème un peu plus général consiste à considérer toutes les représentations (1.1.14) de s. Ce problème n'est pas équivalent au problème de diffusion inverse car une fonction contractive Σ peut avoir une entrée $\Sigma_{22} \equiv 0$ (ou, dans le cas matriciel, telle que $\det \Sigma_{22} \equiv 0$), auquel cas elle ne correspond pas à une matrice de transfert de chaîne Θ. Le cas $\sigma = 0$ revient à trouver toutes les matrices Σ dont le coin $1, 2$ est égal à S. Le cas où l'on impose à Σ d'être intérieure correspond au problème de la synthèse de Darlington.

L'algorithme de Schur donne une première famille de solutions du problème de diffusion inverse, et correspond à un problème d'approximation ou de synthèse du système ; du point de vue mathématique, il correspond à un problème d'interpolation.

Pour le voir, réécrivons la formule (1.1.1) pour $n = 0$ de la manière suivante :

$$s(z) = \frac{s_0 + z\sigma(z)}{1 + zs_0^*\sigma(z)} \tag{1.1.15}$$

(avec $\sigma(z) = s_1(z)$).

L'équation (1.1.15) est de la forme (1.1.11) avec Θ de la forme (1.1.8), et décrit l'ensemble de toutes les fonctions $s \in \mathcal{S}$ telles que $s(0) = s_0$. On a donc, pour $N = 0$, la solution du problème de Carathéodory–Fejér :

Etant donnés $N+1$ nombres complexes $s_0, s_1, \ldots, s_N$, trouver une condition nécessaire et suffisante pour qu'il existe une fonction de Schur s telle que $\dfrac{s^{(n)}(0)}{n!} = s_n$, $n = 0, \ldots, N$, et décrire l'ensemble de toutes les solutions.

De même, la variation (1.1.4) permet de résoudre pour $N = 0$ le problème de Nevanlinna–Pick :

Etant données $N+1$ paires de nombres complexes $(\omega_0, s_0), \ldots, (\omega_N, s_N)$, trouver une condition nécessaire et suffisante pour qu'il existe une fonction de Schur s telle que $s(\omega_n) = s_n$, $n = 0, \ldots, N$, et décrire l'ensemble de toutes les solutions.

Comme nous le verrons par la suite, l'algorithme de Schur permet en fait de résoudre pour tout N ces problèmes d'interpolation. Pour l'instant, notons seulement le fait suivant. De l'équation (1.1.15) on obtient :

$$s'(0) = (1 - |s(0)|^2)\sigma(0).$$

Donc, se fixer la dérivée de s au point 0 est équivalent à se fixer la valeur de σ au point 0, et donc à résoudre à nouveau le problème de Carathéodory–Fejér pour $N = 0$, mais cette fois pour σ. Ce phénomène de récursivité est la clé de l'étude du cas général.

Le cas où $N = \infty$ et d'autres cas (tels que les cas où des points ω_j sont de module égal à 1) sont bien sûr importants et ne doivent pas être ignorés.

Le problème de Carathéodory–Fejér est lié au problème des moments trigonométriques. En effet, soit s une fonction de Schur différente de la fonction $s(z) = -1$. La fonction $\phi(z) = \dfrac{1 - s(z)}{1 + s(z)}$ est analytique dans le disque unité et a une partie réelle positive. Par le théorème de Herglotz [**111**, Chapitre 1], une telle fonction ϕ s'écrit

$$\phi(z) = i\alpha + \int_0^{2\pi} \frac{e^{it} + z}{e^{it} - z} d\mu(t) \tag{1.1.16}$$

où α est un nombre réel et où μ est une mesure positive portée par $[0, 2\pi[$. Ces fonctions s'appellent fonctions de Carathéodory, et leur ensemble sera noté par $\mathcal{C}$. Écrivant (1.1.16) sous la forme

$$\phi(z) = i\alpha + \int_0^{2\pi} d\mu(t) + 2\sum_1^\infty z^n \int_0^{2\pi} e^{-int} d\mu(t), \tag{1.1.17}$$