Université de Lille

École doctorale **MADIS-631**

Unité de recherche **Laboratoire Paul Painlevé**

Thèse présentée par **Axel Renard**

Soutenue le **22 novembre 2024**

En vue de l'obtention du grade de docteur de l'Université de Lille

Discipline **Mathématiques**

Espaces modèles, ensembles spectraux et inégalités de type Schwarz-Pick

Thèse dirigée par Catalin Badea

Composition du jury

Rapporteurs	Stéphane Charpentier	MCF HDR à Aix-Marseille Université	
	Pamela Gorkin	professeure à Bucknell University	
Examinateurs	Emmanuel Fricain	professeur à l'Université de Lille	
	Sophie Grivaux	directrice de recherche au CNRS, Université de Lille	présidente du jury
	Jani Virtanen	professeur à University of Reading	
Directeur de thèse	Catalin Badea	professeur à l'Université de Lille	

UNIVERSITÉ DE LILLE

Doctoral School **MADIS-631**

University Department **Laboratoire Paul Painlevé**

Thesis defended by **Axel RENARD**

Defended on **November 22, 2024**

In order to become Doctor from Université de Lille

Academic Field **Mathematics**

Model spaces, spectral sets and Schwarz-Pick type inequalities

Thesis supervised by Catalin BADEA

Committee members

Referees	Stéphane CHARPENTIER	HDR Associate Professor at Aix-Marseille Université	
	Pamela GORKIN	Professor at Bucknell University	
Examiners	Emmanuel FRICAIN	Professor at Université de Lille	
	Sophie GRIVAUX	Senior Researcher at CNRS, Université de Lille	Committee President
	Jani VIRTANEN	Professor at University of Reading	
Supervisor	Catalin BADEA	Professor at Université de Lille	

This thesis has been prepared at

Laboratoire Paul Painlevé
CNRS UMR 8524
Cité Scientifique, Bâtiment M2
59655 Villeneuve d'Ascq Cedex

☎ +33 3 20 43 48 50
Web Site https://math.univ-lille.fr

Laboratoire
Paul Painlevé

Model spaces, spectral sets and Schwarz-Pick type inequalities

Abstract

The Schwarz-Pick inequality for holomorphic functions of one variable is a classical topic in complex analysis and, more specifically, in hyperbolic geometry. Schwarz-Pick type inequalities for three points and, subsequently, for n points in the open unit disc have been established by Beardon and Minda, and by Baribeau, Rivard, and Wegert, respectively. The aim of this PhD thesis is to study Schwarz-Pick type inequalities from an operator-theoretic perspective. By employing the von Neumann inequality for operators on Hilbert spaces and its generalizations, the Schwarz-Pick inequality and the Beardon-Minda inequality are derived in a unified manner. The connection with model operators is explored, and new Schwarz-Pick type inequalities for functions in one or several complex variables are established. The thesis also focuses on operator versions of Schwarz-Pick type inequalities, building on the work of Ky Fan and D. Joćić. To this end, we discuss related topics such as explicit holomorphic functional calculus, non-commutative divided differences, and contractivity criteria for matrices. The manuscript concludes with a brief exploration of some other spectral sets.

Keywords: functional analysis, operator theory, mathematical analysis, complex analysis, holomorphic functions, spectral theory, Hilbert spaces, hyperbolic geometry, Schwarz-Pick lemma

---

Espaces modèles, ensembles spectraux et inégalités de type Schwarz-Pick

Résumé

L'inégalité de Schwarz-Pick pour une fonction holomorphe d'une variable est un sujet classique en analyse complexe et, plus spécifiquement, en géométrie hyperbolique. Des inégalités de type Schwarz-Pick pour trois points et, ensuite, pour n points du disque unité ouvert ont été établies respectivement par Beardon et Minda, et par Baribeau, Rivard, et Wegert. L'objectif de cette thèse de doctorat est d'étudier les inégalités de type Schwarz-Pick dU point de vue de la théorie des opérateurs. En utilisant l'inégalité de von Neumann pour les opérateurs sur les espaces de Hilbert et ses généralisations, les inégalités de Schwarz-Pick et de Beardon-Minda sont obtenues de manière unifiée. Le lien avec les opérateurs modèles est mis en évidence, et de nouvelles inégalités de type Schwarz-Pick en une ou plusieurs variables complexes sont établies. Nous nous concentrons également sur les versions opérateurs des inégalités de type Schwarz-Pick, en suivant les travaux de Ky Fan et D. Joćić. Pour ce faire, nous discutons d'autres sujets d'intérêt, tels que le calcul fonctionnel holomorphe explicite, les différences divisées non commutatives et les critères de contractivité pour les matrices. Le manuscrit se termine par un aperçu rapide d'autres ensembles spectraux.

Mots clés : analyse fonctionnelle, théorie des opérateurs, analyse mathématique, analyse complexe, fonctions holomorphes, théorie spectrale, espaces de Hilbert, géométrie hyperbolique, lemme de Schwarz-Pick

---

Laboratoire Paul Painlevé
CNRS UMR 8524 – Cité Scientifique, Bâtiment M2 – 59655 Villeneuve d'Ascq Cedex

Remerciements

J'aimerais tout d'abord remercier Catalin Badea, sans qui rien de tout cela n'aurait été possible : merci d'avoir accepté de m'encadrer pour cette thèse et de m'avoir guidé tout au long de ces trois années.

Je remercie également Pamela Gorkin et Stéphane Charpentier d'avoir accepté le rôle de rapporteurs, ainsi que Sophie Grivaux, Emmanuel Fricain et Jani Virtanen d'avoir accepté de faire partie du jury de soutenance.

Je tiens ensuite à remercier les membres des équipes d'analyse fonctionnelle lilloise, lensoise, calaisienne et montoise pour ces moments partagés lors des groupes de travail et des séminaires, qui m'ont beaucoup apporté. Merci également aux personnels administratifs du laboratoire, aux bibliothécaires — Catherine, Hélène, Omar et Pascal —, ainsi qu'aux autres doctorants avec qui j'ai pu échanger durant ces trois années. Je pense en particulier à Vincent, Marvin, Valentin, Romain, Micheline et Pierre.

Je remercie aussi Maeva pour le temps qu'elle m'a consacré en amont de ma soutenance, et pour ses précieux conseils visant à améliorer ma présentation.

Merci également à Denis Bitouzé pour sa formation de perfectionnement à LaTeXet pour toutes les astuces qui m'auront permis de gagner un temps précieux dans la rédaction des articles et de ce manuscrit.

Je tiens en outre à exprimer ma reconnaissance envers tous les enseignants qui m'ont guidé pendant ma scolarité. Je pense en particulier à Léa Blanc-Centi, Emmanuel Fricain, Vincent Thilliez et Etienne Matheron, dont les cours de grande qualité m'ont fait aimer l'analyse fonctionnelle, à Julien Hauseux, qui m'a accompagné dans la rédaction de mon tout premier mémoire de recherche, à Carlos Ribeiro, rencontré en première année de licence et avec qui j'ai eu plaisir à échanger pendant le reste de mon parcours universitaire, ainsi qu'à Benoît Dussart, qui m'a transmis la passion des mathématiques et encouragé à poursuivre mes études dans ce domaine.

J'ai également une pensée pour mes grands-parents paternels, qui m'ont transmis le goût d'apprendre et m'ont toujours encouragé dans mes études : bien que vous ne soyez plus là pour partager ce moment, je sais que vous seriez fiers de me voir accomplir ce chemin.

Enfin, bien qu'elle ne puisse guère lire ces quelques lignes, je ne pourrais conclure ce paragraphe sans remercier ma chère Fidji – fidèle compagne de route – pour sa présence calme et réconfortante tout au long de ces trois années.

Table des matières

Notation

$\mathbb{N}$	denotes the set of non-negative integers, *i.e.* $\mathbb{N} = \{0, 1, 2, 3, \dots\}$.
$\mathbb{D}$	denotes the open unit disk
$\mathbb{T}$	denotes the unit circle, $\mathbb{T} = \overline{\mathbb{D}} \backslash \mathbb{D}$
$\lVert\cdot\rVert_\infty$	for $\omega = (\omega_1, \dots, \omega_n) \in \mathbb{C}^n$, we denote $\lVert\omega\rVert_\infty = \sup\{\lvert\omega_i\rvert : 1 \leq i \leq n\}$
$\mathbb{B}_n$	denotes the open unit ball of $\mathbb{C}^n$, *i.e.* $\mathbb{B}_n = \left\{(z_1, \dots, z_n) \in \mathbb{C}^n : \sum_{i=1}^n \lvert z_i\rvert^2 < 1\right\}$
$\mathcal{C}(X, Y)$	denotes the set of continuous functions mapping X to Y, where X and Y are metric spaces
$\lVert \cdot \rVert_\infty$	for $f \in \mathcal{C}\left(\overline{\mathbb{D}}\right)$, we denote $\lVert f\rVert_\infty = \sup\{\lvert f(z)\rvert : \lvert z\rvert \leq 1\}$
$\mathcal{H}(\mathcal{U}, \mathcal{V})$	is the set of holomorphic functions mapping $\mathcal{U}$ to $\mathcal{V}$, where $\mathcal{U}$ and $\mathcal{V}$ are subsets of $\mathbb{C}$. When $\mathcal{V} = \mathbb{C}$, we sometimes write $\mathcal{H}(\mathcal{U})$ instead of $\mathcal{H}(\mathcal{U}, \mathbb{C})$.
$H^2(\mathcal{U})$	is the Hilbert-Hardy space of $\mathcal{U}$, where $\mathcal{U}$ is an open subset of $\mathbb{C}$. When $\mathcal{U} = \mathbb{D}$, we sometimes write H^2 instead of $H^2(\mathbb{D})$
$H^\infty(\mathcal{U})$	is the set of bounded holomorphic functions on $\mathcal{U}$, where $\mathcal{U}$ is an open subset of $\mathbb{C}$. When $\mathcal{U} = \mathbb{D}$, we sometimes write H^∞ instead of $H^\infty(\mathbb{D})$
$\mathcal{A}\left(\overline{\mathbb{D}}\right)$	is the disc algebra, *i.e.* the set of functions that are holomorphic on $\mathbb{D}$ and continuous on $\overline{\mathbb{D}}$
$[f(z_0), \dots, f(z_n)]$	denotes the $n+1$ order divided difference of f at the points $z_0, \dots, z_n$
(z, w)	denotes the complex pseudo-hyperbolic distance $(z, w) := \frac{z-w}{1-\overline{w}z}$
$\rho(z, w)$	denotes the pseudo-hyperbolic distance $\rho(z, w) := \lvert(z, w)\rvert$
$\mathrm{d}(z, w)$	denotes the hyperbolic distance $\mathrm{d}(z, w) = \tanh^{-1} \frac{1+\rho(z,w)}{1-\rho(z,w)}$
$f^*(z, w)$	denotes the hyperbolic divided difference $f^*(z, w) := \frac{(f(z), f(w))}{(z,w)}$
$\mathcal{B}(H, K)$	is the set of bounded linear operators from H to K, where H and K are two complex Hilbert spaces. $\mathcal{B}(H)$ is a short for $\mathcal{B}(H, H)$. In this manuscript, all the operators are supposed to be linear and bounded
$\lVert \cdot \rVert$	denotes the norm of an element of the Banach space under consideration. When $T \in \mathcal{B}(H, K)$, $\lVert T\rVert$ denotes the operator norm

T^*	denotes the adjoint of T, where T is a Hilbert space operator
A^{T}	denotes the transpose of the matrix A
$\langle \cdot, \cdot \rangle$	denotes the dot product, with the convention that $\langle \cdot, \cdot \rangle$ is linear with respect to the first argument, and anti-linear with respect to the second one
D_T	denotes the defect operator of a contraction $T \in \mathcal{B}(H)$, i.e. $\|T\| \leq 1$. Thus $D_T = (\mathrm{Id} - T^*T)^{1/2}$, where Id is the identity operator
$A \leq B$	means that the bounded linear operator $B - A$, acting on a Hilbert space H, is positive. If $B = \lambda \mathrm{Id}$, for some $\lambda \in \mathbb{C}$, we may simply write $A \leq \lambda$.
$\sigma(T)$	denotes the spectrum of $T \in \mathcal{B}(H)$
$r(T)$	denotes the spectral radius $r(T) = \sup\{\lvert\lambda\rvert : \lambda \in \sigma(T)\}$ of T
$\mathrm{Ker}(T)$	denotes the kernel of T
$\mathrm{Im}(T)$	denotes the range (image) of T
$[\![1, n]\!]$	denotes the set of all integers j with $1 \leq j \leq n$
$\mathcal{S}_n$	denotes the set of permutations of $[\![0, n-1]\!]$.

In the entire manuscript, we use the convention that an empty sum is equal to 0 and that an empty product is equal to 1.

Introduction (en français)

Motivation

Le point de départ de ce travail est l'observation quelque peu surprenante que le lemme de Schwarz-Pick, un résultat classique en analyse complexe, peut être obtenu comme un cas particulier de l'inégalité de von Neumann, qui est un résultat classique en théorie spectrale des opérateurs. Commençons par rappeler les énoncés de ces deux résultats :

Théorème 0.1 (Inégalité de von Neumann)**.** *Soit T une contraction opérant sur un espace de Hilbert H, c'est-à-dire un opérateur linéaire borné sur H tel que $\|T\| \leq 1$, et soit $p \in \mathbb{C}[X]$ un polynôme.*

On a :

$$\|p(T)\| \leq \|p\|_\infty := \sup\{|p(z)| \,:\, |z| \leq 1\}.$$

Théorème 0.2 (Lemme de Schwarz-Pick)**.**

(i) Soit $f : \mathbb{D} \to \mathbb{D}$ une fonction holomorphe du disque unité dans lui-même, et soient $\omega_1, \omega_2 \in \mathbb{D}$ deux points distincts du disque unité. Alors, on a :

$$\left|\frac{f(\omega_1) - f(\omega_2)}{1 - \overline{f(\omega_1)}f(\omega_2)}\right| \leq \left|\frac{\omega_1 - \omega_2}{1 - \overline{\omega_1}\omega_2}\right| \tag{0.1}$$

Si l'on note $\rho(\omega_1, \omega_2) := \left|\frac{\omega_1 - \omega_2}{1 - \overline{\omega_1}\omega_2}\right|$ la distance pseudo-hyperbolique sur le disque unité, cette inégalité devient :

$$\rho(f(\omega_1), f(\omega_2)) \leq \rho(\omega_1, \omega_2).$$

En d'autres termes, toute fonction holomorphe du disque unité est contractante par rapport à la métrique pseudo-hyperbolique.

(ii) Si $\omega_1 = \omega_2 =: \omega \in \mathbb{D}$, cette inégalité devient :

$$|f'(\omega)| \leq \frac{1 - |f(\omega)|^2}{1 - |\omega|^2} \tag{0.2}$$

Il est connu (voir, par exemple, [46, Ex. 2.17–2.18] ou [52, p. 17]) que l'inégalité de Schwarz-Pick peut être obtenue comme un cas particulier de l'inégalité de von Neumann, appliquée à une matrice 2×2 bien choisie. L'idée est la suivante : soient $\omega_1, \omega_2 \in \mathbb{D}$ deux points distincts du disque unité ouvert, et soit

$$T_2 = \begin{pmatrix} \omega_1 & \alpha \\ 0 & \omega_2 \end{pmatrix} \in \mathcal{M}_2(\mathbb{C}).$$

On commence, dans un premier temps, par prouver les deux lemmes suivants :

Lemme 0.3. *T_2 est une contraction si et seulement si*

$$|\alpha|^2 \leq (1-|\omega_1|^2)(1-|\omega_2|^2).$$

Lemme 0.4. *Pour tout polynôme $p \in \mathbb{C}[X]$, on a :*

$$p(T_2) = \begin{pmatrix} p(\omega_1) & \alpha \cdot \left(\frac{p(\omega_1)-p(\omega_2)}{\omega_1-\omega_2}\right) \\ 0 & p(\omega_2) \end{pmatrix}.$$

On pose ensuite $\alpha = \sqrt{1-|\omega_1|^2}\sqrt{1-|\omega_2|^2}$, et on applique l'inégalité de von Neumann à T_2. On obtient alors une inégalité qui, après quelques transformations, s'avère être équivalente à l'inégalité de Schwarz-Pick pour les polynômes. On peut alors conclure la démonstration par un argument d'approximation. Davantage de détails seront donnés dans le Chapitre 1, après avoir passé en revue quelques concepts fondamentaux de théorie spectrale.

Dans cette thèse, nous étendons cette observation à des matrices de plus grande taille et à des matrices à coefficients opérateurs. Ce faisant, nous sommes confrontés aux deux problèmes suivants :

Problème 0.5. *Étant donnée une matrice de taille $n \times n$, notée T_n, et un polynôme – ou une fonction notée f – comment exprimer explicitement $f(T_n)$?*

Problème 0.6. *Étant donnée une matrice T_n de taille $n \times n$, comment estimer sa norme spectrale ?*

Au vu du théorème de décomposition de Schur – qui stipule que toute matrice est unitairement équivalente à une matrice triangulaire supérieure – l'étude peut être limitée aux matrices triangulaires supérieures.

Fonction d'une matrice : un calcul fonctionnel explicite

Le premier problème a déjà été étudié sous différents aspects dans plusieurs références (voir par exemple [33, 36, 16, 31, 40]). Dans ce manuscrit, nous allons développer une approche légèrement différente. Dans le Chapitre 3, nous introduisons la notion de *différences divisées non commutatives* pour les opérateurs – une généralisation des différences divisées *usuelles* bien connues pour les scalaires. Nous commençons par les polynômes, puis nous étendons ce concept aux fonctions rationnelles, et enfin aux applications holomorphes.

Rappelons que, dans le cas scalaire, si $x_0 \ldots, x_n$ sont $n+1$ points deux à deux distincts de $\mathbb{C}$ et si f est une application définie sur un sous-ensemble de $\mathbb{C}$ contenant $x_0, \ldots, x_n$, les *différences divisées* sont définies par récurrence comme suit :

$$[f(x_k)] := f(x_k), \ \forall\, k \in [\![0, n]\!]$$
$$[f(x_k), \ldots, f(x_{k+j})] := \frac{[f(x_{k+1}), \ldots, f(x_{k+j})] - [f(x_k), \ldots, f(x_{k+j-1})]}{x_{k+j} - x_k}, \ \forall k \in [\![0, n-j]\!], \forall\, j \in [\![1, n]\!]$$

Selon la régularité de f, on peut alors obtenir plusieurs formules explicites pour les différences divisées. Nous discuterons ce point plus en détails dans la Section 3.1.

Dans le chapitre 3, on étend ensuite cette notion aux opérateurs, et on définit les *différences divisées non commutatives* comme suit (voir Définitions 3.2.1, 3.2.11, 3.2.13 et 3.2.20) :

Définition 0.7. Soient $H_0, \dots, H_n$ des espaces de Hilbert. Étant donnés des opérateurs linéaires bornés $W_i \in \mathcal{B}(H_i)$, pour $i \in [\![0,n]\!]$, et $C_j \in \mathcal{B}\left(H_{j+1}, H_j\right)$, pour $j \in [\![0,n-1]\!]$, on définit :

(i) $[W_0^k, \dots, W_n^k]_{(C_0,\dots,C_{n-1})} := \sum_{\substack{i_0+\dots+i_n=k-n\\ i_0,\dots,_n \geq 0}} W_0^{i_0} C_0 W_1^{i_1} \dots C_{n-1} W_n^{i_n}$, $\forall k \geq 0$.

Cette définition peut être étendue par linéarité aux polynômes : pour un polynôme $p(X) = \sum_{k=0}^n a_k X^k \in \mathbb{C}[X]$, on pose :

$$[p(W_0), \dots, p(W_n)]_{(C_0,\dots,C_{n-1})} = \sum_{k=0}^n a_k [W_0^k, \dots, W_n^k]_{(C_0,\dots,C_{n-1})}.$$

Ensuite, pour les fractions rationnelles, on définit :

(ii) $\left[W_0^{-k}, \dots, W_n^{-k}\right]_{(C_0,\dots,C_{n-1})} := (-1)^n \sum_{\substack{i_0+\dots+i_n=k-n\\ i_0,\dots,i_n \geq 1}} W_0^{-i_0} C_0 W_1^{-i_1} \dots C_{n-1} W_n^{-i_n}$, $\forall k \geq 1$;

En utilisant la décomposition en éléments simples, cette définition peut également être étendue par linéarité à toutes les fractions rationnelles.

Enfin, pour toute fonction f holomorphe sur un voisinage de $\bigcup_{i=0}^n \sigma(W_i)$, on définit :

(iii) $[f(W_0), \dots, f(W_n)]_{(C_0,\dots,C_{n-1})} := \int_\Gamma f(\xi)(\xi \mathrm{Id} - W_0)^{-1} C_0 (\xi \mathrm{Id} - W_1)^{-1} \dots C_{n-1} (\xi \mathrm{Id} - W_n)^{-1} \mathrm{d}\xi$, où Γ est un système fini de courbes de Jordan rectifiables, orientées positivement, entourant $\bigcup_{i=0}^n \sigma(W_i)$.

Nous verrons que – par analogie avec les différences divisées classiques – les *différences divisées non commutatives* satisfont la relation de récurrence suivante :

Proposition 0.8. *Soit $n \geq 1$, et soient $H_0, \dots, H_n$ des espaces de Hilbert. Étant donnés des opérateurs linéaires bornés $W_i \in \mathcal{B}(H_i)$, pour $i \in [\![0,n]\!]$, et $C_j \in \mathcal{B}(H_j, H_{j-1})$, pour $j \in [\![1,n]\!]$, et étant donnée une fonction f analytique au voisinage de $\bigcup_{i=0}^n \sigma(W_i)$, on a :*

$$\begin{aligned}
&[f(W_0)] = f(W_0)\,;\\
&W_0[f(W_0), \dots, f(W_n)]_{(C_0,\dots,C_{n-1})} - [f(W_0), \dots, f(W_n)]_{(C_0,\dots,C_{n-1})} W_n\\
&\quad = [f(W_0), \dots, f(W_{n-1})]_{(C_1,\dots,C_{n-2})} C_{n-1} - C_0 [f(W_1), \dots, f(W_n)]_{(C_1,\dots,C_{n-1})}.
\end{aligned}$$

Cette notion de *différences divisées non commutatives* nous permettra d'établir un calcul fonctionnel explicite pour les matrices triangulaires supérieures à coefficients opérateurs (voir Théorèmes 3.2.6, 3.2.16 et 3.2.22) :

Théorème 0.9. *Soit $n \geq 2$, soient $H_1, \ldots, H_n$ des espaces de Hilbert, et soit*

$$\widetilde{T_n} = \begin{pmatrix} W_1 & C_1^{(1)} & \cdots & \cdots & C_1^{(n-1)} \\ 0 & W_2 & C_2^{(1)} & \cdots & C_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & C_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & W_n \end{pmatrix} : \bigoplus_{i=1}^{n} H_i \to \bigoplus_{i=1}^{n} H_i$$

un opérateur linéaire borné. Soit f une fonction analytique dans un voisinage de $\bigcup_{i=0}^{n} \sigma(W_i)$.

Alors, on a :

$$f\left(\widetilde{T_n}\right) = \begin{pmatrix} f(W_1) & D_1^{(1)} & \cdots & \cdots & D_1^{(n-1)} \\ 0 & f(W_2) & D_2^{(1)} & \cdots & D_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & D_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & f(W_n) \end{pmatrix},$$

$$\text{où} \quad D_i^{(j)} = \sum_{\substack{1 \leq l \leq j \\ i=k_0<\cdots<k_l=j+i}} \left[f\left(W_{k_0}\right), \ldots, f\left(W_{k_l}\right)\right]_{\left(C_{k_0}^{(k_1-k_0)}, \ldots, C_{k_{l-1}}^{(k_l-k_{l-1})}\right)}.$$

L'analogie entre les différences divisées non commutatives pour les opérateurs et les différences divisées *usuelles* pour les scalaires nous permettra alors d'obtenir le corollaire suivant :

Corollaire 0.10. *Soit*

$$T_n = \begin{pmatrix} \omega_1 & \alpha_1^{(1)} & \cdots & \cdots & \alpha_1^{(n-1)} \\ 0 & \omega_2 & \alpha_2^{(1)} & \cdots & \alpha_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \alpha_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & \omega_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$$

une matrice telle que $\omega_i \neq \omega_j$, pour tout $i \neq j$, et soit f une fonction holomorphe au voisinage de $\omega_1, \ldots, \omega_n$.

On a alors :

$$f(T_n) = \begin{pmatrix} f(\omega_1) & \beta_1^{(1)} & \cdots & \cdots & \beta_1^{(n-1)} \\ 0 & f(\omega_2) & \beta_2^{(1)} & \cdots & \beta_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \beta_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & f(\omega_n) \end{pmatrix},$$

$$\text{où} \quad \beta_i^{(j)} = \sum_{\substack{1 \leq l \leq j \\ i=k_0<\cdots<k_l=j+i}} \left(\prod_{s=0}^{l-1} \alpha_{k_s}^{(k_{s+1}-k_s)}\right) \left[f\left(\omega_{k_0}\right), \ldots, f\left(\omega_{k_l}\right)\right].$$

La relation de récurrence de la Proposition 0.8 nous amènera ensuite à étudier l'équation de Sylvester :

$$AX - XB = Y, \tag{0.3}$$

où A, X, B, Y sont des opérateurs linéaires bornés agissant sur un espace de Banach.

Dans la Section 3.3, nous verrons que, dans le cas où les spectres respectifs des W_i sont deux à deux disjoints, la relation de récurrence de la Proposition 0.8 définit complètement les *différences divisées non commutatives*. Cela nous mènera au résultat suivant :

Théorème 0.11. *Soit $n \geq 2$, soient $H_1, \dots, H_n$ des espaces de Hilbert, et soit*

$$\widetilde{T_n} = \begin{pmatrix} W_1 & C_1^{(1)} & \cdots & C_1^{(n-1)} \\ 0 & W_2 & C_2^{(1)} \cdots & C_2^{(n-2)} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & C_{n-1}^{(1)} \\ 0 & \cdots & 0 & W_n \end{pmatrix} : \bigoplus_{i=1}^{n} H_i \to \bigoplus_{i=1}^{n} H_i$$

un opérateur linéaire borné. Soit f une fonction analytique au voisinage de $\bigcup_{i=0}^{n} \sigma(W_i)$. Supposons que $\sigma(W_i) \cap \sigma(W_j) = \emptyset$, pour tout $i \neq j$.

Alors, on a :

$$f\left(\widetilde{T_n}\right) = \begin{pmatrix} \mathrm{Id} & -X_1^{(1)} & \cdots & -X_1^{(n-1)} \\ 0 & \mathrm{Id} & -X_2^{(1)} \cdots & -X_2^{(n-2)} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -X_{n-1}^{(1)} \\ 0 & \cdots & 0 & \mathrm{Id} \end{pmatrix} \begin{pmatrix} f(W_1) & 0 & \cdots & 0 \\ 0 & f(W_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & f(W_n) \end{pmatrix} \begin{pmatrix} \mathrm{Id} & Y_1^{(1)} & \cdots & Y_1^{(n-1)} \\ 0 & \mathrm{Id} & Y_2^{(1)} \cdots & Y_2^{(n-2)} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & Y_{n-1}^{(1)} \\ 0 & \cdots & 0 & \mathrm{Id} \end{pmatrix}$$

où, pour tout $j \in [\![1, n-1]\!]$, pour tout $i \in [\![1, n+1-j]\!]$, $X_i^{(j)}$ est l'unique solution de l'équation

$$W_i X_i^{(j)} - X_i^{(j)} W_{i+j} = C_i^{(j)} - \sum_{k=1}^{j-1} C_i^{(k)} X_{k+i}^{(j-k)} \tag{0.4}$$

et

$$Y_i^{(j)} = \sum_{\substack{1 \leq l \leq j \\ i=k_0 < \cdots < k_l = j+i}} X_{k_0}^{(k_1-k_0)} \dots X_{k_{l-1}}^{(k_l-k_{l-1})} \tag{0.5}$$

Matrices contractantes

Le second problème (Problème 0.6), consistant à estimer la norme spectrale d'une matrice, sera abordé dans le Chapitre 4. Une approche directe pour estimer la norme d'une matrice consisterait – au moins pour les matrices à coefficients scalaires – à tenter d'exploiter la formule $\|T_n\|^2 = \|T_n^* T_n\| = r(T_n^* T_n)$, où $r(\cdot)$ désigne le rayon spectral. Cette méthode fonctionne très bien pour les matrices 2×2 (voir Lemme 1.2.14). Cependant, pour $n \geq 3$, les calculs deviennent trop complexes pour fournir un critère utilisable en pratique. Dans ce manuscrit, nous utiliserons le raffinement suivant du théorème de Parrott sur la complétion de matrices à coefficients opérateurs :

Théorème 0.12 (Parrott)**.** *Soient H_1, H_2, K_1, K_2 des espaces de Hilbert. On suppose que les opérateurs $\begin{bmatrix} A \\ C \end{bmatrix} \in \mathcal{B}(H_1, K_1 \oplus K_2)$ et $\begin{bmatrix} C & D \end{bmatrix} \in \mathcal{B}(H_1 \oplus H_2, K_2)$ sont des contractions.*

Alors,

$$T = \begin{bmatrix} A & B \\ C & D \end{bmatrix} : H_1 \oplus H_2 \to K_1 \oplus K_2$$

est une contraction si et seulement s'il existe une contraction $W \in \mathcal{B}(H_2, K_1)$ telle que :

$$B = D_{Z^*} W D_Y - Z C^* Y,$$

où $Z \in \mathcal{B}(H_1, K_1)$ et $Y \in \mathcal{B}(H_2, K_2)$ sont des contractions telles que $D = D_{C^*} Y$ et $A = Z D_C$, et où D_S désigne l'opérateur de défaut associé à la contraction S.

De plus,

(i) Y et Z peuvent être choisis comme étant (respectivement) les uniques solutions Y_0 et Z_0 de norme minimale parmi toutes les solutions des équations $D = D_{C^*} Y$ et $A = Z D_C$;

(ii) Si T est une contraction, il existe une unique contraction W_0 telle que :

$$B = D_{Z_0^*} W_0 D_{Y_0} - Z_0 C^* Y_0 \text{ et } Im\left(D_{Z_0^*}\right)^{\perp} \subset Ker(W_0^*).$$

Cet opérateur satisfait :

$$\|W_0\| = \inf\{\|W\| : B = D_{Z_0^*} W D_{Y_0} - Z_0 C^* Y_0\}.$$

Nous déduirons de ce théorème les résultats suivants pour les matrices de taille 3×3 et 4×4 (voir les Théorèmes 4.2.2 et 4.3.1 pour une version plus détaillée de ces résultats) :

Théorème 0.13. *Soient $\omega_1, \omega_2, \omega_3 \in \mathbb{D}$. Alors, $T = \begin{pmatrix} \omega_1 & \alpha_1 & \beta \\ 0 & \omega_2 & \alpha_2 \\ 0 & 0 & \omega_3 \end{pmatrix} \in \mathcal{M}_3(\mathbb{C})$ est une contraction (vue comme un opérateur agissant sur l'espace de Hilbert $\mathbb{C}^3$) si et seulement si :*

$$\begin{cases} |\alpha_i|^2 \leq (1 - |\omega_i|^2)(1 - |\omega_{i+1}|^2), \quad i = 1, 2, \\ \left|\beta(1 - |\omega_2|^2) + \alpha_1 \alpha_2 \overline{\omega_2}\right|^2 \leq \left[(1 - |\omega_1|^2)(1 - |\omega_2|^2) - |\alpha_1|^2\right] \cdot \left[(1 - |\omega_2|^2)(1 - |\omega_3|^2) - |\alpha_2|^2\right] \end{cases}$$

Théorème 0.14. *Soient $\omega_1, \omega_2, \omega_3, \omega_4 \in \mathbb{D}$. Alors, $T = \begin{pmatrix} \omega_1 & \alpha_1 & \beta_1 & \gamma \\ 0 & \omega_2 & \alpha_2 & \beta_2 \\ 0 & 0 & \omega_3 & \alpha_3 \\ 0 & 0 & 0 & \omega_4 \end{pmatrix} \in \mathcal{M}_4(\mathbb{C})$ est une con-*

traction (vue comme un opérateur agissant sur l'espace de Hilbert $\mathbb{C}^4$) si et seulement si :

$$\begin{cases} |\alpha_i|^2 \leq (1-|\omega_i|^2)(1-|\omega_{i+1}|^2),\ i=1,3 \quad \& \quad |\alpha_2|^2 < (1-|\omega_2|^2)(1-|\omega_3|^2) \\ |\beta_i(1-|\omega_{i+1}|^2)+\alpha_i\alpha_{i+1}\overline{\omega_{i+1}}|^2 \leq \Big[(1-|\omega_i|^2)(1-|\omega_{i+1}|^2)-|\alpha_i|^2\Big]\times \\ \qquad\qquad\qquad \Big[(1-|\omega_{i+1}|^2)(1-|\omega_{i+2}|^2)-|\alpha_{i+1}|^2\Big], \quad i=1,2 \\ \Big|\gamma[(1-|\omega_2|^2)(1-|\omega_3|^2)-|\alpha_2|^2]+\alpha_1\beta_2\overline{\omega_2}(1-|\omega_3|^2)+\alpha_3\beta_1\overline{\omega_3}(1-|\omega_2|^2) \\ \qquad\qquad\qquad +\beta_1\beta_2\overline{\alpha_2}+\alpha_1\alpha_2\alpha_3\overline{\omega_2\omega_3}\Big|^2(1-|\omega_2|^2)(1-|\omega_3|^2) \\ \leq \Big[\Big((1-|\omega_1|^2)(1-|\omega_2|^2)-|\alpha_1|^2\Big)\Big((1-|\omega_2|^2)(1-|\omega_3|^2)-|\alpha_2|^2\Big)-\Big|\alpha_1\alpha_2\overline{\omega_2}+\beta_1(1-|\omega_2|^2)\Big|^2\Big] \\ \times\Big[\Big((1-|\omega_2|^2)(1-|\omega_3|^2)-|\alpha_2|^2\Big)\Big((1-|\omega_3|^2)(1-|\omega_4|^2)-|\alpha_3|^2\Big)-\Big|\alpha_2\alpha_3\overline{\omega_3}+\beta_2(1-|\omega_3|^2)\Big|^2\Big] \end{cases}$$

$$ou \begin{cases} |\alpha_i|^2 \leq (1-|\omega_i|^2)(1-|\omega_{i+1}|^2),\ i=1,3 \quad \& \quad |\alpha_2|^2 = (1-|\omega_2|^2)(1-|\omega_3|^2) \\ \beta_i = \dfrac{-\alpha_i\alpha_{i+1}\omega_{i+1}}{1-|\omega_{i+1}|^2},\ i=1,2 \\ \left|\gamma - \dfrac{\overline{\omega_2\omega_3}\alpha_1\alpha_2\alpha_3}{(1-|\omega_2|^2)(1-|\omega_3|^2)}\right|^2 (1-|\omega_2|^2)(1-|\omega_3|_2) \\ \quad \leq \Big[(1-|\omega_1|^2)(1-|\omega_2|^2)-|\alpha_1|^2\Big]\times\Big[(1-|\omega_3|^2)(1-|\omega_4|^2)-|\alpha_3|^2\Big] \end{cases}$$

Nous verrons que dans le cas des matrices de taille 4×4, les calculs deviennent plus laborieux. Le calcul fonctionnel explicite du Théorème 0.9 nous permettra de simplifier ces derniers en supposant que l'un des coefficients diagonaux est égal à 0.

Nous obtiendrons également la généralisation suivante du Théorème 0.13 pour les matrices à coefficients opérateurs :

Théorème 0.15. *Soient H_1, H_2, H_3 trois espaces de Hilbert. Soient $W_i \in \mathcal{B}(H_i)$, $1 \leq i \leq 3$, trois contractions, et soit*

$$T = \begin{bmatrix} W_1 & A_1 & B \\ 0 & W_2 & A_2 \\ 0 & 0 & W_3 \end{bmatrix} \in \mathcal{B}(H_1 \oplus H_2 \oplus H_3).$$

Alors, T est une contraction si et seulement s'il existe trois contractions $V_1 \in \mathcal{B}(H_2, H_1), V_2 \in \mathcal{B}(H_3, H_2), V_3 \in \mathcal{B}(H_3, H_1)$ telles que :

$$\begin{cases} A_1 = D_{W_1^*} V_1 D_{W_2}, \\ A_2 = D_{W_2^*} V_2 D_{W_3}, \\ B = \Big[D_{W_1^*}(\mathrm{Id} - V_1V_1^*)D_{W_1^*}\Big]^{1/2} V_3 \Big[D_{W_3}(\mathrm{Id} - V_2^*V_2)D_{W_3}\Big]^{1/2} - D_{W_1^*}V_1W_2^*V_2D_{W_3}. \end{cases}$$

De plus, en tentant de généraliser les Théorèmes 0.13 et 0.14 aux matrices de taille $n\times n$, nous établirons le théorème suivant, qui caractérise les contractions dont la diagonale et la sur-diagonale sont prescrites :

Théorème 0.16. *Soit $n \in \mathbb{N}^*$, soient $\omega_1, \dots, \omega_n \in \mathbb{D}$ et soit*

$$T_n = \begin{pmatrix} \omega_1 & \alpha_1^{(1)} & \alpha_1^{(2)} & \cdots & \cdots & \alpha_1^{(n-1)} \\ 0 & \omega_2 & \alpha_2^{(1)} & \alpha_2^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \alpha_{n-2}^{(2)} \\ \vdots & & \ddots & \ddots & \ddots & \alpha_{n-1}^{(1)} \\ 0 & \cdots & \cdots & \cdots & 0 & \omega_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{C}).$$

Supposons que

$$\alpha_i^{(1)} = \sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+1}|^2}, \text{ pour tout } 1 \le i \le n-1.$$

Alors, T_n est une contraction si et seulement si

$$\alpha_i^{(j)} = \prod_{k=i+1}^{j+i-1} (-\overline{\omega_k})\sqrt{1-|w_i|^2}\sqrt{1-|\omega_{i+j}|^2}, \text{ pour tout } 1 \le i \le n,\ 1 \le j \le n-i.$$

À nouveau, pour prouver ce théorème, le calcul fonctionnel explicite du Théorème 0.9 nous permettra de simplifier les calculs en supposant que l'un des coefficients diagonaux est égal à 0.

Opérateurs modèles

Nous verrons que le Théorème 0.16 fournit une caractérisation précise des opérateurs modèles (voir Chapitre 2 puis Section 4.4) :

Soit $H^\infty = H^\infty(\mathbb{D})$ l'ensemble des fonctions holomorphes bornées sur $\mathbb{D}$, et soit $H^2 = H^2(\mathbb{D})$ l'espace de Hardy-Hilbert du disque unité, c'est-à-dire l'espace des fonctions f holomorphes sur $\mathbb{D}$ telles que

$$\sup_{0<r<1} \int_{\mathbb{T}} |f(\zeta)|^2 \mathrm{d}m(\zeta) < \infty$$

ou, de manière équivalente, telles que

$$\sum_{n=0}^{\infty} |a_n|^2 < \infty, \quad \text{en écrivant } f(z) = \sum_{n=0}^{\infty} a_n z^n.$$

Si u est une fonction intérieure, *i.e.* une fonction de H^∞ telle que $\lim_{r\to 1^-} |u(r\xi)| = 1$, pour presque tout $\xi \in \mathbb{T}$, on définit l'espace l'espace modèle correspondant $\mathcal{K}_u$ par :

$$\mathcal{K}_u := \left(uH^2(\mathbb{D})\right)^\perp = \left\{f \in H^2(\mathbb{D}) : \langle f, uh\rangle = 0,\ \forall h \in H^2(\mathbb{D})\right\}.$$

Soit maintenant $S : H^2(\mathbb{D}) \to H^2(\mathbb{D})$ le *shift unilatéral*, défini par $S(f)(z) = zf(z)$ ou, en raisonnant avec les coefficients de Taylor $(a_n)_{n\ge 0}$ de f, par $S(a_0, a_1, \dots) = (0, a_0, a_1, \dots)$.

On définit le *shift compressé* associé par $S_u := P_u S|_{\mathcal{K}_u}$, où P_u est la projection orthogonale de $H^2(\mathbb{D})$ sur $\mathcal{K}_u$.

Si Θ_n est un produit de Blaschke fini dont les zéros $\omega_1,\ldots,\omega_n \in \mathbb{D}$ sont deux à deux distincts, et si l'on note $b_{\omega_k}(z) = \frac{z-\omega_k}{1-\overline{\omega_k}z}$ le facteur de Blaschke associé à ω_k, pour $k \in [\![1,n]\!]$, on peut prouver que $\dim\left(\mathcal{K}_{\Theta_n}\right) = n$. Plus précisément, on obtient une base orthonormée de $\mathcal{K}_{\Theta_n}$ en posant :

$$\phi_1(z) = \frac{\sqrt{1-|\omega_1|^2}}{1-\overline{\omega_1}z} \quad \text{et} \quad \phi_k(z) = \left(\prod_{j=1}^{k-1} b_{\omega_j}\right)\frac{\sqrt{1-|\omega_k|^2}}{1-\overline{\omega_k}z} \quad k = 2,\ldots,n.$$

La base $(\phi_1(z),\ldots,\phi_n(z))$ est appelée la base orthonormée de Takenaka–Malmquist–Walsh de $\mathcal{K}_{\Theta_n}$.

En notant M_{Θ_n} la représentation matricielle de S_{Θ_n} dans la base de Takenaka–Malmquist, on a :

$$\left[M_{\Theta_n}\right]_{i,j} = \begin{cases} \omega_j & \text{si } i = j \\ \prod_{k=i+1}^{j-1}(-\overline{\omega_k})\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_j|^2} & \text{si } i < j \\ 0 & \text{si } i > j \end{cases}.$$

Ainsi, avec cette notation, le Théorème 0.16 peut être reformulé comme suit :

Théorème 0.17. *Soit $n \in \mathbb{N}^*$, soient $\omega_1,\ldots,\omega_n \in \mathbb{D}$ et soit*

$$T_n = \begin{pmatrix} \omega_1 & \alpha_1^{(1)} & \alpha_1^{(2)} & \cdots & \cdots & \alpha_1^{(n-1)} \\ 0 & \omega_2 & \alpha_2^{(1)} & \alpha_2^{(2)} & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \alpha_{n-2}^{(2)} \\ \vdots & & & \ddots & \ddots & \alpha_{n-1}^{(1)} \\ 0 & \cdots & \cdots & \cdots & 0 & \omega_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{C}).$$

Supposons que

$$\alpha_i^{(1)} = \sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+1}|^2},\ \textit{pour tout } 1 \le i \le n-1.$$

Alors, T_n est une contraction si et seulement si $T_n = M_{\Theta_n}$.

Inégalités de type Schwarz-Pick

Dans le Chapitre 5, nous utiliserons des versions de l'inégalité de von Neumann pour obtenir des inégalités de type Schwarz-Pick. L'idée principale sera d'appliquer l'inégalité de von Neumann à une contraction (ou un tuple de contractions) *bien choisi* T. Nos inégalités découleront alors du fait que $f(T)$ demeure une contraction, pour toute fonction f de l'algèbre du disque telle que $\|f\|_\infty \le 1$. Pour ce faire, nous utiliserons le calcul fonctionnel explicite du Chapitre 3 ainsi que les caractérisations des matrices contractantes du Chapitre 4.

Nous commencerons par nous concentrer les matrices de taille 3×3, en établissant une preuve opératorielle du *lemme Schwarz-Pick à trois points* obtenu par Beardon et Minda [8] :

Théorème 0.18 (Beardon-Minda)**.** *Soit $f \in \mathcal{H}(\mathbb{D},\mathbb{D})$ une fonction holomorphe de $\mathbb{D}$ dans lui-même,* *et soient ω_1, ω_2, ω_3 trois points distincts de $\mathbb{D}$. Alors, on a :*

$$\mathrm{d}\,(f^*(\omega_1,\omega_2), f^*(\omega_3,\omega_2)) \leq \mathrm{d}(\omega_1,\omega_3). \tag{0.6}$$

où d désigne la distance hyperbolique du disque unité et où $f^*(z,\omega)$ désigne la différence divisée hyperbolique de f aux points z et ω.

De plus, l'égalité a lieu si et seulement si f est un produit de Blaschke de degré ≤ 2.

De même, nous donnerons une preuve opératorielle du résultat suivant, établi par Yamashita [65], qui peut être vu comme un analogue de l'inégalité de Beardon-Minda dans le cas où les trois points sont confondus :

Théorème 0.19. *Soit $f \in \mathcal{H}(\mathbb{D},\mathbb{D})$ une fonction holomorphe sur le disque unité et soit $\Gamma(z,f) := \frac{(1-|z|^2)|f'(z)|}{1-|f(z)|^2}$. Alors, pour tout $\omega \in \mathbb{D}$, on a :*

$$\left|\frac{\partial\Gamma(\omega,f)}{\partial\omega}\right| \leq \frac{1-|\Gamma(\omega,f)|^2}{1-|\omega|^2}. \tag{0.7}$$

De plus, l'égalité a lieu si et seulement si f est un produit de Blaschke de degré ≤ 2.

Cette inégalité peut être reformulée en termes de *dérivées invariantes de Peschl*. Soit $f : \mathbb{D} \to \mathbb{D}$ une fonction holomorphe, et soit $\omega \in \mathbb{D}$. Considérons l'application :

$$g : z \in \mathbb{D} \mapsto \frac{f\left(\frac{z+\omega}{1+\overline{\omega}z}\right) - f(\omega)}{1-\overline{f(\omega)}f\left(\frac{z+\omega}{1+\overline{\omega}z}\right)} \in \mathbb{C}. \tag{0.8}$$

Alors, g est analytique sur $\mathbb{D}$ et $g(0) = 0$. On a alors $g(z) = \sum_{n=1}^{\infty} \frac{D_n f(\omega)}{n!} z^n$, avec $D_n f(z_0) := g^{(n)}(0)$. Les quantités $D_n f(\omega)$ sont appelées *dérivées invariantes de Peschl* (voir par exemple [37]). Les deux premières dérivées invariantes de Peschl sont calculées explicitement comme suit :

$$\begin{aligned} D_1 f(\omega) &= \frac{(1-|\omega|^2)f'(\omega)}{1-|f(\omega)|^2}, \\ D_2 f(\omega) &= \frac{(1-|\omega|^2)^2}{1-|f(\omega)|^2}\left[f''(\omega) - \frac{2\overline{\omega}f'(\omega)}{1-|\omega|^2} + \frac{2\overline{f(\omega)}f'(\omega)^2}{1-|f(\omega)|^2}\right]. \end{aligned}$$

Avec ces notations, l'inégalité de Schwarz-Pick pour les dérivées (0.2) s'écrit $|D_1 f(\omega)| \leq 1$, tandis que l'inégalité (0.7) s'écrit (voir [12, Proposition 3.4]) $|D_2 f(\omega)| \leq 2(1-|D_1 f(\omega)|^2)$.

Par la suite, nous examinerons les inégalités de Schwarz-Pick à n points du point de vue de la théorie des opérateurs. Dans la Section 5.2, nous nous pencherons sur le lien entre l'inégalité Schwarz-Pick à n points de Baribeau-Rivard-Wegert [7] – qui fait intervenir des différences divisées hyperboliques –, le problème d'interpolation de Nevanlinna-Pick et l'inégalité suivante :

$$\|f(M_{\Theta_n})\| \leq 1, \quad \text{pour toute fonction holomorphe } f : \mathbb{D} \to \mathbb{D}, \tag{0.9}$$

qui est une conséquence de l'inégalité de von Neumann.

Nous établirons également *l'inégalité de Schwarz-Pick d'ordre supérieur* suivante, prolongeant ainsi les travaux de Wiener et Ruscheweyh [47, 56] :

Théorème 0.20. *Soit f une fonction analytique de $\mathbb{D}$ dans $\overline{\mathbb{D}}$. Si l'on note $f(z) = \sum_{n=0}^{\infty} a_n z^n$, pour $z \in \mathbb{D}$, alors, pour tout $n \geq 1$ et pour tout $k \geq 1$, on a :*

$$\left|a_{n+k}(1-|a_0^2|) + a_n a_k \overline{a}_0\right|^2 \leq \left[(1-|a_0|^2)^2 - |a_n|^2\right] \cdot \left[(1-|a_0|^2)^2 - |a_k|^2\right]. \tag{0.10}$$

On obtient en particulier le corollaire suivant :

Corollaire 0.21. *Soit f une fonction analytique de $\mathbb{D}$ dans $\overline{\mathbb{D}}$ avec $f(z) = \sum_{n=0}^{\infty} a_n z^n$ pour $z \in \mathbb{D}$. Alors*

$$1 - |a_0|^2 - |a_n| \geq \frac{\left|a_{2n}(1-|a_0^2|) + a_n^2 \overline{a}_0\right|}{2(1-|a_0|^2)}.$$

Ce résultat est une amélioration de l'inégalité de Wiener :

$$|a_k| \leq 1 - |a_0|^2, \quad \forall\, k \geq 1.$$

Par la suite, dans la Section 5.4, nous nous concentrerons sur le polydisque, en exploitant les résultats de Drury [20] et Knese [39] – qui stipulent (respectivement) que l'inégalité de von Neumann s'applique aux n–uplets commutatifs de matrices de taille 2×2 et 3×3 (voir Théorèmes 1.2.11 et 1.2.12) – pour obtenir plusieurs inégalités de type Schwarz-Pick pour le polydisque. Tout d'abord, nous donnerons une preuve opératorielle du résultat suivant :

Théorème 0.22 (Rudin, [54])**.**

(i) Soit $f \in \mathcal{H}(\mathbb{D}^n, \mathbb{D})$ et soient $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_n) \in \mathbb{D}^n$.

On a alors :

$$\left|\frac{f(a_1,\cdots,a_n) - f(b_1,\cdots,b_n)}{1 - \overline{f(a_1,\cdots,a_n)} f(b_1,\cdots,b_n)}\right| \leq \max_{1\leq i\leq n} \left|\frac{a_i - b_i}{1 - \overline{a_i} b_i}\right|. \tag{0.11}$$

(ii) Soit $f \in \mathcal{H}(\mathbb{D}^n, \mathbb{D})$ et soit $a = (a_2, \ldots, a_n) \in \mathbb{D}^n$.

On a alors :

$$\sum_{i=1}^{n} (1-|a_i|^2) \left|\frac{\partial f(a)}{\partial z_i}\right| \leq 1 - |f(a)|^2. \tag{0.12}$$

Ensuite, nous établirons, dans la Section 5.4.2, la généralisation suivante – en plusieurs variables – du Théorème 0.19 :

Théorème 0.23. *Soit $n \in \mathbb{N}^*$, soit $w = (\omega_1, \ldots, \omega_n) \in \mathbb{D}^n$, et soit $f : \mathbb{D}^n \to \mathbb{D}$ une fonction holomorphe. On a alors :*

$$|D_2 f(\omega)| \leq 2(1 - |D_1 f(\omega)|^2), \tag{0.13}$$

où $D_1 f$ et $D_2 f$ désignent les dérivées invariantes de Peschl de f.

Dans le cas particulier du bidisque, nous utiliserons la notion de *variétés distinguées* – introduite par Agler and McCarthy dans [1] – pour obtenir un raffinement du Théorème 0.22.

Une *variété distinguée* est un ensemble de la forme $V \cap \overline{\mathbb{D}}^2$, où V est une variété algébrique dans $\mathbb{C}^2$ telle que $\overline{V} \cap \partial(\mathbb{D}^2) = \overline{V} \cap \mathbb{T}^2$.

Agler et McCarthy ont prouvé dans [1] le résultat suivant :

Théorème 0.24 (Agler-McCarthy, [1]). *Pour toute paire de contractions (T_1,T_2) telle que $T_1T_2 = T_2T_1$ et telle que ni T_1 ni T_2 n'aient de valeurs propres de module 1, il existe une variété distinguée $V\cap\mathbb{D}^2$ telle que l'inégalité de von Neumann soit vérifiée sur $V\cap\mathbb{D}^2$ pour tout polynôme p dans $\mathbb{C}[X_1,X_2]$, c'est-à-dire :*

$$\|p(T_1,T_2)\| \le \sup\{|p(z_1,z_2)| : (z_1,z_2)\in V\cap\mathbb{D}^2\}.$$

Ce théorème est un raffinement du théorème d'Ando, qui affirme que l'inégalité de von Neumann est valable pour toute paire de contractions (T_1,T_2) agissant sur un espace de Hilbert et telles que $T_1T_= T_2T_1$.

En exploitant ce résultat, nous établirons le raffinement suivant du Théorème 0.22 dans le cas du bidisque :

Théorème 0.25.

(i) Soient $(a_1,a_2),(b_1,b_2)\in\mathbb{D}^2$ deux points du bidisque. Alors, il existe une variété distinguée $V\cap\mathbb{D}^2$ telle que l'inégalité de Schwarz-Pick

$$\left|\frac{f(a_1,a_2)-f(b_1,b_2)}{1-\overline{f(a_1,a_2)}f(b_1,b_2)}\right| \le \max\left\{\left|\frac{a_1-b_1}{1-\overline{a_1}b_1}\right|,\left|\frac{a_2-b_2}{1-\overline{a_2}b_2}\right|\right\} \tag{0.14}$$

soit vérifiée pour toute fonction f holomorphe sur $\mathbb{D}^2$ et continue sur $\overline{\mathbb{D}}^2$ telle

$$\sup\{|f(z_1,z_2)| \le 1 : (z_1,z_2)\in V\cap\mathbb{D}^2\}.$$

(ii) Soient $(a_1,a_2),(b_1,b_2)\in\mathbb{D}^2$ deux points du bidisque. Alors, il existe une variété distinguée $V\cap\mathbb{D}^2$ telle que l'inégalité de Schwarz-Pick soit vérifiée pour toute fonction f holomorphe sur $\mathbb{D}^2$ pour laquelle il existe une suite de réels $(r_n)_{n\in\mathbb{N}^*}\subset]0,1[$ convergeant vers 1 telle que :

$$\sup\{|f(r_nz_1,r_nz_2)| \le 1 : n\ge 1,(z_1,z_2)\in V\cap\mathbb{D}^2\}.$$

Enfin, pour conclure le Chapitre 5, nous nous pencherons sur des inégalités de type Schwarz-Pick pour les opérateurs, en prolongeant le travail de Fan [22] et, plus récemment, de Jocić [33], qui ont établi des versions opératorielles du lemme de Schwarz. Nous établirons ainsi les deux résultat suivants à partir du théorème de Parrott et du Théorème 0.15 :

Théorème 0.26. *Soient H_1,H_2 deux espaces de Hilbert, et soient $W_1\in\mathcal{B}(H_1)$, $W_2\in\mathcal{B}(H_2)$, $V\in\mathcal{B}(H_2,H_1)$ trois contractions. On suppose que $\sigma(W_1)\cap\sigma(W_2)=\emptyset$, et que $f\in\mathcal{H}(\mathbb{D},\mathbb{D})$ est holomorphe sur un voisinage ouvert de $\sigma(W_1)\cup\sigma(W_2)$. On note $X = X_{W_1,W_2,V}$ l'unique solution de l'équation*

$$W_1X - XW_2 = D_{W_1^*}VD_{W_2}.$$

Alors, il existe une contraction $Y\in\mathcal{B}(H_2,H_1)$ telle que :

$$f(W_1)X - Xf(W_2) = D_{f(W_1)^*}YD_{f(W_2)}.$$

Théorème 0.27. *Soient H_1,H_2,H_3 trois espaces de Hilbert, et soient $W_1\in\mathcal{B}(H_1)$, $W_2\in\mathcal{B}(H_2)$ et $W_3\in\mathcal{B}(H_3)$ trois contractions. On se donne également trois contractions $V_1\in\mathcal{B}(H_2,H_1)$, $V_2\in\mathcal{B}(H_3,H_2)$, et $V_3\in\mathcal{B}(H_3,H_1)$. On suppose que $\sigma(W_i)\cap\sigma(W_j)=\emptyset$, pour tous $1\le i<j\le 3$. Soit $f:\mathbb{D}\to\mathbb{D}$ une fonction holomorphe sur un voisinage ouvert de $\sigma(W_1)\cup\sigma(W_2)\cup\sigma(W_3)$. On note respectivement*

X_1, X_2, X_3 les uniques solutions (respectives) des équations de Sylvester :

$$\begin{aligned} W_1X_1 - X_1W_2 &= D_{W_1^*}V_1D_{W_2}, \\ W_2X_2 - X_2W_3 &= D_{W_2^*}V_2D_{W_3}, \\ \textit{et } W_1X_3 - X_3W_3 &= B - W_1X_1X_2 + X_1W_2X_2, \end{aligned}$$

où

$$B = \left[D_{W_1^*}(\mathrm{Id} - V_1V_1^*)D_{W_1^*}\right]^{1/2} V_3 \left[D_{W_3}(\mathrm{Id} - V_2^*V_2)D_{W_3}\right]^{1/2} - D_{W_1^*}V_1W_2^*V_2D_{W_3}.$$

Alors, il existe trois contractions $Y_1 \in \mathcal{B}(H_2, H_1)$, $Y_2 \in \mathcal{B}(H_3, H_2)$, $Y_3 \in \mathcal{B}(H_3, H_1)$ telles que :

$$\begin{cases} f(W_1)X_1 - X_1f(W_2) = D_{f(W_1)^*}Y_1D_{f(W_2)}, & (0.15) \\ f(W_2)X_2 - X_2f(W_3) = D_{f(W_2)^*}Y_2D_{f(W_3)}, & (0.16) \\ f(W_1)X_3 - X_3f(W_3) = X_1f(W_2)X_2 - f(W_1)X_1X_2 & \\ \quad + \left[D_{f(W_1)^*}(\mathrm{Id} - Y_1Y_1^*)D_{f(W_1)^*}\right]^{1/2} Y_3 \left[D_{f(W_3)}(\mathrm{Id} - Y_2^*Y_2)D_{f(W_3)}\right]^{1/2} & (0.17) \\ \qquad\qquad - D_{f(W_1)^*}Y_1f(W_2)^*Y_2D_{f(W_3)}. & \end{cases}$$

On note que, dans le cas scalaire, ces deux résultats sont respectivement équivalents à l'inégalité de Schwarz-Pick (Théorème 0.2) et à l'inégalité de Beardon-Minda (Théorème 0.18).

Ensembles spectraux

En 1951, von Neumann a introduit dans [63] la notion d'ensemble spectral, afin d'estimer la norme des fonctions de matrices en fonction de la norme *sup* de la fonction :

Définition 0.28. Soit $T \in \mathcal{B}(H)$, soit $S \subset \mathbb{C}$ une partie du plan complexe contenant $\sigma(T)$, et soit $K > 0$ une constante réelle.

On dit que S est un *ensemble K-spectral* pour T si, pour toute fonction rationnelle f dont les pôles sont en dehors de S, on a :

$$\|f(T)\| \leq K \cdot \|f\|_S,$$

où $\|f\|_S = \sup\{|f(z)| : z \in S\}$.

Si $K = 1$, on dit que S est un *ensemble spectral* pour T.

Par exemple, l'inégalité de von Neumann affirme que le disque unité fermé est un ensemble spectral pour toute contraction.

Une notion analogue d'ensemble K-spectral peut être définie pour un n-uplet d'opérateurs commutant deux à deux.

Dans le dernier chapitre de ce manuscrit, nous examinerons brièvement deux autres ensembles spectraux : la boule unité ouverte de $\mathbb{C}^n$ et la couronne $\mathbb{A}_r := \{z \in \mathbb{C} : r < |z| < 1\}$, avec $r \in [0, 1[$.

Tout d'abord, concernant la boule unité de $\mathbb{C}^n$, notée $\mathbb{B}_n$, Hartz, Richter et Shalit ont prouvé dans [30] l'inégalité suivante :

Définition 0.29. Soit H un espace de Hilbert, et soit $T = (T_1, \ldots, T_n)$ un n-uplet d'opérateurs linéaires bornés sur H. On dit que T est une *contraction en rang* si $\|T\|^2 = \|\sum_{i=1}^n T_iT_i^*\| \leq 1$.

Théorème 0.30 (Hartz-Richter-Shalit). *Pour tout entier $d \geq 1$, il existe une constante C_d telle que, pour tout $n \in \mathbb{N}^*$, pour tout n-uplet $(T_1, \ldots, T_n)$ de matrices de taille $d \times d$ commutant deux à deux tel que $\|T\|^2 = \|\sum_{i=1}^n T_i T_i^*\| \leq 1$, et pour tout polynôme $p \in \mathbb{C}[X_1, \ldots, X_n]$, on ait :*

$$\|p(T)\| \leq C_d \cdot \sup\{|p(z)| : z \in \mathbb{B}_n\}.$$

De plus, on peut choisir $C_2 = 1$.

Nous commencerons donc par établir un critère explicite pour qu'un n-uplet $(T_1, \ldots, T_n)$ de matrices triangulaires supérieures de taille 2×2 commutant deux à deux soit une *contraction en rang* puis, en exploitant l'inégalité de Hartz-Richter-Shalit, nous obtiendrons le lemme de type Schwarz-Pick suivant pour la boule unité, qui est un cas particulier du *lemme de Schwarz pour la boule unité* obtenu par Rudin [55, Theorem 8.1.4] :

Théorème 0.31. *Soit $f : \mathbb{B}_n \to \mathbb{D}$ une fonction holomorphe, et soient $\omega, z \in \mathbb{B}_n$.*

Alors, on a :

$$\frac{\left|1 - \overline{f(\omega)} f(z)\right|^2}{(1 - |f(z)|^2)(1 - |f(\omega)|^2)} \leq \frac{|1 - \langle z, \omega \rangle|^2}{(1 - \|z\|^2)(1 - \|\omega\|^2)}.$$

Ensuite, concernant la couronne $\mathbb{A}_r$, nous nous concentrerons sur les travaux de Tsikalas [61], qui a introduit la classe d'opérateurs linéaires bornés

$$\mathcal{F}_r := \left\{T \in \mathcal{B}(H) : r^2 T^{-1} \left(T^{-1}\right)^* + T T^* \leq r^2 + 1, \sigma(T) \subset \mathbb{A}_r\right\},$$

où H désigne un espace de Hilbert et $r \in [0, 1[$.

Tsikalas a prouvé que, pour tout $r \in [0, 1[$, $\mathbb{A}_r$ est un ensemble $\sqrt{2}$-spectral pour les éléments de $\mathcal{F}_r$:

Théorème 0.32 (Tsikalas). *Pour toute fonction holomorphe bornée $\phi \in H^\infty(\mathbb{A}_r)$, et pour tout $T \in \mathcal{F}_r$, on a :*

$$\|\phi(T)\| \leq \sqrt{2} \sup\{|\phi(z)| : z \in \mathbb{A}_r\},$$

où $\sqrt{2}$ est la meilleure constante possible.

Pour conclure cet ultime chapitre, nous établirons la caractérisation suivante de la classe $\mathcal{F}_r$ pour les matrices triangulaires supérieures de taille 2×2 :

Proposition 0.33. *Soit $T = \begin{pmatrix} \omega_1 & \alpha \\ 0 & \omega_2 \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$, avec $\omega_1, \omega_2 \in \mathbb{A}_r$.*

$T \in \mathcal{F}_r$ si et seulement si

$$|\alpha|^2 \leq \frac{(1 - |\omega_1|^2)(1 - |\omega_2|^2)(|\omega_1|^2 - r^2)(|\omega_2|^2 - r^2)}{|r^2 - \overline{\omega_2}\omega_1|^2 + r^2(1 - |\omega_1|^2)(1 - |\omega_2|^2)}.$$

Introduction (in English)

Motivation

The starting point of this manuscript is the somewhat surprising observation that the Schwarz-Pick lemma, a classical result in complex analysis, can be derived as a special case of the von Neumann inequality, a result concerning Hilbert space operators. Let us first recall the statements of these two results:

Theorem 0.1 (von Neumann's inequality)**.** *Let T be a contraction acting on a Hilbert space H, i.e. T is a bounded linear operator on H such that $\|T\| \leq 1$.*

Then, for every polynomial $p \in \mathbb{C}[X]$, we have:

$$\|p(T)\| \leq \|p\|_\infty := \sup\{|p(z)| \,:\, |z| \leq 1\}.$$

Theorem 0.2 (Schwarz-Pick's lemma)**.**

(i) Let $f : \mathbb{D} \to \mathbb{D}$ be an holomorphic function, and let $\omega_1, \omega_2 \in \mathbb{D}$ be two distinct points of the unit disc. Then, we have:

$$\left|\frac{f(\omega_1) - f(\omega_2)}{1 - \overline{f(\omega_1)} f(\omega_2)}\right| \leq \left|\frac{\omega_1 - \omega_2}{1 - \overline{\omega_1}\omega_2}\right| \tag{0.1}$$

If we denote by $\rho(\omega_1, \omega_2) := \left|\frac{\omega_1 - \omega_2}{1 - \overline{\omega_1}\omega_2}\right|$ the pseudo-hyperbolic distance on the unit disc, this inequality can be written as:

$$\rho(f(\omega_1), f(\omega_2)) \leq \rho(\omega_1, \omega_2)$$

In other words, every holomorphic function mapping $\mathbb{D}$ to $\mathbb{D}$ is contractive with respect to the pseudo-hyperbolic metric.

(ii) If $\omega_1 = \omega_2 =: \omega \in \mathbb{D}$, this inequality becomes:

$$|f'(\omega)| \leq \frac{1 - |f(\omega)|^2}{1 - |\omega|^2} \tag{0.2}$$

It is known (see for example [46, Ex. 2.17–2.18] or [52, p. 17]) that the Schwarz-Pick inequality can be obtained as a special case of the von Neumann inequality, applied to a suitably chosen 2×2 matrix. The idea is the following: let $\omega_1, \omega_2 \in \mathbb{D}$ be two distinct points of the open unit disc, and let $T_2 = \begin{pmatrix} \omega_1 & \alpha \\ 0 & \omega_2 \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$. We first prove the two following lemmas:

Lemma 0.3. *T_2 is a contraction if and only if*

$$|\alpha|^2 \leq (1-|\omega_1|^2)(1-|\omega_2|^2).$$

Lemma 0.4. *For every polynomial $p \in \mathbb{C}[X]$, we have:*

$$p(T_2) = \begin{pmatrix} p(\omega_1) & \alpha \cdot \left(\frac{p(\omega_1)-p(\omega_2)}{\omega_1-\omega_2}\right) \\ 0 & p(\omega_2) \end{pmatrix}.$$

Then, we let $\alpha = \sqrt{1-|\omega_1|^2}\sqrt{1-|\omega_2|^2}$, and we apply the von Neumann inequality to T_2. After some computation, we get the Schwarz-Pick inequality for polynomials. We can then conclude the proof by an approximation argument. Further details will be given in Chapter 1, after reviewing some fundamental concepts in spectral theory.

In this thesis, we extend this observation to matrices of larger sizes and to matrices with operator coefficients. In doing so, we encounter the following two problems:

Problem 0.5. *Given a $n \times n$ matrix T_n and a polynomial – or a more general function f – how can we explicitly express $f(T_n)$?*

Problem 0.6. *Given a $n \times n$ matrix T_n, how can we estimate its spectral norm?*

In light of the Schur decomposition theorem – which states that every matrix is unitarily equivalent to an upper-triangular matrix – the study can be limited to upper-triangular matrices.

Function of a matrix : an explicit functional calculus

The first issue has already been studied in several references (see *e.g.* [33, 36, 16, 31, 40]), from different perspectives. In this manuscript, we develop a slightly different approach. In Chapter 3, we introduce the notion of *non-commutative divided differences* for operators– a generalization of the well-known divided differences for scalars. We start with polynomials and then extend this concept to rational functions and, finally, to holomorphic maps.

Recall that, in the scalar case, if $x_0 \ldots, x_n$ are $n+1$ pairwise distinct points of $\mathbb{C}$ and if f is a map defined on any subset of $\mathbb{C}$ containing $x_0, \ldots, x_n$, the *divided differences* are defined recursively as follows:

$$[f(x_k)] := f(x_k),\ \forall\, k \in [\![0,n]\!]$$
$$[f(x_k), \ldots, f(x_{k+j})] := \frac{[f(x_{k+1}), \ldots, f(x_{k+j})] - [f(x_k), \ldots, f(x_{k+j-1})]}{x_{k+j} - x_k}, \forall\, j \in [\![1,n]\!],\ \forall k \in [\![0, n-j]\!].$$

Depending on the regularity of f, we can then obtain several explicit formula for the divided differences. This will be discussed in Section 3.1.

Now, we extend this notion to operators and we define the *non-commutative divided differences* as follows (see Definitions 3.2.1, 3.2.11, 3.2.13 and 3.2.20):

Definition 0.7. Let $H_0,\dots,H_n$ be Hilbert spaces. Given bounded linear operators $W_i \in \mathcal{B}(H_i)$, for $i \in [\![0,n]\!]$, and $C_j \in \mathcal{B}\left(H_{j+1},H_j\right)$, for $j \in [\![0,n-1]\!]$, we define:

(i) $[W_0^k,\dots,W_n^k]_{(C_0,\dots,C_{n-1})} := \sum_{\substack{i_0+\dots+i_n=k-n\\ i_0,\dots,i_n\geq 0}} W_0^{i_0}C_0W_1^{i_1}\dots C_{n-1}W_n^{i_n}$, $\forall\, k \geq 0$.

This definition can be extended by linearity for polynomials. Given a polynomial $p(X) = \sum_{k=0}^n a_kX^k \in \mathbb{C}[X]$, we set:

$$[p(W_0),\dots,p(W_n)]_{(C_0,\dots,C_{n-1})} = \sum_{k=0}^n a_k[W_0^k,\dots,W_n^k]_{(C_0,\dots,C_{n-1})}.$$

Then, for rational fractions, we define:

(ii) $\left[W_0^{-k},\dots,W_n^{-k}\right]_{(C_0,\dots,C_{n-1})} := (-1)^n\sum_{\substack{i_0+\dots+i_n=k-n\\ i_0,\dots,i_n\geq 1}} W_0^{-i_0}C_0W_1^{-i_1}\dots C_{n-1}W_n^{-i_n}$, $\forall\, k \geq 1$;

Using the partial fraction decomposition, this definition can also be extend by linearity for arbitrary rational fractions.

Finally, for any function f holomorphic in the neighborhood of $\bigcup_{i=0}^n \sigma(W_i)$, we define:

(iii) $[f(W_0),\dots,f(W_n)]_{(C_0,\dots,C_{n-1})} := \int_\Gamma f(\xi)(\xi\mathrm{Id}-W_0)^{-1}C_0(\xi\mathrm{Id}-W_1)^{-1}\dots C_{n-1}(\xi\mathrm{Id}-W_n)^{-1}\mathrm{d}\xi$, where Γ is a finite system of rectifiable, positively oriented, Jordan curves surrounding $\bigcup_{i=0}^n \sigma(W_i)$.

We will see that – in analogy with the classical divided differences – the *non-commutative divided differences* satisfy the following recursive relation:

Proposition 0.8. *Let $n \geq 1$, and let $H_0,\dots,H_n$ be Hilbert spaces. Given bounded linear operators $W_i \in \mathcal{B}(H_i)$, for $i \in [\![0,n]\!]$, and $C_j \in \mathcal{B}(H_j,H_{j-1})$, for $j \in [\![1,n]\!]$, and given a map f which is analytic in the neighbourhood of $\bigcup_{i=0}^n \sigma(W_i)$, we have:*

$$\begin{aligned}
&[f(W_0)] = f(W_0);\\
&W_0[f(W_0),\dots,f(W_n)]_{(C_0,\dots,C_{n-1})} - [f(W_0),\dots,f(W_n)]_{(C_0,\dots,C_{n-1})}W_n\\
&\quad = [f(W_0),\dots,f(W_{n-1})]_{(C_1,\dots,C_{n-2})}C_{n-1} - C_0[f(W_1),\dots,f(W_n)]_{(C_1,\dots,C_{n-1})}.
\end{aligned}$$

This notion of *non-commutative divided differences* will allow us to develop an explicit functional calculus for upper-triangular matrices with operator coefficients (see Theorem 3.2.6, Theorem 3.2.16 and Theorem 3.2.22):

Theorem 0.9. *Let $n \geq 2$, let $H_1,\dots,H_n$ be Hilbert spaces, and let*

$$\widetilde{T}_n = \begin{pmatrix} W_1 & C_1^{(1)} & \cdots & C_1^{(n-1)} \\ 0 & W_2 & C_2^{(1)} \cdots & C_2^{(n-2)} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & C_{n-1}^{(1)} \\ 0 & \cdots & 0 & W_n \end{pmatrix} : \bigoplus_{i=1}^n H_i \to \bigoplus_{i=1}^n H_i$$

be a bounded linear operator. Let f be a map which is analytic in the neighborhood of $\bigcup_{i=0}^n \sigma(W_i)$.

Then, we have :

$$f\left(\widetilde{T}_n\right) = \begin{pmatrix} f(W_1) & D_1^{(1)} & \cdots & \cdots & D_1^{(n-1)} \\ 0 & f(W_2) & D_2^{(1)} & \cdots & D_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & D_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & f(W_n) \end{pmatrix},$$

$$\text{where} \quad D_i^{(j)} = \sum_{\substack{1\le l\le j \\ i=k_0<\cdots<k_l=j+i}} \left[f\left(W_{k_0}\right),\ldots,f\left(W_{k_l}\right)\right]_{\left(C_{k_0}^{(k_1-k_0)},\ldots,C_{k_{l-1}}^{(k_l-k_{l-1})}\right)},$$

$$\text{and where the} \quad \left[f\left(W_{k_0}\right),\ldots,f\left(W_{k_l}\right)\right]_{\left(C_{k_0}^{(k_1-k_0)},\ldots,C_{k_{l-1}}^{(k_l-k_{l-1})}\right)}\text{'s}$$

denote the non-commutative divided differences.

The analogy between the non-commutative divided differences for operators and the usual divided differences for the scalars will then allow us to deduce the following corollary:

Corollary 0.10. *Let*

$$T_n = \begin{pmatrix} \omega_1 & \alpha_1^{(1)} & \cdots & \cdots & \alpha_1^{(n-1)} \\ 0 & \omega_2 & \alpha_2^{(1)} & \cdots & \alpha_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \alpha_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & \omega_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$$

be a matrix such that $w_i \neq w_j$, for all $i \neq j$, and let f be a function holomorphic in the neighborhood of $\omega_1,\ldots,\omega_n$.

Then, we have:

$$f(T_n) = \begin{pmatrix} f(\omega_1) & \beta_1^{(1)} & \cdots & \cdots & \beta_1^{(n-1)} \\ 0 & f(\omega_2) & \beta_2^{(1)} & \cdots & \beta_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \beta_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & f(\omega_n) \end{pmatrix},$$

$$\text{where} \quad \beta_i^{(j)} = \sum_{\substack{1\le l\le j \\ i=k_0<\cdots<k_l=j+i}} \left(\prod_{s=0}^{l-1} \alpha_{k_s}^{(k_{s+1}-k_s)}\right)\left[f\left(\omega_{k_0}\right),\ldots,f\left(\omega_{k_l}\right)\right],$$

$$\text{and where the} \quad \left[f\left(\omega_{k_0}\right),\ldots,f\left(\omega_{k_l}\right)\right]'s$$

denote the divided differences.

Furthermore, the recursive relation of Proposition 0.8 will lead us to focus on the Sylvester equation:

$$AX - XB = Y, \tag{0.3}$$

where A, X, B, Y are bounded linear operators (acting on a Banach space).

In Section 3.3, we will see that, in the case where the spectra of the W_i's are pairwise disjoint, the recursive relation of Proposition 0.8 defines completely the *non-commutative divided differences*. This will lead to the following result:

Theorem 0.11. *Let $n \geq 2$, let $H_1, \ldots, H_n$ be Hilbert spaces, and let*

$$\widetilde{T_n} = \begin{pmatrix} W_1 & C_1^{(1)} & \cdots & \cdots & C_1^{(n-1)} \\ 0 & W_2 & C_2^{(1)} & \cdots & C_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & C_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & W_n \end{pmatrix} : \bigoplus_{i=1}^{n} H_i \to \bigoplus_{i=1}^{n} H_i$$

be a bounded linear operator. Let f be a map which is analytic in the neighborhood of $\bigcup_{i=0}^{n} \sigma(W_i)$. Assume that $\sigma(W_i) \cap \sigma(W_j) = \emptyset$, for all $i \neq j$.

Then, we have:

$$f\left(\widetilde{T_n}\right) = \begin{pmatrix} \mathrm{Id} & -X_1^{(1)} & \cdots & \cdots & -X_1^{(n-1)} \\ 0 & \mathrm{Id} & -X_2^{(1)} & \cdots & -X_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & -X_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & \mathrm{Id} \end{pmatrix} \begin{pmatrix} f(W_1) & 0 & \cdots & \cdots & 0 \\ 0 & f(W_2) & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & f(W_n) \end{pmatrix} \begin{pmatrix} \mathrm{Id} & Y_1^{(1)} & \cdots & \cdots & Y_1^{(n-1)} \\ 0 & \mathrm{Id} & Y_2^{(1)} & \cdots & Y_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & Y_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & \mathrm{Id} \end{pmatrix}$$

where, for all $j \in [\![1, n-1]\!]$, for all $i \in [\![1, n+1-j]\!]$, $X_i^{(j)}$ is the unique solution of the equation

$$W_i X_i^{(j)} - X_i^{(j)} W_{i+j} = C_i^{(j)} - \sum_{k=1}^{j-1} C_i^{(k)} X_{k+i}^{(j-k)} \tag{0.4}$$

and

$$Y_i^{(j)} = \sum_{\substack{1 \leq l \leq j \\ i = k_0 < \cdots < k_l = j+i}} X_{k_0}^{(k_1 - k_0)} \ldots X_{k_{l-1}}^{(k_l - k_{l-1})} \tag{0.5}$$

Contractive matrices

The second issue (Problem 0.6, which is about estimating the spectral norm of a matrix) will be dealt with in Chapter 4. A straightforward approach for estimating the spectral norm of a matrix would be – at least for matrices with scalars coefficients – to try to exploit the formula $\|T_n\|^2 = \|T_n^* T_n\| = r(T_n^* T_n)$, where $r(\cdot)$ denotes the spectral radius. This method works very well for 2×2 matrices (see Lemma 1.2.14). However, for $n \geq 3$, the computations become too intricate to provide a practical criterion. In this manuscript, we will use the following sharpened version of Parrott's theorem on matrix completion:

Theorem 0.12 (Parrott). *Let H_1, H_2, K_1, K_2 be Hilbert spaces, and assume that the operators* $\begin{bmatrix} A \\ C \end{bmatrix} \in \mathcal{B}(H_1, K_1 \oplus K_2)$ *and* $\begin{bmatrix} C & D \end{bmatrix} \in \mathcal{B}(H_1 \oplus H_2, K_2)$ *are contractions.*

Then,

$$T = \begin{bmatrix} A & B \\ C & D \end{bmatrix} : H_1 \oplus H_2 \to K_1 \oplus K_2$$

is a contraction if and only if there exists a contraction $W \in \mathcal{B}(H_2, K_1)$ such that:

$$B = D_{Z^*} W D_Y - Z C^* Y,$$

where $Z \in \mathcal{B}(H_1, K_1)$ and $Y \in \mathcal{B}(H_2, K_2)$ are contractions such that $D = D_{C^*} Y$ and $A = Z D_C$, and where D_S denotes the defect operator of a contraction S.

Moreover,

(i) Y and Z can be chosen to be (respectively) the unique solutions Y_0 and Z_0 of minimal operator norm among all solutions of the operator equations $D = D_{C^*} Y$ and $A = Z D_C$;

(ii) If T is a contraction, there exists a unique contraction W_0 such that:

$$B = D_{Z_0^*} W_0 D_{Y_0} - Z_0 C^* Y_0 \text{ and } Im\left(D_{Z_0^*}\right)^{\perp} \subset Ker(W_0^*).$$

This operator satisfies:

$$\|W_0\| = \inf\{\|W\| : B = D_{Z_0^*} W D_{Y_0} - Z_0 C^* Y_0\}.$$

We will derive from this theorem the following results for 3×3 and 4×4 matrices (see Theorem 4.2.2 and Theorem 4.3.1 for a more detailed version of these results):

Theorem 0.13. *Let $\omega_1, \omega_2, \omega_3 \in \mathbb{D}$. Then,* $T = \begin{pmatrix} \omega_1 & \alpha_1 & \beta \\ 0 & \omega_2 & \alpha_2 \\ 0 & 0 & \omega_3 \end{pmatrix} \in \mathcal{M}_3(\mathbb{C})$ *is a contraction when acting on the Hilbert space $\mathbb{C}^3$ if and only if:*

$$\begin{cases} |\alpha_i|^2 \le (1 - |\omega_i|^2)(1 - |\omega_{i+1}|^2), \quad i = 1, 2, \\ \left|\beta(1 - |\omega_2|^2) + \alpha_1 \alpha_2 \overline{\omega_2}\right|^2 \le \left[(1 - |\omega_1|^2)(1 - |\omega_2|^2) - |\alpha_1|^2\right] \cdot \left[(1 - |\omega_2|^2)(1 - |\omega_3|^2) - |\alpha_2|^2\right]. \end{cases}$$

Theorem 0.14. *Let $\omega_1, \omega_2, \omega_3, \omega_4 \in \mathbb{D}$. Then,* $T = \begin{pmatrix} \omega_1 & \alpha_1 & \beta_1 & \gamma \\ 0 & \omega_2 & \alpha_2 & \beta_2 \\ 0 & 0 & \omega_3 & \alpha_3 \\ 0 & 0 & 0 & \omega_4 \end{pmatrix} \in \mathcal{M}_4(\mathbb{C})$ *is a contraction* *when acting on the Hilbert space $\mathbb{C}^4$ if and only if:*

$$\begin{cases} |\alpha_i|^2 \leq (1-|\omega_i|^2)(1-|\omega_{i+1}|^2),\ i=1,3 \quad \& \quad |\alpha_2|^2 < (1-|\omega_2|^2)(1-|\omega_3|^2) \\ |\beta_i(1-|\omega_{i+1}|^2)+\alpha_i\alpha_{i+1}\overline{\omega_{i+1}}|^2 \leq \Big[(1-|\omega_i|^2)(1-|\omega_{i+1}|^2)-|\alpha_i|^2\Big]\times \\ \qquad\qquad \Big[(1-|\omega_{i+1}|^2)(1-|\omega_{i+2}|^2)-|\alpha_{i+1}|^2\Big], \quad i=1,2 \\ \Big|\gamma[(1-|\omega_2|^2)(1-|\omega_3|^2)-|\alpha_2|^2]+\alpha_1\beta_2\overline{\omega_2}(1-|\omega_3|^2)+\alpha_3\beta_1\overline{\omega_3}(1-|\omega_2|^2) \\ \qquad\qquad +\beta_1\beta_2\overline{\alpha_2}+\alpha_1\alpha_2\alpha_3\overline{\omega_2\omega_3}\Big|^2(1-|\omega_2|^2)(1-|\omega_3|^2) \\ \quad \leq \Big[\Big((1-|\omega_1|^2)(1-|\omega_2|^2)-|\alpha_1|^2\Big)\Big((1-|\omega_2|^2)(1-|\omega_3|^2)-|\alpha_2|^2\Big)-\Big|\alpha_1\alpha_2\overline{\omega_2}+\beta_1(1-|\omega_2|^2)\Big|^2\Big] \\ \quad \times\Big[\Big((1-|\omega_2|^2)(1-|\omega_3|^2)-|\alpha_2|^2\Big)\Big((1-|\omega_3|^2)(1-|\omega_4|^2)-|\alpha_3|^2\Big)-\Big|\alpha_2\alpha_3\overline{\omega_3}+\beta_2(1-|\omega_3|^2)\Big|^2\Big] \end{cases}$$

or

$$\begin{cases} |\alpha_i|^2 \leq (1-|\omega_i|^2)(1-|\omega_{i+1}|^2),\ i=1,3 \quad \& \quad |\alpha_2|^2 = (1-|\omega_2|^2)(1-|\omega_3|^2) \\ \beta_i = \dfrac{-\alpha_i\alpha_{i+1}\omega_{i+1}}{1-|\omega_{i+1}|^2},\ i=1,2 \\ \left|\gamma-\dfrac{\overline{\omega_2\omega_3}\alpha_1\alpha_2\alpha_3}{(1-|\omega_2|^2)(1-|\omega_3|^2)}\right|^2\Big(1-|\omega_2|^2\Big)\Big(1-|\omega_3|^2\Big) \\ \quad \leq \Big[(1-|\omega_1|^2)(1-|\omega_2|^2)-|\alpha_1|^2\Big]\times\Big[(1-|\omega_3|^2)(1-|\omega_4|^2)-|\alpha_3|^2\Big] \end{cases}$$

We will see that in the case of 4×4 matrices, the computations become more complex. The explicit functional calculus of Theorem 0.9 will allow us to simplify them, by assuming that one of the diagonal entries is equal to zero.

Then, we will also obtain the following generalization of Theorem 0.13 for matrices with operators coefficients:

Theorem 0.15. *Let H_1, H_2, H_3 be three Hilbert spaces. Let $W_i \in \mathcal{B}(H_i)$, $1\leq i\leq 3$, be three contractions and let*

$$T = \begin{bmatrix} W_1 & A_1 & B \\ 0 & W_2 & A_2 \\ 0 & 0 & W_3 \end{bmatrix} \in \mathcal{B}(H_1\oplus H_2\oplus H_3).$$

Then, T is a contraction if and only if there exist three contractions $V_1 \in \mathcal{B}(H_2,H_1)$, $V_2\in\mathcal{B}(H_3,H_2)$, $V_3 \in \mathcal{B}(H_3,H_1)$ such that:

$$\begin{cases} A_1 = D_{W_1^*}V_1D_{W_2}, \\ A_2 = D_{W_2^*}V_2D_{W_3}, \\ B = \Big[D_{W_1^*}(\mathrm{Id}-V_1V_1^*)D_{W_1^*}\Big]^{1/2}V_3\Big[D_{W_3}(\mathrm{Id}-V_2^*V_2)D_{W_3}\Big]^{1/2}-D_{W_1^*}V_1W_2^*V_2D_{W_3}. \end{cases}$$

Furthermore, in attempting to extend Theorem 0.13 and Theorem 0.14 to $n\times n$ matrices, we will establish the following theorem, which characterizes contractions with prescribed main and upper diagonals:

Theorem 0.16. *Let $n \in \mathbb{N}^*$, let $\omega_1, \ldots, \omega_n \in \mathbb{D}$ and let*

$$T_n = \begin{pmatrix} \omega_1 & \alpha_1^{(1)} & \alpha_1^{(2)} & \cdots & \cdots & \alpha_1^{(n-1)} \\ 0 & \omega_2 & \alpha_2^{(1)} & \alpha_2^{(2)} & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \alpha_{n-2}^{(2)} \\ \vdots & & \ddots & \ddots & \ddots & \alpha_{n-1}^{(1)} \\ 0 & \cdots & \cdots & \cdots & 0 & \omega_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{C}).$$

Assume that

$$\alpha_i^{(1)} = \sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+1}|^2}, \textit{ for all } 1 \le i \le n-1.$$

Then, T_n is a contraction if and only if

$$\alpha_i^{(j)} = \prod_{k=i+1}^{j+i-1} (-\overline{\omega_k})\sqrt{1-|w_i|^2}\sqrt{1-|\omega_{i+j}|^2}, \textit{ for all } 1 \le i \le n,\ 1 \le j \le n-i.$$

Again, to prove this theorem, we will use the explicit functional calculus of Theorem 0.9 to simplify the computations, by assuming that one of the diagonal entries is equal to zero.

Model operators

In fact, we will see that Theorem 0.16 provides a precise characterization of model operators (see Chapter 2 and, then, Section 4.4):

Let $H^\infty = H^\infty(\mathbb{D})$ be the set of all holomorphic functions that are bounded on $\mathbb{D}$, and let $H^2 = H^2(\mathbb{D})$ be the Hardy-Hilbert space of $\mathbb{D}$, which is the space of all holomorphic functions $f \in \mathcal{H}(\mathbb{D})$ such that

$$\sup_{0<r<1} \int_{\mathbb{T}} |f(\zeta)|^2 \mathrm{d}m(\zeta) < \infty$$

or, equivalently, such that

$$\sum_{n=0}^{\infty} |a_n|^2 < \infty \quad \text{if } f(z) = \sum_{n=0}^{\infty} a_n z^n.$$

If u is an inner function, *i.e.* a function of H^∞ such that $\lim_{r\to 1^-} |u(r\xi)| = 1$, for almost every $\xi \in \mathbb{T}$, the corresponding *model space* $\mathcal{K}_u$ is defined to be

$$\mathcal{K}_u := \left(uH^2(\mathbb{D})\right)^\perp = \left\{f \in H^2(\mathbb{D}) : \langle f, uh\rangle = 0, \forall h \in H^2(\mathbb{D})\right\}.$$

Now, let $S : H^2(\mathbb{D}) \to H^2(\mathbb{D})$ be the *unilateral shift*, defined by $S(f)(z) = zf(z)$ or, in terms of the Taylor coefficients $(a_n)_{n\ge 0}$ of f, by $S(a_0, a_1, \ldots) = (0, a_0, a_1, \ldots)$.

We define the associated *compressed shift* by $S_u := P_u S|_{\mathcal{K}_u}$, where P_u is the orthogonal projection from $H^2(\mathbb{D})$ onto $\mathcal{K}_u$.

If Θ_n be a finite Blaschke product with pairwise distinct zeros $\omega_1, \ldots, \omega_n \in \mathbb{D}$, and if we denote by $b_{\omega_k}(z) = \frac{z-\omega_k}{1-\overline{\omega_k}z}$ the Blaschke factors, for $k \in [\![1,n]\!]$, we can prove that $\dim\left(\mathcal{K}_{\Theta_n}\right) = n$ and, more precisely, we obtain an orthonormal basis of $\mathcal{K}_{\Theta_n}$ if we set:

$$\phi_1(z) = \frac{\sqrt{1-|\omega_1|^2}}{1-\overline{\omega_1}z} \quad \text{and} \quad \phi_k(z) = \left(\prod_{j=1}^{k-1} b_{\omega_j}\right)\frac{\sqrt{1-|\omega_k|^2}}{1-\overline{\omega_k}z} \quad k = 2,\ldots,n.$$

The basis $(\phi_1(z),\ldots,\phi_n(z))$ is called the Takenaka–Malmquist–Walsh orthonormal basis of $\mathcal{K}_{\Theta_n}$.

Writing S_{Θ_n} with respect to the Takenaka–Malmquist basis gives the matrix representation M_{Θ_n} with entries

$$\left[M_{\Theta_n}\right]_{i,j} = \begin{cases} \omega_j & \text{if } i{=}j \\ \prod_{k=i+1}^{j-1}(-\overline{\omega_k})\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_j|^2} & \text{if } i{<}j \\ 0 & \text{if } i{>}j \end{cases}.$$

Thus, with this notation, Theorem 0.16 can be restated as follows:

Theorem 0.17. *Let $n \in \mathbb{N}^*$, let $\omega_1,\ldots,\omega_n \in \mathbb{D}$ and let*

$$T_n = \begin{pmatrix} \omega_1 & \alpha_1^{(1)} & \alpha_1^{(2)} & \cdots & \cdots & \alpha_1^{(n-1)} \\ 0 & \omega_2 & \alpha_2^{(1)} & \alpha_2^{(2)} & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \alpha_{n-2}^{(2)} \\ \vdots & & \ddots & \ddots & \ddots & \alpha_{n-1}^{(1)} \\ 0 & \cdots & \cdots & \cdots & 0 & \omega_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{C}).$$

Assume that

$$\alpha_i^{(1)} = \sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+1}|^2}, \text{ for all } 1 \le i \le n-1.$$

Then, T_n is a contraction if and only if $T_n = M_{\Theta_n}$.

Schwarz-Pick type inequalities

In Chapter 5, we will use versions of the von Neumann inequality to derive Schwarz-Pick type inequalities. The key idea will be to apply the von Neumann inequality to a *well chosen* contractive matrix (or tuple of contractive matrices) T. Our inequalities will then follow from the fact that $f(T)$ remains a contraction, for every function f on the disc algebra such that $\|f\|_\infty \le 1$. In order to do this, we will use the explicit functional calculus from Chapter 3 together with the characterizations of contractive matrices from Chapter 4.

We will first focus on 3×3 matrices, and we will use Theorem 0.13 to provide an operator-theoretic proof of the *three points Schwarz-Pick lemma* obtained by Beardon and Minda [8]:

Theorem 0.18 (Beardon-Minda)**.** *Let $f \in \mathcal{H}(\mathbb{D},\mathbb{D})$ be an holomorphic function from $\mathbb{D}$ to itself, and let ω_1, ω_2,ω_3 be pairwise distinct points in $\mathbb{D}$. Then,*

$$\mathrm{d}\left(f^*(\omega_1,\omega_2), f^*(\omega_3,\omega_2)\right) \le \mathrm{d}(\omega_1,\omega_3), \tag{0.6}$$

where d denotes the hyperbolic distance of the open unit disc, and where $f^*(z,\omega)$ denotes the hyperbolic *divided difference of f at the points z and ω.*

Moreover, equality holds if and only if f is a Blaschke product of degree ≤ 2.

Similarly, we will develop an operator-theoretic proof of the following result proved by Yamashita in [65], which can be seen as an analog of Beardon-Minda's inequality in the case where the three points are coincident:

Theorem 0.19. *Let $f \in \mathcal{H}(\mathbb{D},\mathbb{D})$ and let $\Gamma(z,f) = \frac{(1-|z|^2)|f'(z)|}{1-|f(z)|^2}$. Then, for every $\omega \in \mathbb{D}$,*

$$\left|\frac{\partial\Gamma(\omega,f)}{\partial\omega}\right| \leq \frac{1-|\Gamma(\omega,f)|^2}{1-|\omega|^2}. \tag{0.7}$$

Moreover, equality holds if and only if f is a Blaschke product of degree ≤ 2.

This inequality can be rephrased in terms of *Peschl's invariant derivatives*. Let $f : \mathbb{D} \to \mathbb{D}$ be an holomorphic function, let $\omega \in \mathbb{D}$, and consider the mapping

$$g : z \in \mathbb{D} \mapsto \frac{f\left(\frac{z+\omega}{1+\overline{\omega}z}\right) - f(\omega)}{1-\overline{f(\omega)}f\left(\frac{z+\omega}{1+\overline{\omega}z}\right)} \in \mathbb{C}. \tag{0.8}$$

Then g is analytic on $\mathbb{D}$ and $g(0) = 0$. We have $g(z) = \sum_{n=1}^{\infty} \frac{D_n f(\omega)}{n!} z^n$, with $D_n f(z_0) := g^{(n)}(0)$. The quantities $D_n f(\omega)$ are called *Peschl's invariant derivatives* (see *e.g.* [37]). The first two values of Peschl's invariant derivatives are explicitely computed as:

$$\begin{aligned}
D_1 f(\omega) &= \frac{(1-|\omega|^2)f'(\omega)}{1-|f(\omega)|^2},\\
D_2 f(\omega) &= \frac{(1-|\omega|^2)^2}{1-|f(\omega)|^2}\left[f''(\omega) - \frac{2\overline{\omega}f'(\omega)}{1-|\omega|^2} + \frac{2\overline{f(\omega)}f'(\omega)^2}{1-|f(\omega)|^2}\right].
\end{aligned}$$

With these notations, the Schwarz-Pick inequality for derivatives (0.2) can be restated as $|D_1 f(\omega)| \leq 1$, while (0.7) can be written (see [12, Proposition 3.4]) as $|D_2 f(\omega)| \leq 2(1-|D_1 f(\omega)|^2)$.

Afterwards, we will examine n–points Schwarz-Pick type inequalities from an operator-theoretic perspective. In Section 5.2, we will explore the connection between the n points Schwarz-Pick inequality of Baribeau, Rivard, and Wegert [7] – which involves hyperbolic divided differences –, the Nevanlinna-Pick interpolation problem and the inequality:

$$\|f(M_{\Theta_n})\| \leq 1, \quad \text{for all holomorphic map } f : \mathbb{D} \to \mathbb{D}, \tag{0.9}$$

which is a consequence of the von Neumann inequality.

We will also establish the following *higher order* Schwarz-Pick inequality, extending the work of Wiener and Ruscheweyh [47, 56]:

Theorem 0.20. *Let f be a analytic function of $\mathbb{D}$ into $\overline{\mathbb{D}}$ with $f(z) = \sum_{n=0}^{\infty} a_n z^n$ for $z \in \mathbb{D}$. Then, for each $n \geq 1$ and each $k \geq 1$ we have:*

$$\left|a_{n+k}\left(1-|a_0|^2\right) + a_n a_k \overline{a_0}\right|^2 \leq \left[(1-|a_0|^2)^2 - |a_n|^2\right] \cdot \left[(1-|a_0|^2)^2 - |a_k|^2\right]. \tag{0.10}$$

In particular, this gives:

Corollary 0.21. *Let f be a analytic function of $\mathbb{D}$ into $\overline{\mathbb{D}}$ with $f(z) = \sum_{n=0}^{\infty} a_n z^n$ for $z \in \mathbb{D}$. Then*

$$1 - |a_0|^2 - |a_n| \geq \frac{\left| a_{2n}\left(1 - |a_0|^2\right) + a_n^2 \overline{a_0} \right|}{2(1 - |a_0|^2)}.$$

The last result is an improvement of Wiener's inequality:

$$|a_k| \leq 1 - |a_0|^2, \quad \forall\, k \geq 1.$$

Subsequently, in Section 5.4, we will focus on the polydisc, exploiting Drury's [20] and Knese's [39] results – which state (respectively) that the von Neumann inequality holds for commuting tuples of 2×2 and 3×3 matrices (see Theorem 1.2.11 and Theorem 1.2.12) – to derive several Schwarz-Pick type inequalities for the polydisc. First of all, we will provide an operator-theoretic proof of the following result:

Theorem 0.22 (Rudin, [54])**.**

(i) Let $f : \mathbb{D}^n \to \mathbb{D}$ be an holomorphic function, and let $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_n) \in \mathbb{D}^n$.

Then,

$$\left| \frac{f(a_1, \cdots, a_n) - f(b_1, \cdots, b_n)}{1 - \overline{f(a_1, \cdots, a_n)} f(b_1, \cdots, b_n)} \right| \leq \max_{1 \leq i \leq n} \left| \frac{a_i - b_i}{1 - \overline{a_i} b_i} \right|. \tag{0.11}$$

(ii) Let $f : \mathbb{D}^n \to \mathbb{D}$ be an holomorphic function, and let $a = (a_2, \ldots, a_n) \in \mathbb{D}^n$.

Then,

$$\sum_{i=1}^{n} (1 - |a_i|^2) \left| \frac{\partial f(a)}{\partial z_i} \right| \leq 1 - |f(a)|^2. \tag{0.12}$$

Then, we will establish in Section 5.4.2 the following generalization of Theorem 0.19 in several variables:

Theorem 0.23. *For $n \in \mathbb{N}^*$ let $w = (\omega_1, \ldots, \omega_n) \in \mathbb{D}^n$ and consider $f \in \mathcal{H}(\mathbb{D}^n, \mathbb{D})$.*

Then, we have:

$$|D_2 f(\omega)| \leq 2(1 - |D_1 f(\omega)|^2), \tag{0.13}$$

where $D_1 f$ and $D_2 f$ denote the Peschl invariant derivatives of f.

Furthermore, in the particular case of the bidisc, we use the notion of *distinguished varieties* – introduced by Agler and McCarthy in [1] – to obtained an enhanced version of Theorem 0.22.

A *distinguished variety* is a set of the form $V \cap \overline{\mathbb{D}}^2$, where V is an algebraic set in $\mathbb{C}^2$ with the property that $\overline{V} \cap \partial(\mathbb{D}^2) = \overline{V} \cap \mathbb{T}^2$.

Agler and McCarthy proved in [1] the following result:

Theorem 0.24 (Agler and McCarthy, [1])**.** *For any pair of commuting contractive matrices (T_1, T_2) without unimodular eigenvalues, there is a distinguished variety $V \cap \mathbb{D}^2$ such that the von-Neumann* *inequality holds on $V \cap \mathbb{D}^2$ for any polynomial p in $\mathbb{C}[X_1, X_2]$, i.e.*

$$\|p(T_1, T_2)\| \leq \sup\{|p(z_1, z_2)| : (z_1, z_2) \in V \cap \mathbb{D}^2\}.$$

This theorem is a sharpening of Ando's theorem, which states that the von Neumann inequality holds for any pair of commuting contractions acting on a Hilbert space.

Exploiting this result, we establish the following sharpening of Theorem 0.22 for the bidisc:

Theorem 0.25.

(i) Let (a_1, a_2) and (b_1, b_2) be two points in the bidisc $\mathbb{D}^2$. Then there is a distinguished variety $V \cap \mathbb{D}^2$ such that the Schwarz-Pick inequality

$$\left|\frac{f(a_1,a_2) - f(b_1,b_2)}{1 - \overline{f(a_1,a_2)} f(b_1,b_2)}\right| \leq \max\left\{\left|\frac{a_1 - b_1}{1 - \overline{a_1} b_1}\right|, \left|\frac{a_2 - b_2}{1 - \overline{a_2} b_2}\right|\right\} \tag{0.14}$$

holds for any function f which is holomorphic on the bidisc $\mathbb{D}^2$ and continuous on $\overline{\mathbb{D}}^2$ with

$$\sup\{|f(z_1, z_2)| \leq 1 : (z_1, z_2) \in V \cap \mathbb{D}^2\}.$$

(ii) Let (a_1, a_2) and (b_1, b_2) be two points in the bidisc $\mathbb{D}^2$. Then there is a distinguished variety $V \cap \mathbb{D}^2$ such that the Schwarz-Pick inequality holds for any function f which is holomorphic in the bidisc $\mathbb{D}^2$ and for which there is a sequence of positive real number (r_n) convergent to 1 with $r_n < 1$ such that

$$\sup\{|f(r_n z_1, r_n z_2)| \leq 1 : n \geq 1, (z_1, z_2) \in V \cap \mathbb{D}^2\}.$$

Last but not least, in Chapter 5, we will also explore operator versions of Schwarz-Pick type inequalities, extending the work of Fan [22] and, more recently, of Jocić [33], who established operator versions of the Schwarz lemma. We will thus derive the two following results from (respectively) Parrott's Theorem 0.12 and Theorem 0.15:

Theorem 0.26. *Let H_1, H_2 be two Hilbert spaces. Consider three contractions $W_1 \in \mathcal{B}(H_1)$, $W_2 \in \mathcal{B}(H_2)$ and $V \in \mathcal{B}(H_2, H_1)$. Assume that $\sigma(W_1) \cap \sigma(W_2) = \emptyset$, and that $f \in \mathcal{H}(\mathbb{D}, \mathbb{D})$ is holomorphic on an open neighborhood of $\sigma(W_1) \cup \sigma(W_2)$. We denote by $X = X_{W_1, W_2, V}$ the unique solution of the equation*

$$W_1 X - X W_2 = D_{W_1^*} V D_{W_2}.$$

Then, there exists a contraction $Y \in \mathcal{B}(H_2, H_1)$ such that:

$$f(W_1)X - X f(W_2) = D_{f(W_1)^*} Y D_{f(W_2)}.$$

Theorem 0.27. *Let H_1, H_2, H_3 be three Hilbert spaces. Consider three contractions $W_1 \in \mathcal{B}(H_1)$, $W_2 \in \mathcal{B}(H_2)$ and $W_3 \in \mathcal{B}(H_3)$. Let $V_1 \in \mathcal{B}(H_2, H_1)$, $V_2 \in \mathcal{B}(H_3, H_2)$, and $V_3 \in \mathcal{B}(H_3, H_1)$ be contractions. Assume that $\sigma(W_i) \cap \sigma(W_j) = \emptyset$, for all $1 \leq i < j \leq 3$, and that $f \in \mathcal{H}(\mathbb{D}, \mathbb{D})$ is holomorphic on an open neighborhood of $\sigma(W_1) \cup \sigma(W_2) \cup \sigma(W_3)$. Let X_1, X_2, X_3 be respectively the unique solution of*

Sylvester's equations:

$$W_1X_1 - X_1W_2 = D_{W_1^*}V_1D_{W_2},$$
$$W_2X_2 - X_2W_3 = D_{W_2^*}V_2D_{W_3},$$
$$\textit{and } W_1X_3 - X_3W_3 = B - W_1X_1X_2 + X_1W_2X_2,$$

where

$$B = \left[D_{W_1^*}(\mathrm{Id} - V_1V_1^*)D_{W_1^*}\right]^{1/2} V_3 \left[D_{W_3}(\mathrm{Id} - V_2^*V_2)D_{W_3}\right]^{1/2} - D_{W_1^*}V_1W_2^*V_2D_{W_3}.$$

Then, there exist three contractions $Y_1 \in \mathcal{B}(H_2, H_1), Y_2 \in \mathcal{B}(H_3, H_2), Y_3 \in \mathcal{B}(H_3, H_1)$ *such that:*

$$\begin{cases} f(W_1)X_1 - X_1f(W_2) = D_{f(W_1)^*}Y_1D_{f(W_2)}, & (0.15)\\ f(W_2)X_2 - X_2f(W_3) = D_{f(W_2)^*}Y_2D_{f(W_3)}, & (0.16)\\ f(W_1)X_3 - X_3f(W_3) = X_1f(W_2)X_2 - f(W_1)X_1X_2 & \\ + \left[D_{f(W_1)^*}(\mathrm{Id} - Y_1Y_1^*)D_{f(W_1)^*}\right]^{1/2} Y_3 \left[D_{f(W_3)}(\mathrm{Id} - Y_2^*Y_2)D_{f(W_3)}\right]^{1/2} & (0.17)\\ \qquad\qquad - D_{f(W_1)^*}Y_1f(W_2)^*Y_2D_{f(W_3)}. & \end{cases}$$

Observe that in the scalar case, the first theorem is equivalent to the Schwarz-Pick lemma (Theorem 0.2), and the second one is equivalent to the Beardon-Minda inequality (Theorem 0.18).

Spectral sets

In 1951, von Neumann introduced in [63] the notion of spectral set, in order to estimate the norm of functions of matrices in terms of the sup-norm of the function:

Definition 0.28. Let $T \in \mathcal{B}(H)$, let $S \subset \mathbb{C}$ be a subset of the complex plane containing $\sigma(T)$, and let $K > 0$ be a real constant.

We say that S is a K-*spectral set* for T if, for every rational function f whose poles are outside S, we have:

$$\|f(T)\| \le K \cdot \|f\|_S,$$

where $\|f\|_S = \sup\{|f(z)| : z \in S\}$.

If $K = 1$, we say that S is a *spectral set* for T.

For instance, the von Neumann inequality states that the closed unit disc is a spectral set for every contraction.

A similar notion of K-spectral set can be defined for commuting n-tuples of operators.

In the last chapter of this manuscript, we will have a quick look at two other K-spectral sets: the open unit ball of $\mathbb{C}^n$ and the annulus $\mathbb{A}_r := \{z \in \mathbb{C} : r < |z| < 1\}$, with $r \in [0, 1[$.

First of all, concerning the open unit ball $\mathbb{B}_n$ of $\mathbb{C}^n$, Hartz, Richter, and Shalit proved in [30] the following von Neumann type inequality:

Definition 0.29. Let H be an Hilbert space, and let $T = (T_1, \ldots, T_n)$ be a n-tuple of bounded linear operators on H. We say that T is a *row contraction* if $\|T\|^2 = \|\sum_{i=1}^n T_iT_i^*\| \le 1$.

Theorem 0.30 (Hartz-Richter-Shalit)**.** *For every integer $d \geq 1$, there exists a constant C_d such that, for all $n \in \mathbb{N}^*$, for every commuting row contractive n-tuple of $d \times d$ matrices $T = (T_1, \dots, T_n)$, and for every polynomial $p \in \mathbb{C}[X_1, \dots, X_n]$, we have:*

$$\|p(T)\| \leq C_d \cdot \sup_{z \in \mathbb{B}_n} |p(z)|.$$

Moreover, we can choose $C_2 = 1$.

We will establish an explicit criteria for a commuting row contractive n-tuple of 2×2 upper-triangular matrices to be row contractive and, then, exploiting this inequality, we can obtain the following Schwarz-Pick type lemma for the unit ball, which is particular case of *Rudin's Schwarz lemma for the unit ball* (take $m = 1$ in [55, Theorem 8.1.4]):

Theorem 0.31. *Let $f : \mathbb{B}_n \to \mathbb{D}$ be an holomorphic map, and let $\omega, z \in \mathbb{B}_n$.*

Then,

$$\frac{\left|1 - \overline{f(\omega)} f(z)\right|^2}{(1 - |f(z)|^2)(1 - |f(\omega)|^2)} \leq \frac{|1 - \langle z, \omega \rangle|^2}{(1 - \|z\|^2)(1 - \|\omega\|^2)}.$$

Finally, concerning the annulus, we will focus on the work of Tsikalas [61], who introduced the class of bounded linear operators:

$$\mathcal{F}_r := \left\{T \in \mathcal{B}(H) : r^2 T^{-1} \left(T^{-1}\right)^* + TT^* \leq r^2 + 1, \sigma(T) \subset \mathbb{A}_r\right\},$$

where H denotes a Hilbert space and $r \in [0, 1[$.

He proved that, for all $r \in [0, 1[$, $\mathbb{A}_r]$ is a $\sqrt{2}$-spectral set for the elements of $\mathcal{F}_r$:

Theorem 0.32 (Tsikalas)**.** *For every bounded holomorphic function $\phi \in H^\infty(\mathbb{A}_r)$, for every $T \in \mathcal{F}_r$,*

$$\|\phi(T)\| \leq \sqrt{2} \sup\{|\phi(z)| : z \in \mathbb{A}_r\},$$

where the constant $\sqrt{2}$ is the best possible.

In the end of this manuscript, we will establish the following characterization of the class $\mathcal{F}_r$ for 2×2 upper-triangular matrix:

Proposition 0.33. *Let $T = \begin{pmatrix} \omega_1 & \alpha \\ 0 & \omega_2 \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$, with $\omega_1, \omega_2 \in \mathbb{A}_r$.*

$T \in \mathcal{F}_r$ if and only if

$$|\alpha|^2 \leq \frac{(1 - |\omega_1|^2)(1 - |\omega_2|^2)(|\omega_1|^2 - r^2)(|\omega_2|^2 - r^2)}{|r^2 - \overline{\omega_2}\omega_1|^2 + r^2(1 - |\omega_1|^2)(1 - |\omega_2|^2)}.$$

Chapter 1

The von Neumann inequality

In this chapter, we recall some fundamental concepts in operator theory.

In all this chapter, X denotes a (complex) Banach space and H denotes a (complex) Hilbert space.

1.1 Basics of spectral theory

We state here – without proof – some very classical results in spectral theory. The interested reader can refer to [35, Chap. 3–4], [14, Chap. VII] or [21, Chap. VII] for a more complete overview.

1.1.1 General definitions and properties

We start by recalling the definition of the spectrum of an operator:

Definition 1.1.1. Let $T \in \mathcal{B}(X)$.

(i) The *spectrum* of T is the set $\sigma(T) = \{\lambda \in \mathbb{C} : T - \lambda\mathrm{Id} \text{ is not invertible}\}$;

(ii) The *spectral radius* is the quantity $r(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}$;

(iii) The *resolvent* of T is the map $R(\,\cdot\,, T) : \mathbb{C}\backslash\sigma(T) \to \mathcal{B}(X)$ defined by $R(z, T) = (T - z\mathrm{Id})^{-1}$.

Remark 1.1.2. If H is of finite dimension, $\sigma(T)$ is just the set of all the eigenvalues of T.

Remark 1.1.3. If $\mathcal{A}$ is a unital subalgebra of $\mathcal{B}(X)$, we denote by

$$\sigma_{\mathcal{A}}(T) = \{\lambda \in \mathbb{C} : T - \lambda\mathrm{Id} \text{ is not invertible in } \mathcal{A}\}.$$

We have $\sigma(T) \subset \sigma_{\mathcal{A}}(T)$.

The two following properties can be easily deduced from the definition:

Proposition 1.1.4. *With the above notation:*

(i) $r(T) \leq \|T\|$;

(ii) *$\sigma(T)$ is a compact subset of $\mathbb{C}$.*

With a little more work, one can prove the following important result giving the spectrum of $p(T)$, where p is a polynomial:

Proposition 1.1.5. *Let $T \in \mathcal{B}(X)$ and let $p \in \mathbb{C}[X]$.*

Then,
$$\sigma(p(T)) = p(\sigma(T)) = \{p(\lambda) : \lambda \in \sigma(T)\}.$$

The resolvent has the following properties:

Lemma 1.1.6. (i) *Resolvent's identity :*

$$\forall\, \lambda, \mu \in \mathbb{C},\ R(\lambda, T) - R(\mu, T) = (\lambda - \mu) R(\lambda, T) R(\mu, T);$$

(ii) *The resolvent is holomorphic on $\mathbb{C} \backslash \sigma(T)$ and satisfies $\lim_{|z| \to +\infty} R(z, T) = 0$;*

(iii) *For all $\lambda \in \mathbb{C}$ such that $|\lambda| > r(T)$, $R(\lambda, T) = -\sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}$. Moreover, the series converges in the uniform operator topology.*

Those properties of the resolvent enable us to prove the following proposition:

Proposition 1.1.7. (i) *$\sigma(T)$ is non empty;*

(ii) *$r(T) = \lim_{n \to \infty} \|T^n\|^{1/n} = \inf_{n \geq 1} \|T^n\|^{1/n}$ (spectral radius formula).*

1.1.2 Spectral theory in an Hilbertian framework

The last identity of Proposition 1.1.7 has the following crucial consequence when we work in an Hilbertian framework:

Corollary 1.1.8. *Let $T \in \mathcal{B}(H)$.*

(i) *If T is normal, then, $r(T) = \|T\|$;*

(ii) *In the general case, we have $\|T\|^2 = r(T^*T)$.*

We have also the following useful properties:

Proposition 1.1.9. *Let $T \in \mathcal{B}(H)$.*

(i) *If T is self-adjoint, then, $\sigma(T) \subset \mathbb{R}$;*

(ii) *If T is positive, then, $\sigma(T) \subset \mathbb{R}_+$;*

(iii) *If T is unitary, then, $\sigma(T) \subset \mathbb{T}$.*

Finally, one of the main tool that we will use in this manuscript is the existence of functional calculus for bounded linear operators acting on Hilbert spaces. We start with the *continuous functional calculus for self-adjoint operators*:

Theorem 1.1.10. *Let $T \in \mathcal{B}(H)$ be a self-adjoint operator. There exists a unique continuous homomorphism $\Psi_T : \mathcal{C}(\sigma(T)) \to \mathcal{B}(H)$ such that $\Psi_T(p) = p(T)$, for every polynomial $p \in \mathbb{C}[X]$. Moreover, if we denote $f(T) = \Psi_T(f)$, for $f \in \mathcal{C}(\sigma(T))$, we have the following properties :*

(i) *There exists a sequence of polynomials $(p_n)_{n \in \mathbb{N}}$ such that $\lim_{n \to \infty} p_n(T) = f(T)$;*

(ii) *$\|f(T)\| = \|f\|_{\infty, \sigma(T)} := \sup\{|f(t)| : t \in \sigma(T)\}$;*

(iii) $f(T)^* = \overline{f}(T)$;

(iv) $f(T)A = Af(T)$ *for every* $A \in \mathcal{B}(H)$ *such that* $TA = AT$;

(v) $\sigma(f(T)) = f(\sigma(T))$;

(vi) *If* $f \in \mathcal{C}(\sigma(T), \mathbb{R})$ *and* $g \in \mathcal{C}(f(\sigma(T)))$, $(g \circ f)(T) = g(f(T))$.

Ψ_T *is called the* continuous functional calculus *associated to* T.

A noteworthy application of the continuous functional calculus is the following:

Definition 1.1.11. An operator $T \in \mathcal{B}(H)$ is called a *contraction* if $\|T\| \leq 1$. If $\|T\| < 1$, we say that T is a *strict contraction*.

Proposition 1.1.12 (Defect operator)**.** *Let* $T \in \mathcal{B}(H)$ *be a contraction.*

(i) There exists a unique positive operator $D_T := (\mathrm{Id} - T^*T)^{1/2}$ *such that* $D_T^2 = \mathrm{Id} - T^*T$.

This operator is called defect operator *associated to* T.

Moreover:

(ii) There exists a sequence of polynomials $(p_n)_{n\in\mathbb{N}}$ *such that* $\lim_{n\to\infty} p_n(T^*T) = D_T$;

(iii) The defect operator satisfies the equation $TD_T = D_{T^*}T$.

In the more general case where the operator T is not necessarily self-adjoint, one can define an *holomorphic functional calculus*, also known as *Riesz-Dunford's functional calculus*:

Definition 1.1.13. Let $T \in \mathcal{B}(H)$ and let $\mathcal{H}(T)$ be the set of functions that are holomorphic on some neighborhood of $\sigma(T)$. Let $f \in \mathcal{H}(T)$ and let $\mathcal{U}$ be an open set containing $\sigma(T)$ whose boundary $\partial\mathcal{U}$ consists of a finite number of rectifiable Jordan curves oriented in the positive sens, and such that f is holomorphic on $\overline{\mathcal{U}}$.

Then, the operator $f(T)$ is defined by the equation :

$$f(T) = \frac{1}{2i\pi}\int_{\partial\mathcal{U}} f(\xi)(\xi\mathrm{Id} - T)^{-1}d\xi.$$

The holomorphic functional calculus satisfies the following properties:

Theorem 1.1.14. *Let* $T \in \mathcal{B}(H)$.

(i) The mapping $f \mapsto f(T)$ *is a homomorphism of algebra from* $\mathcal{H}(T)$ *to* $\mathcal{B}(H)$;

(ii) If $f \in \mathcal{H}(T)$ *has the power series expansion* $f(z) = \sum_{k=0}^{\infty} a_k z^k$, *valid in a neighborhood of* $\sigma(T)$, *then,* $f(T) = \sum_{k=0}^{\infty} a_k T^k$;

(iii) If $f \in \mathcal{H}(T)$, *then,* $f \in \mathcal{H}(T^*)$ *and* $f(T^*) = f(T)^*$;

(iv) If $f \in \mathcal{H}(T)$, *then,* $\sigma(f(T)) = f(\sigma(T))$;

(v) If $f \in \mathcal{H}(T)$ *and* $g \in \mathcal{H}(f(\sigma(T)))$, *then,* $(g \circ f)(T) = g(f(T))$.

1.1.3 Gelfand representation

The Gelfand representation theory relies on the following lemma :

Lemma 1.1.15 (Gelfand-Mazur)**.** *Let $\mathcal{A}$ be an unital Banach algebra such that every non zero element is invertible. We denote by* 1 *its unital element.*
Then, $\mathcal{A}$ is isometrically isomorphic to $\mathbb{C}$.

In the following, we place ourselves within the framework of a unital commutative Banach algebra $\mathcal{C}$. For instance, $\mathcal{C}$ can be a commutative subalgebra of $\mathcal{B}(X)$, where X is a Banach space.

A unital commutative Banach algebra is, in particular, a unital commutative ring and, as such, admits (maximal) ideals. Given a closed ideal $I \subsetneq \mathcal{C}$, we equip the vector space $\mathcal{C}/I$ with the usual quotient norm $\|\overline{x}\| = \inf_{y \in \overline{x}} \|y\|$. Together with this norm, $\mathcal{C}/I$ is a Banach space.

Lemma 1.1.16. *Let M be a maximal ideal of $\mathcal{C}$.*
Then, M is closed in $\mathcal{C}$ and $\mathcal{C}/M \simeq \mathbb{C}$.

Now, we introduce the key notion of this paragraph:

Definition 1.1.17. A non-zero algebra homomorphism $\chi : \mathcal{C} \to \mathbb{C}$ is called a *character* of $\mathcal{C}$.
The set of all characters of $\mathcal{C}$ is called the *Gelfand spectrum* of $\mathcal{C}$, and is denoted by $\widehat{\mathcal{C}}$.

Proposition 1.1.18. *Let $\mathcal{M}_{\mathcal{C}}$ be the set of maximal ideals of $\mathcal{C}$.*
The map $\Phi : \chi \in \widehat{\mathcal{C}} \mapsto \ker \chi \in \mathcal{M}_{\mathcal{C}}$ is a bijection.

This allows us to state the following crucial theorem, which establishes a connection between the spectrum of an element $x \in \mathcal{C}$ and the Gelfand spectrum of $\mathcal{C}$:

Theorem 1.1.19. *Let $\chi \in \widehat{\mathcal{C}}$ and let $x \in \mathcal{C}$.*

(i) χ is continuous on $\mathcal{C}$ and $\|\chi\| = 1$;

(ii) $\sigma_{\mathcal{C}}(x) = \left\{\chi(x) : \chi \in \widehat{\mathcal{C}}\right\}$;

(iii) x is invertible in $\mathcal{C}$ if and only if, for all $\chi \in \widehat{\mathcal{C}}$, $\chi(x) \neq 0$.

1.1.4 Joint spectrum

Definition 1.1.20. Let $T = (T_1, \ldots, T_n)$ be a n-tuple of pairwise commuting operators acting on a Banach space X. The *(algebraic) joint spectrum* of T is the set

$$\sigma(T) = \left\{(\chi(T_1), \ldots, \chi(T_n)) : \chi \in \widehat{\mathcal{C}}\right\},$$

where $\mathcal{C}$ is the commutative subalgebra of $\mathcal{B}(X)$ generated by $T_1, \ldots, T_n$.

In this manuscript, the notion of joint spectrum will only appear in the finite-dimensional case. In this context, the notion joint spectrum is linked with the notion of joint triangularization (see [13]).

Recall that if $T = (T_1, \ldots, T_n)$ is a n-tuple of $d \times d$ pairwise commuting matrices with complex coefficients, there exists a unitary matrix U such that, for every $i \in [\![1, n]\!]$, $U^* T_i U$ is upper-triangular with diagonal entries $\lambda_1^{(i)}, \ldots, \lambda_d^{(i)}$, where $\left\{\lambda_1^{(i)}, \ldots, \lambda_d^{(i)}\right\}$ denotes the spectrum of T_i (perhaps with repetitions).

Proposition 1.1.21. *Let $T = (T_1, \ldots, T_n)$ be a n-tuple of $d \times d$ pairwise commuting matrices over $\mathbb{C}$, and let $\left\{\lambda_1^{(i)}, \ldots, \lambda_d^{(i)}\right\}$ be the spectrum of T_i (perhaps with repetitions), for all $i \in [\![1, n]\!]$.*

Then, we have:

$$\sigma(T) = \left\{ \left(\lambda_k^{(1)}, \ldots, \lambda_k^{(n)} \right) : k = 1, \ldots, d \right\}.$$

Remark 1.1.22. For a n-tuple of pairwise commuting operators $T = (T_1, \ldots, T_n)$ acting on a Banach space, the *accurate* notion of spectrum is the Taylor spectrum $\sigma_{\text{Tay}}(T)$, whose construction is more elaborated. For further details on the Taylor spectrum, see [41, 60]. In general, we have the inclusion $\sigma_{\text{Tay}}(T) \subset \sigma(T)$. However, in the finite-dimensional case, it can be shown (see [13]) that $\sigma_{\text{Tay}}(T) = \sigma(T)$.

1.2 The von Neumann inequality

1.2.1 von Neumann's inequality in one variable

From now, for $f \in \mathcal{C}\left(\overline{\mathbb{D}}\right)$, we will denote

$$\|f\|_\infty = \|f\|_{\infty, \overline{\mathbb{D}}} = \sup\{|f(z)| : |z| \leq 1\}.$$

This work is mainly based on the following crucial inequality, obtained by von Neumann in 1951:

Theorem 1.2.1 (von Neumann [63])**.** *Let $T \in \mathcal{B}(H)$ be a contraction and let $p \in \mathbb{C}[X]$ be a polynomial. Then,*

$$\|p(T)\| \leq \|p\|_\infty.$$

Remark 1.2.2. If we set $p(z) = \sum_{k=0}^n a_k z^k$, this inequality is a refinement of the rough majorization:

$$\|p(T)\| = \left\| \sum_{k=0}^n a_k T^k \right\| \leq \sum_{k=0}^n |a_k| \cdot \|T\|^k \leq \sum_{k=0}^n |a_k| =: \|p\|_1.$$

Indeed, for all $z \in \overline{\mathbb{D}}$, $|p(z)| \leq \sum_{k=0}^n |a_k| \cdot |z|^k \leq \|p\|_1$, which implies that $\|p\|_\infty \leq \|p\|_1$.

Remark 1.2.3. Although we have $\|\cdot\|_\infty \leq \|\cdot\|_1$, the norms $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are not equivalent on $\mathbb{C}[X]$. To see this, let us consider the sequence $(p_n)_{n \in \mathbb{N}^*}$, where $p_n(z) = \sum_{k=1}^n \frac{1}{k}\left(z^{n+k} - z^{n-k}\right)$, for all $n \geq 1$. On the one hand, it is easy to see that $\lim_{n \to \infty} \|p_n\|_1 = +\infty$. On the other hand, for $z = e^{i\theta} \in \mathbb{T}$, $\left|p\left(e^{i\theta}\right)\right| = \left|\sum_{k=1}^n \frac{\sin(k\theta)}{k}\right|$. Using some results on convergence of Fourier series, we can prove (see *e.g.* [64, Chap. 2, Ex. 7.26]) that, for all $n \in \mathbb{N}^*$, for all $\theta \in \mathbb{R}$, $\left|\sum_{k=1}^n \frac{\sin(k\theta)}{k}\right| \leq \frac{\pi}{2} + 1$. It follows that, for all $n \in \mathbb{N}^*$, $\|p_n\|_\infty \leq \frac{\pi}{2} + 1$.

There exist several proofs of this famous inequality. One of the most usual proofs uses Sz.-Nagy's dilation theorem, which enables to restrict oneself to the case where T is unitary:

Theorem 1.2.4 (Sz.-Nagy)**.** *Let T be a contraction operator on a Hilbert space H. Then, there is a Hilbert space K containing H as a subspace and a unitary operator U on K such that, for all $n \in \mathbb{N}$:*

$$T^n = P_H U^n_{|H},$$

where $P_H : K \to H$ is the orthogonal projection onto H.

We say that U is an unitary dilation *of T.*

The interested reader can, for instance, refer to [50, Chap. 1] or to [46, Chap. 1] in order to obtain more information about those two results.

Remark 1.2.5. We can also prove (see *e.g.* [50, Th. 1.9]) that the von Neumann inequality actually characterizes Hilbert spaces in the following sense: if X is a complex Banach space such that, for all $T \in \mathcal{B}(X)$ with $\|T\| = 1$, we have $\|p(T)\| \leq \|p\|_\infty$ for all polynomials p, then, X is isometrically isomorphic to a Hilbert space.

A fundamental application of the von Neumann inequality is the existence of functional calculus for contractions on the disc algebra:

Theorem 1.2.6. *Let $T \in \mathcal{B}(H)$ be a contraction. There exists a continuous homomorphism $\Psi_T : \mathcal{A}\left(\overline{\mathbb{D}}\right) \to \mathcal{B}(H)$ such that $\Psi_T(p) = p(T)$, for every polynomial $p \in \mathbb{C}[X]$. Moreover, if we denote $f(T) = \Psi_T(f)$, for $f \in \mathcal{A}\left(\overline{\mathbb{D}}\right)$, we have the following properties:*

(i) There exists a sequence of polynomials $(p_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} p_n(T) = f(T)$;

(ii) $\|f(T)\| \leq \|f\|_\infty$.

1.2.2 von Neumann's inequality in several variables

The von Neumann inequality can be generalized to polynomials of two complex variables: indeed, for pairs of commuting contractions, there is an analogue of Sz.-Nagy's dilation theorem due to Ando ([5]) and, consequently, a two-variable analogue of von Neumann's inequality (see [50, Chap. 1] or [46, Chap. 5]):

Theorem 1.2.7 (Ando)**.** *Let $T_1, T_2 \in \mathcal{B}(H)$ be two commuting contractions. Then, there exist a Hilbert space K containing H as a subspace, and two commuting unitaries U_1, U_2 on K such that, for all $n, m \in \mathbb{N}$,*

$$T_1^n T_2^m = P_H U_1^n U_2^m{}_{|H},$$

where $P_H : K \to H$ is the orthogonal projection onto H.

Corollary 1.2.8. *Let $T_1, T_2 \in \mathcal{B}(H)$ be two commuting contractions, and let $p \in \mathbb{C}[X_1, X_2]$ be a polynomial in two variables. Then,*

$$\begin{aligned}\|p(T_1, T_2)\| &\leq \sup\{|p(z_1, z_2)| : |z_1| \leq 1, |z_2| \leq 1\} \\ &= \max\{|p(z_1, z_2)| : |z_1| = |z_2| = 1\}\end{aligned}$$

This can also be generalized to n-tuples of commuting isometries (see [46, Th. 5.1]):

Theorem 1.2.9. *Let $n \in \mathbb{N}^*$ and let $V_1, \cdots, V_n$ be n commuting isometries on a Hilbert space H. Then, there exists a Hilbert space K containing H and a n-tuple of commuting unitaries $(U_1, \cdots, U_n)$ on K such that, for all $m_1, \cdots m_n \in \mathbb{N}$:*

$$V_1^{m_1} \cdots V_n^{m_n} = P_H U_1^{m_1} \cdots U_n^{m_n}{}_{|H},$$

where $P_H : K \to H$ is the orthogonal projection onto H.

Corollary 1.2.10. *Let $n \in \mathbb{N}^*$, let $V_1, \cdots, V_n$ be n commuting isometries on a Hilbert space H and let*

$p \in \mathbb{C}[X_1, \cdots, X_n]$ be a polynomial in n variables. Then,

$$\begin{aligned}\|p(T_1, \cdots, T_n)\| &\leq \sup\{|p(z_1, \cdots, z_n)| : |z_1| \leq 1, \cdots, |z_n| \leq 1\} \\ &= \max\{|p(z_1, \cdots, z_n)| : |z_1| = \cdots = |z_n| = 1\}\end{aligned}$$

More surprisingly, the analogue of von Neumann's inequality fails for three or more commuting contractions. There are several counterexamples in the literature: the two most known are probably due to Crabb and Davie, who found in [15] an example of three commuting contractions acting on a 8-dimensional Hilbert space who do not satisfy the von Neumann inequality, and to Varopoulos, who found in [62] another counterexample involving three commuting contractions acting on a 5-dimensional Hilbert space. More recently, Holbrook found in [32] a 3-tuple of 4×4 matrices which fails the von Neumann inequality.

Of course, this implies that Ando's dilation theorem is not valid for three (or more) mutually commuting contractions. Nevertheless, Parrott presented in [45] an example of three mutually commuting contractions T_1, T_2, T_3 which do not dilate to three commuting unitaries, but which satisfy the von Neumann inequality, *i.e.* such that, for every polynomial $p \in \mathbb{C}[X]$, we have:

$$\|p(T_1, T_2, T_3)\| \leq \sup\{|p(z_1, z_2, z_3)| : |z_1| = \cdots = |z_3| = 1\}.$$

This shows that the results on dilations of commuting contractions are stronger than the von Neumann inequalities.

However, the von Neumann inequality still holds for n-tuples of 2×2 or 3×3 contractive matrices:

Theorem 1.2.11 (Drury [20] ; Fu and Russo [25])**.** *Let $T_1, ..., T_n$ be mutually commuting 2×2 contractive matrices, and let $p \in \mathbb{C}[X_1, \cdots, X_n]$ be a polynomial in n variables. Then, we have:*

$$\|p(T_1, ..., T_n)\| \leq \sup\{|p(z_1, \cdots z_n)| : |z_1| = \cdots = |z_n| = 1\}.$$

Theorem 1.2.12 (Knese [39])**.** *Let $T_1, ..., T_n$ be mutually commuting 3×3 contractive matrices, and let $p \in \mathbb{C}[X_1, \cdots, X_n]$ be a polynomial in n variables. Then, we have:*

$$\|p(T_1, ..., T_n)\| \leq \sup\{|p(z_1, \cdots z_n)| : |z_1| = \cdots = |z_n| = 1\}.$$

1.2.3 Link between the von Neumann inequality and the hyperbolic geometry of the unit disc

Another application of von Neumann's inequality, which can be considered as the starting point of this work, is the fact that von Neumann's inequality provides an alternative proof of Schwarz-Pick's lemma:

Theorem 1.2.13 (Schwarz-Pick)**.**

(i) Let $f \in \mathcal{H}(\mathbb{D}, \mathbb{D})$ be an holomorphic function mapping $\mathbb{D}$ to $\mathbb{D}$ and let $\omega_1, \omega_2 \in \mathbb{D}$ be two distinct points of the unit disc. Then, we have:

$$\left| \frac{f(\omega_1) - f(\omega_2)}{1 - \overline{f(\omega_1)} f(\omega_2)} \right| \leq \left| \frac{\omega_1 - \omega_2}{1 - \overline{\omega_1} \omega_2} \right| \tag{1.1}$$

If we denote by $\rho(\omega_1, \omega_2) := \left| \frac{\omega_1 - \omega_2}{1 - \overline{\omega_1} \omega_2} \right|$ the pseudo-hyperbolic distance on the unit disc, this *inequality becomes:*

$$\rho(f(\omega_1), f(\omega_2)) \leq \rho(\omega_1, \omega_2).$$

In other words, every holomorphic function mapping $\mathbb{D}$ to $\mathbb{D}$ is contractive with respect to the pseudo-hyperbolic metric.

(ii) *If $\omega_1 = \omega_2 =: \omega \in \mathbb{D}$, this inequality becomes:*

$$|f'(\omega)| \leq \frac{1-|f(\omega)|^2}{1-|\omega|^2} \tag{1.2}$$

The fact that von Neumann's inequality provides an alternative proof of Schwarz-Pick's lemma is mentioned in [46, Ex. 2.17–2.18] or in [52, p. 17] but, as this observation plays a crucial role in this work, we write down the details in this manuscript.

The idea is to apply von Neumann's inequality to a *well chosen* 2×2 matrix. In order to do this, we start with the following lemmas:

Lemma 1.2.14. *Let $T = \begin{pmatrix} a & c \\ 0 & b \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$, and let $N = |a|^2 + |b|^2 + |c|^2$, $\delta = ab$.*

Then,

(i) $\|T\|^2 = \frac{1}{2}\left(N + \sqrt{N^2 - 4|\delta|^2}\right)$;

(ii) $\|T\| \leq 1$ *if and only if* $|c|^2 \leq (1-|a|^2)(1-|b|^2)$.

Proof.

(i) We will use the identity $\|T\|^2 = \|T^*T\| = r(T^*T)$.

If we denote by χ_{T^*T} the characteristic polynomial of T^*T, a straightforward computation gives:

$$T^*T = \begin{pmatrix} |a|^2 & \overline{a}c \\ a\overline{c} & |c|^2 + |b|^2 \end{pmatrix}$$

and

$$\chi_{T^*T}(\lambda) = \lambda^2 - N\lambda + |\delta|^2.$$

The discriminant of this polynomial is:

$$\begin{aligned} \Delta &= N^2 - 4|\delta|^2 \\ &= \left[|a|^2 + |b|^2 + |c|^2 + 2|ab|\right]\left[(|a| - |b|)^2 + |c|^2\right] \end{aligned}$$

Except in the trivial case where $T = a \cdot \mathrm{Id}$, we have $\Delta > 0$ and, therefore, T^*T has two distinct real eigenvalues equal to $\frac{1}{2}\left(N \pm \sqrt{N^2 - 4|\delta|^2}\right)$. Hence, $\|T\|^2 = r(T^*T) = \frac{1}{2}\left(N + \sqrt{N^2 - 4|\delta|^2}\right)$.

(ii) This is a straightforward consequence of the preceding item.

□

Lemma 1.2.15. *Let $T = \begin{pmatrix} a & c \\ 0 & b \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$, and let $p \in \mathbb{C}[X]$.*

(i) *If* $a \neq b$, *then,* $p(T) = \begin{pmatrix} p(a) & c\left(\frac{p(a)-p(b)}{a-b}\right) \\ 0 & p(b) \end{pmatrix}$;

(ii) *If* $a = b$, *then,* $p(T) = \begin{pmatrix} p(a) & cp'(a) \\ 0 & p(a) \end{pmatrix}$.

Proof.

(i) We can prove, by induction on n, that for all $n \in \mathbb{N}$,

$$T^n = \begin{pmatrix} a^n & c\left(\frac{a^n-b^n}{a-b}\right) \\ 0 & b^n \end{pmatrix}.$$

Thus, we obtain that, for every polynomial $p \in \mathbb{C}[X]$,

$$p(T) = \begin{pmatrix} p(a) & c\left(\frac{p(a)-p(b)}{a-b}\right) \\ 0 & p(b) \end{pmatrix}.$$

(ii) We proceed as in the proof of the previous item.

□

Remark 1.2.16. If $a \neq b \in \overline{\mathbb{D}}$, then, we can use the density of polynomials in the disc algebra $\mathcal{A}\left(\overline{\mathbb{D}}\right)$ to get:

$$f(T) = \begin{pmatrix} f(a) & c\left(\frac{f(a)-f(b)}{a-b}\right) \\ 0 & f(b) \end{pmatrix}, \text{ for all functions } f \in \mathcal{A}\left(\overline{\mathbb{D}}\right).$$

Similarly, if $a = b \in \mathbb{D}$, we get:

$$f(T) = \begin{pmatrix} f(a) & cf'(a) \\ 0 & f(a) \end{pmatrix}, \text{ for all functions } f \in \mathcal{A}\left(\overline{\mathbb{D}}\right).$$

We will also need the following key identity, that will play a crucial role throughout this manuscript:

Lemma 1.2.17. *For all* $u, v \in \mathbb{C}$, *we have:*

$$|1 - \overline{u}v|^2 = \left(1 - |u|^2\right)\left(1 - |v|^2\right) + |u - v|^2 \tag{1.3}$$

Now, we are ready for the proof of Theorem 1.2.13:

Proof of Theorem 1.2.13.

We only deal with the case where $\omega_1 \neq \omega_2$ (the case where $\omega_1 = \omega_2$ is similar). If $f(\omega_1) = f(\omega_2)$, the result is obvious. In the following, we will thus assume that $f(\omega_1) \neq f(\omega_2)$. Let $\omega_1, \omega_2 \in \mathbb{D}$ and let $T = \begin{pmatrix} \omega_1 & \alpha \\ 0 & \omega_2 \end{pmatrix}$, with $\alpha = \sqrt{1 - |\omega_1|^2}\sqrt{1 - |\omega_2|^2}$. By Lemma 1.2.14, T is a contraction. Now, let $f \in \mathcal{A}\left(\overline{\mathbb{D}}\right)$.

By Lemma 1.2.15, $f(T) = \begin{pmatrix} f(\omega_1) & \alpha\left(\frac{f(\omega_1)-f(\omega_2)}{\omega_1-\omega_2}\right) \\ 0 & f(\omega_2) \end{pmatrix}$ and, by von Neumann's inequality, $f(T)$ is a contraction too. Thus, by Lemma 1.2.14, we have:

$$\left(1-|\omega_1|^2\right)\left(1-|\omega_2|^2\right)\left|\frac{f(\omega_1)-f(\omega_2)}{\omega_1-\omega_2}\right|^2 \leq \left(1-|f(\omega_1)|^2\right)\left(1-|f(\omega_2)|^2\right)$$

which is equivalent to:

$$\frac{\left(1-|\omega_1|^2\right)\left(1-|\omega_2|^2\right)}{|\omega_1-\omega_2|^2} \leq \frac{\left(1-|f(\omega_1)|^2\right)\left(1-|f(\omega_2)|^2\right)}{|f(\omega_1)-f(\omega_2)|^2} \tag{1.4}$$

Then, by (1.3), (1.4) is equivalent to

$$\left|\frac{1-\overline{\omega_1}\omega_2}{\omega_1-\omega_2}\right|^2 - 1 \leq \left|\frac{1-\overline{f(\omega_1)}f(\omega_2)}{f(\omega_1)-f(\omega_2)}\right|^2 - 1,$$

which is equivalent to Schwarz-Pick's inequality. Finally, for an holomorphic function $f \in \mathcal{H}(\mathbb{D},\mathbb{D})$, we define the net $(f_r)_{0<r<1}$, where $f_r(z) = f(rz)$, for all $z \in \overline{\mathbb{D}}$. For all $r \in]0,1[$, $f_r \in \mathcal{A}\left(\overline{\mathbb{D}}\right)$ and, thus, we have:

$$\left|\frac{f(r\omega_1)-f(r\omega_2)}{1-\overline{f(r\omega_1)}f(r\omega_2)}\right| \leq \left|\frac{\omega_1-\omega_2}{1-\overline{\omega_1}\omega_2}\right|.$$

Schwarz-Pick's lemma is then obtained by making $r \to 1^-$ in this inequality. □

As it will be the common thread of this work, let us sum up quickly the scheme of the proof: first of all, we apply the von Neumann inequality to a *well chosen* contractive matrix T (we will see that this matrix is in fact the *model matrix* of size 2 with diagonal entries ω_1, ω_2). Then, we use the explicit formula for $f(T)$, where f is a function of the disc algebra, together with the characterization of contractive 2×2 matrices to obtain an inequality that turns out to be, after some transformations, equivalent to the Schwarz-Pick lemma. Once we get the inequality for the functions of the disc algebra, we conclude with a classical approximation argument.

Chapter 2

Hardy spaces and model operators

In this chapter, we recall – without proof – some elementary facts about Hardy spaces and model operators. For general notions on Hardy spaces and model spaces, we refer to [27, 28, 26, 2, 24] for a more complete overview about the topic.

2.1 Hardy spaces.

2.1.1 Hardy-Hilbert space on the open unit disc.

In this section, if f is an holomorphic function on $\mathbb{D}$, we denote by $(a_n(f))_{n\in\mathbb{N}}$ its Taylor coefficients, *i.e.* $f(z) = \sum_{n=0}^{\infty} a_n(f) z^n$.

Definition 2.1.1. The Hardy-Hilbert space $H^2 = H^2(\mathbb{D})$ is the set of holomorphic functions on $\mathbb{D}$ whose Taylor coefficients are square summable, *i.e.*

$$H^2 = H^2(\mathbb{D}) := \left\{ f \in \mathcal{H}(\mathbb{D}) : \sum_{n=0}^{\infty} |a_n(f)|^2 < \infty \right\}.$$

For $f, g \in H^2(\mathbb{D})$, we define the inner product by:

$$\langle f, g \rangle_{H^2} = \sum_{n=0}^{\infty} a_n(f) \overline{a_n(g)},$$

and the norm $\|\cdot\|_{H^2}$ by:

$$\|f\|_{H^2}^2 := \langle f, f \rangle_{H^2} = \sum_{n=0}^{\infty} |a_n(f)|^2.$$

Proposition 2.1.2. *$H^2(\mathbb{D})$ is an Hilbert space which is isometrically isomorphic to $\ell^2(\mathbb{N})$. Moreover, if we denote $e_n : z \mapsto z^n$, for all $n \in \mathbb{N}$, then, $(e_n)_{n\in\mathbb{N}}$ is an orthonormal basis of H^2.*

An important property of the Hardy-Hilbert is the existence of *reproducing kernels*. We start with the following proposition:

Proposition 2.1.3. *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions of H^2, and let $f \in H^2$.*

(i) If $f_n \xrightarrow[n\to\infty]{} f$ in H^2, then, $f_n \xrightarrow[n\to\infty]{} f$ uniformly on every compact of $\mathbb{D}$.

(ii) For every $a \in \mathbb{D}$, the linear form $\delta_a : f \mapsto f(a)$ is continuous on H^2.

Applying Riesz's representation theorem, we obtain the following:

Corollary 2.1.4. *For every $a \in \mathbb{D}$, there exists an unique function $k_a \in H^2$ such that, for every $f \in H^2$, $f(a) = \langle f, k_a\rangle_{H^2}$.*

We say that k_a is the reproducing kernel *of H^2 at the point a.*

A straightforward computation gives:

Proposition 2.1.5. *For every $a \in \mathbb{D}$, for every $z \in \mathbb{D}$, $k_a(z) = \frac{1}{1-\overline{a}z}$.*

This function is usually called the Szegö kernel *or the* Cauchy kernel.

Moreover, for $n \geq 0$, we have $f^{(n)}(a) = \left\langle f, k_a^{(n)}\right\rangle$, where $k_a^{(n)}(z) = \frac{n!z^n}{(1-\overline{a}z)^{n+1}}$.

We can also characterize the Hardy-Hilbert space using integral means:

Proposition 2.1.6. *Let $f \in \mathcal{H}(\mathbb{D})$ be an holomorphic function on $\mathbb{D}$.*

$f \in H^2$ if and only if $\sup_{0\leq r<1} \int_0^{2\pi} |f(re^{i\theta})|^2 \frac{d\theta}{2\pi} < \infty$.

In this case, we have:

$$\|f\|_{H^2}^2 = \lim_{r\to 1^-} \int_0^{2\pi} |f(re^{i\theta})|^2 \frac{d\theta}{2\pi}.$$

Remark 2.1.7. From this characterization, we obtain quite easily that every bounded holomorphic function on $\mathbb{D}$ belongs to H^2.

2.1.2 Hardy-Hilbert space on the open unit circle

In this section, if $f \in L^2(\mathbb{T})$ and $n \in \mathbb{Z}$, we denote by $\widehat{f}(n)$ the n^{th} Fourrier coefficient of f, *i.e.* $\widehat{f}(n) = \int_0^{2\pi} f(e^{i\theta})e^{-in\theta}d\theta$.

Definition 2.1.8. The Hardy-Hilbert space on $\mathbb{T}$ is the set of integrable functions on $\mathbb{T}$ whose negative Fourrier coefficients are all zero, *i.e.*

$$H^2(\mathbb{T}) := \left\{f \in L^2(\mathbb{T}) : \widehat{f}(n) = 0, \forall\, n < 0\right\}.$$

As it is closed subspace of $L^2(\mathbb{T})$, it is a Hilbert space. More precisely, one can identify the two spaces $H^2(\mathbb{D})$ and $H^2(\mathbb{T})$:

Proposition 2.1.9. *The two spaces $H^2(\mathbb{D})$ and $H^2(\mathbb{T})$ are isometrically isomorphic through the application:*

$$\begin{array}{cccc} \Phi & : & H^2(\mathbb{T}) & \longrightarrow & H^2(\mathbb{D}) \\ & & f & \longmapsto & \left(z \mapsto \sum_{n=0}^{\infty} \widehat{f}(n) z^n\right) \end{array}$$

whose inverse is:

$$\begin{array}{cccc} \Psi & : & H^2(\mathbb{D}) & \longrightarrow & H^2(\mathbb{T}) \\ & & f & \longmapsto & \left(f^* : e^{i\theta} \mapsto \sum_{n=0}^{\infty} a_n(f) e^{in\theta}\right) \end{array}.$$

The map f^* is called the boundary value *of f and is defined almost everywhere on $\mathbb{T}$.*

Corollary 2.1.10. *Let $f,g \in H^2(\mathbb{D})$.*

$$\langle f,g\rangle_{H^2} = \langle f^*,g^*\rangle_{L^2} = \int_{\mathbb{T}} f^*(\xi)\overline{g^*(\xi)}d\xi.$$

The boundary value have the following properties:

Proposition 2.1.11. *Let $f \in H^2(\mathbb{D})$, $r \in [0,1[$ and $f_r : e^{i\theta} \in \mathbb{T} \mapsto f(re^{i\theta}) = \sum_{n=0}^{\infty} a_n(f) r^n e^{in\theta}$.*

(i) $f_r \in H^2(\mathbb{T}) \cap \mathcal{C}(\mathbb{T})$ and $\widehat{f_r}(n) = c_n(f) r^n$, for all $n \geq 0$;

(ii) $f_r \xrightarrow[r\to 1^-]{} f^*$ in $L^2(\mathbb{T})$.

Theorem 2.1.12 (Fatou's theorem)**.** *Let $f \in H^2(\mathbb{D})$. For almost all $\xi \in \mathbb{T}$, we have:*

$$\lim_{r\to 1^-} f(r\xi) = f^*(\xi).$$

Remark 2.1.13. Proposition 2.1.11 together with the Riesz-Fischer theorem provide immediately the following weak version of Fatou's theorem: for every $f \in H^2(\mathbb{D})$ and for every sequence $(r_n)_{n\in\mathbb{N}} \in [0,1[^{\mathbb{N}}$ converging to 1, there exists a subsequence $\left(r_{\phi(n)}\right)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} f(r_{\phi(n)}\xi) = f^*(\xi)$ almost everywhere on $\mathbb{T}$. Fatou's theorem states that there is in fact no need to restrict oneself to a subsequence.

From now, we will identify $H^2(\mathbb{D})$ and $H^2(\mathbb{T})$, and denote similarly a function $f \in H^2(\mathbb{D})$ and its boundary value f^*.

2.1.3 The Riesz projection

Definition 2.1.14. The *Riesz projection* P is the orthogonal projection from $L^2(\mathbb{T})$ to H^2.

In terms of Fourier series, P is given by the formula:

$$P\left(\sum_{n\in\mathbb{Z}} \widehat{f}(n)\xi^n\right) = \sum_{n\geq 0} \widehat{f}(n)\xi^n.$$

The Riesz projection returns the *analytic part* of a Fourier series in L^2. For instance,

$$P(1+2\cos(\theta)) = P(e^{-i\theta}+1+e^{i\theta}) = 1+e^{i\theta}.$$

We also notice that, as an orthogonal projection, P is self-adjoint.

Proposition 2.1.15. *The Riesz projection is given by the formula:*

$$(Pf)(a) = \langle f,k_a\rangle_{H^2} = \int_{\mathbb{T}} \frac{f(\xi)}{1-\overline{\xi}a}d\xi,$$

where k_a is the Cauchy-Szegö kernel.

2.2 Operators on Hardy spaces

2.2.1 Shifts on H^2

Definition 2.2.1. The *forward shift* is the linear operator $S : H^2 \to H^2$ defined by

$$S(f)(z) = zf(z), \forall z \in \mathbb{D}.$$

In terms of Taylor coefficients, identifying H^2 with $\ell^2(\mathbb{N})$, S can be defined as follows:

$$S(a_0, a_1, \dots,) = (0, a_0, a_1, \dots).$$

Proposition 2.2.2. *With respect to the canonical orthonormal basis of H^2, S has the following matrix representation:*

$$\begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Proposition 2.2.3. *The adjoint of the forward shift is the* backward shift $S^* : H^2 \to H^2$ *defined by*

$$S^*(f)(z) = \frac{f(z) - f(0)}{z}.$$

In terms of Taylor coefficients, identifying H^2 with $\ell^2(\mathbb{N})$, S^* can be defined as follows:

$$S(a_0, a_1, \dots,) = (a_1, a_2, \dots).$$

With respect to the canonical orthonormal basis of H^2, S^* has the following matrix representation:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

2.2.2 Multiplication operators

Definition 2.2.4. An analytic function $\phi : \mathbb{D} \to \mathbb{C}$ is a *multiplier* of H^2 if $\phi \cdot H^2 \subset H^2$.

Note that the set of multipliers of H^2 forms an algebra.

Definition 2.2.5. If ϕ is a multiplier of H^2, the *multiplication operator* with symbol ϕ is the operator $M_\phi : H^2 \to H^2$ defined by:

$$M_\phi(f) = \phi f.$$

Proposition 2.2.6. *Let $H^\infty = H^\infty(\mathbb{D})$ be the set of bounded holomorphic functions on $\mathbb{D}$.*

Then, H^∞ is exactly the multiplier algebra of H^2.

Moreover, for all $\phi \in H^\infty$, $\|M_\phi\| = \|\phi\|_\infty$.

Multiplication operators can be characterized as the commutants of the forward shift. More precisely, we have the following:

Proposition 2.2.7. *A bounded linear operator* $A \in \mathcal{B}\left(H^2\right)$ *satisfies* $SA = AS$ *if and only if* $A = M_\phi$, *for some* $\phi \in H^\infty$.

2.2.3 Toeplitz operators

Definition 2.2.8. For $\phi \in L^\infty$, the *Toeplitz operator* $T_\phi : H^2 \to H^2$ with symbol ϕ is defined by

$$T_\phi(f) = P(\phi f),$$

where P is the Riesz projection of L^2 onto H^2.

Using the integral representation of the Riesz projection, one can write an integral formula for T_ϕ:

$$T_\phi(f)(\lambda) = \int_{\mathbb{T}} \frac{f(\xi)\phi(\xi)}{1-\overline{\xi}\lambda}\mathrm{dm}(\xi),\ \forall\, \lambda \in \mathbb{D}, \forall\, f \in H^2.$$

When $\phi \in H^\infty$, we have $T_\phi(f) = \phi f$, and T_ϕ is just the multiplication operator with symbol ϕ.

Proposition 2.2.9. *With respect to the canonical orthonormal basis of* H^2, T_ϕ *has the the following matrix representation:*

$$\begin{pmatrix} \alpha_0 & \alpha_{-1} & \alpha_{-2} & \alpha_{-3} & \alpha_{-4} & \cdots \\ \alpha_1 & \alpha_0 & \alpha_{-1} & \alpha_{-2} & \alpha_{-3} & \cdots \\ \alpha_2 & \alpha_1 & \alpha_0 & \alpha_{-1} & \alpha_{-2} & \cdots \\ \alpha_3 & \alpha_2 & \alpha_1 & \alpha_0 & \alpha_{-1} & \cdots \\ \alpha_4 & \alpha_3 & \alpha_2 & \alpha_1 & \alpha_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where $\alpha_k = \widehat{\phi}(k)$, *for all* $k \in \mathbb{Z}$.

Proposition 2.2.10 (Brown-Halmos). *For all* $\phi \in L^\infty$, $\|T_\phi\| = \|\phi\|_\infty$

Proposition 2.2.11. *For all* $\phi \in L^\infty$, $T_\phi^* = T_{\overline{\phi}}$.

2.3 Model spaces

2.3.1 Model spaces as invariant subspaces

An important issue in functional analysis is the invariant subspace problem, which can be stated as follows: does every bounded linear operator on a complex Banach space sends some non-trivial closed subspace to itself? For the forward shift operator, the existence of invariant subspace is established, and it is possible to give an explicit characterization of these subspaces.

First of all, we start with an important definition:

Definition 2.3.1. A function $u \in H^\infty$ is said to be *inner* if $|u(\xi)| = 1$ for almost every $\xi \in \mathbb{T}$.

Example 2.3.2. The automorphisms of the open unit disc, *i.e.* functions of the form

$$\tau_{\theta,\omega} : z \mapsto e^{i\theta} \frac{z-\omega}{1-\overline{\omega}z}, \quad \omega \in \mathbb{D}, \theta \in \mathbb{R}$$

are inner. Subsequently, if $\omega_1, \dots, \omega_n \in \mathbb{D}$, $\tau = \tau_{1,\omega_1} \tau_{1,\omega_2} \dots \tau_{1,\omega_n}$ is also an inner function.

Now, we can state the following characterization of the forward shift's invariant subspaces:

Theorem 2.3.3 (Beurling)**.** *The non-trivial invariant subspaces of H^2 under the forward shift S are exactly the subspaces of the form $\mathcal{M} = uH^2$, where u is a non-constant inner function.*

As a corollary of Beurling's theorem, we can also characterize the non-trivial invariant subspaces of the backward shift S^*:

Definition 2.3.4. If u is an inner function, the corresponding *model space* $\mathcal{K}_u$ is defined to be

$$\mathcal{K}_u := (uH^2)^\perp = \{f \in H^2 : \langle f, uh \rangle = 0, \forall\, h \in H^2\}.$$

Corollary 2.3.5. *The model spaces $\mathcal{K}_u$ are precisely the proper S^*-invariant subspaces of H^2.*

2.3.2 Reproducing kernel

As a subspace of the reproducing kernel Hilbert space H^2, each model space $\mathcal{K}_u$ has a reproducing kernel. More precisely, we have:

Proposition 2.3.6. *Let u be an inner function. Then, for every $a \in \mathbb{D}$, the function m_a^u defined by*

$$m_a^u(z) = \frac{1-\overline{u(a)}u(z)}{1-\overline{a}z}, \; z \in \mathbb{D}$$

is the reproducing kernel of $\mathcal{K}_u$ at the point a.

In other words, for each $a \in \mathbb{D}$, the function m_a^u belongs to $\mathcal{K}_u$ and satisfies the reproducing formula

$$f(a) = \langle f, m_a^u \rangle, \text{ for every } f \in \mathcal{K}_u.$$

2.3.3 The projection P_u

Using non-tangential boundary values, $\mathcal{K}_u$ can be seen as a closed subspace of $L^2(\mathbb{T})$. Let P_u be the orthogonal projection of L^2 onto $\mathcal{K}_u$. Then, we have the following relations:

Proposition 2.3.7. *Let $f \in L^2$.*

For every $a \in \mathbb{D}$, $(P_u f)(a) = \langle f, m_a^u \rangle$.

Moreover, $m_a^u = P_u k_a$, where k_a denotes the Cauchy-Szegö kernel.

Proposition 2.3.8. *For all $f \in H^2$, we have:*

$$P_u(f) = f - uP(\overline{u}f),$$

where P is the Riesz projection from L^2 onto H^2.

2.3.4 Finite-dimensional model spaces

The simplest examples of model spaces are those corresponding to finite Blaschke products

$$u(z) = \prod_{k=1}^{n} b_{\omega_k}(z), \text{ where } b_{\omega_k}(z) = \frac{z-\omega_k}{1-\overline{\omega_k}z}.$$

In this case, the vectors in $\mathcal{K}_u$ can be completely characterized in an explicit fashion:

Proposition 2.3.9. *If Θ_n is a finite Blaschke product with distinct zeros $\omega_1,\ldots,\omega_n \in \mathbb{D}$ of respective multiplicities $p_1,\ldots,p_n$, then,*

$$\mathcal{K}_{\Theta_n} = \operatorname{span}\left\{k_{\omega_i}^{(l_i-1)} : 1 \leq i \leq n,\, 1 \leq l_i \leq p_i\right\},$$

where $k_\omega^n(z) = \frac{n!z^n}{(1-\overline{\omega}z)^{n+1}}$.

In particular, $\dim \mathcal{K}_{\Theta_n} = p_1 + \cdots + p_n$.

Furthermore, model spaces derived from finite Blaschke products are the only finite-dimensional model spaces:

Proposition 2.3.10. *Let u be an inner function.*

$\dim \mathcal{K}_u < \infty$ *if and only if u is a finite Blaschke product.*

Unlike the Hardy space H^2, which has the natural orthonormal basis $(z^n)_{n\in\mathbb{N}}$, there is no canonical orthonormal basis for a model space $\mathcal{K}_u$. However, when u is a finite Blaschke product, an orthonormal basis can be obtained by orthogonalizing the kernel functions corresponding to the zeros of u. The orthonormal basis that is obtained is usually called the *Takenaka-Mamquist-Walsh basis*:

Theorem 2.3.11. *Let Θ_n be a finite Blaschke product with zeros $\omega_1,\ldots,\omega_n$ (repeated according to multiplicity), let $b_{\omega_i}(z) = \frac{z-\omega_i}{1-\overline{\omega_i}z}$, for all $i \in [\![1,n]\!]$, and let:*

$$\phi_k(z) = \left(\prod_{j=k+1}^{n} b_{\omega_j}(z)\right) \frac{\sqrt{1-|\omega_k|^2}}{1-\overline{\omega_k}z}, \text{ for all } k \in [\![1,n]\!]$$

Then, $(\phi_1,\ldots,\phi_n)$ forms an orthonormal basis of $\mathcal{K}_{\Theta_n}$, which is called the Takenaka-Malmquist-Walsh basis.

2.3.5 The compressed shift

Definition 2.3.12. Let H be a separable complex Hilbert space, let $\mathcal{M}$ be a closed subspace of H, and let $T \in \mathcal{B}(H)$. We define the operator $R \in \mathcal{B}(\mathcal{M})$ by $R = P_{\mathcal{M}} T_{|\mathcal{M}}$, where $P_{\mathcal{M}}$ is the orthogonal projection from H onto $\mathcal{M}$.

According to the orthogonal decomposition $H = \mathcal{M} \oplus \mathcal{M}^\perp$, T has the matrix representation $\begin{bmatrix} R & * \\ * & * \end{bmatrix}$.

We say that R is a *compression* of T to the subspace $\mathcal{M}$ (and that T is a *dilation* of R) if, for every $n \in \mathbb{N}$, we have $R^n = P_{\mathcal{M}} T^n_{|\mathcal{M}}$.

Definition 2.3.13. For an inner function u, the operator

$$\begin{array}{rccc} S_u: & \mathcal{K}_u & \longrightarrow & \mathcal{K}_u \\ & f & \longmapsto & P_u S f \end{array}$$

is called the *compressed shift operator*.

Those two definitions coincide in the following sense:

Theorem 2.3.14. *For an inner function u, the operator S_u is a compression of S.*

Moreover, we can show that S_u^* is the restriction of the backward shift operator S^* to $\mathcal{K}_u$:

Proposition 2.3.15. *For any inner function u and for any function $f \in \mathcal{K}_u$,*

$$S_u^*(f) = \frac{f - f(0)}{z}.$$

One of the main motivation for studying model spaces and compressed shifts is the theory of model operators, developed by Sz.-Nagy and Foias (see [59] or [27, Chapter 9]), where it is shown that certain types of Hilbert space contractions are unitarily equivalent to the compressions of the unilateral shift to a model space. More precisely, we have:

Theorem 2.3.16 (Sz.-Nagy–Foias)**.** *If T is a contraction on a Hilbert space H satisfying:*

(i) $\|T^{*n}x\| \xrightarrow[n\to\infty]{} 0$ *for all* $x \in H$*;*

(ii) $\operatorname{rank}(\mathrm{Id} - T^*T) = \operatorname{rank}(\mathrm{Id} - TT^*) = 1$*;*

then, there exists an inner function u such that T is unitarily equivalent to S_u.

2.3.6 Model matrices

In the case where Θ_n is a finite Blaschke product, we can write the matrix representation of S_{Θ_n} in the Takenaka-Malmquist-Walsh basis. A straightforward computation gives:

Theorem 2.3.17. *Let Θ_n be a finite Blaschke product with distinct zeros $\omega_1, \ldots, \omega_n \in \mathbb{D}$. The matrix representation M_{Θ_n} of S_{Θ_n} in the Takenaka-Malmquist-Walsh basis is given by:*

$$\left[M_{\Theta_n}\right]_{i,j} = \left\langle z\phi_j, \phi_i \right\rangle = \begin{cases} \omega_j & \text{if } i{=}j \\ \prod_{k=i+1}^{j-1} (-\overline{\omega_k})\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_j|^2} & \text{if } i{<}j \\ 0 & \text{if } i{>}j \end{cases} \tag{2.1}$$

If Θ_n has distinct zeros, then, M_{Θ_n} has distinct eigenvalues and, thus, it can be diagonalized (see [10, Section 5]):

Theorem 2.3.18. *Let Θ_n be a finite Blaschke product with distinct zeros $\omega_1, \ldots, \omega_n \in \mathbb{D}$, and let M_{Θ_n} be the matrix representation of S_{Θ_n} in the Takenaka-Malmquist-Walsh basis.*

Then, $M_{\Theta_n} = PDP^{-1}$, where D is the diagonal matrix with diagonal entries $\omega_1,\dots,\omega_n$, and

$$[P]_{i,j} = \begin{cases} 1 & \text{if } i = j \\ \frac{\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_j|^2}}{\omega_j-\omega_i}\prod_{k=i+1}^{j-1}\left(\frac{1-\overline{\omega_k}\omega_j}{\omega_j-\omega_k}\right) & \text{if } i < j \\ 0 & \text{if } i > j \end{cases}$$

$$[P^{-1}]_{i,j} = \begin{cases} 1 & \text{if } i = j \\ \frac{\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_j|^2}}{\omega_i-\omega_j}\prod_{k=i+1}^{j-1}\left(\frac{1-\overline{\omega_k}\omega_i}{\omega_i-\omega_k}\right) & \text{if } i < j \\ 0 & \text{if } i > j \end{cases}.$$

Chapter 3

Function of a matrix

The aim of this chapter is – given a function f and an upper-triangular matrix T – to give an explicit expression for $f(T)$. Recall (see Lemma 1.2.15) that for a 2×2 upper-triangular matrix

$$T_2 = \begin{pmatrix} \omega_1 & \alpha \\ 0 & \omega_2 \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$$

and a polynomial $p \in \mathbb{C}[X]$, a straightforward computation gives

$$p(T_2) = \begin{pmatrix} p(\omega_1) & \alpha\left(\frac{p(\omega_1)-p(\omega_2)}{\omega_1-\omega_2}\right) \\ 0 & p(\omega_2) \end{pmatrix} \tag{3.1}$$

Then, we can obtain a similar formula for $f(T_2)$, where f is a function holomorphic in the neighborhood of ω_1, ω_2, using classical approximation arguments. However, for a matrix T_n of size $n \geq 3$, straightforward computations give more intricate expressions for $p(T_n)$. In this chapter, we will generalize (3.1) to matrices of bigger size.

Similarly, given a polynomial $p \in \mathbb{C}[X]$ and a 2×2 matrix

$$\widetilde{T_2} = \begin{pmatrix} W_1 & C \\ 0 & W_2 \end{pmatrix},$$

whose coefficients are operators acting on a Hilbert space, how can we express $p(T_2)$?

A straightforward computation gives

$$\widetilde{T_2}^{\,k} = \begin{pmatrix} W_1^k & \sum_{j=0}^{k-1} W_1^{k-j-1} C W_2^j \\ 0 & W_2^k \end{pmatrix}, \quad \text{for } k \in \mathbb{N},$$

which generalizes (3.1) in the non-commutative case, in view of the well-known identity

$$(a-b)\sum_{j=0}^{k-1} a^{k-j-1} b^j = a^k - b^k,\ \forall\, a, b \in \mathbb{C},\ \forall\, k \in \mathbb{N}.$$

In this chapter, we will also generalize this formula to operator matrices of arbitrary size.

Observe that those two problems can be seen in a unified way, as a $n \times n$ scalar upper-triangular matrix T_n can be seen as a 2×2 operator upper-triangular matrix using the following block decomposition:

$$T_n = \left(\begin{array}{ccccc|c}
\omega_1 & \alpha_1^{(1)} & \cdots & \cdots & \alpha_1^{(n-2)} & \alpha_1^{(n-1)} \\
0 & \omega_2 & \alpha_2^{(1)} & \cdots & \alpha_2^{(n-3)} & \alpha_2^{(n-2)} \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\
\vdots & & \ddots & \ddots & \alpha_{n-2}^{(1)} & \vdots \\
0 & \cdots & \cdots & 0 & \omega_{n-1} & \alpha_{n-1}^{(1)} \\
\hline
0 & \cdots & \cdots & \cdots & 0 & \omega_n
\end{array}\right) = \left(\begin{array}{c|c}
T_{n-1} & * \\
\hline
0 & \omega_n
\end{array}\right).$$

Moreover, observe that in (3.1), the quantity $\frac{f(\omega_1)-f(\omega_2)}{\omega_1-\omega_2}$ can be seen as a first order divided difference.

The key point of this chapter will be the notion of *non-commutative divided differences*, which generalizes the well-known notion of divided differences for scalars. What follows is in line with the work presented in various references, as *e.g.* [33, 36, 16, 31, 40]. We offer here a slightly different perspective.

3.1 Usual divided differences

Before defining the *non-commutative divided differences*, we will quickly recall some *well-known* facts about *usual* divided differences. One can refer for instance to [18, 58] for a more complete overview about that.

Definition 3.1.1. Let $n \in \mathbb{N}$ and let $(x_0, y_0), \dots, (x_n, y_n)$ be $n+1$ points of $\mathbb{C}^2$ such that $x_i \neq x_j$ when $i \neq j$. The *divided differences* are defined recursively as follows:

$$[y_k] := y_k, \; \forall \, k \in [\![0, n]\!]$$

$$[y_k, \dots, y_{k+j}] := \frac{[y_{k+1}, \dots, y_{k+j}] - [y_k, \dots, y_{k+j-1}]}{x_{k+j} - x_k}, \; \forall k \in [\![0, n-j]\!], \forall \, j \in [\![1, n]\!].$$

In the following, we will always assume that $y_k = f(x_k)$, for some map $f : \mathbb{K} \to \mathbb{K}$.

The divided differences are usually known for being used in Newton-Lagrange's polynomial interpolation :

Theorem 3.1.2. *Let $x_0, \dots, x_n$ be $(n+1)$ distinct points of $\mathbb{K}$, and let $f : \mathbb{K} \to \mathbb{K}$ be a map. Then, the unique polynomial P_n of degree less than or equal to n such that $P_n(x_k) = f(x_k)$, for all $k \in [\![0, n]\!]$, is given by :*

$$P_n(x) = \sum_{k=0}^{n} f(x_k) \prod_{\substack{i=0 \\ i \neq k}}^{n} \frac{x - x_i}{x_k - x_i} \qquad \text{(Lagrange's formula)}$$

$$= \sum_{k=0}^{n} [f(x_0), \ldots, f(x_k)] \prod_{i=0}^{k-1} (x - x_i) \qquad \textit{(Newton's formula)}$$

P_n is said to be the interpolation polynomial of degree n of f.

The following explicit formula for divided differences can be seen as a corollary of the previous theorem (by identifying the coefficient of degree n in both formulas), but it can also be proved very easily by induction, by using only the recursive definition:

Proposition 3.1.3.

$$[f(x_0), \ldots, f(x_n)] = \sum_{k=0}^{n} \frac{f(x_k)}{\prod_{\substack{0 \le i \le n \\ i \ne k}} (x_k - x_i)}$$

From that formula, we can deduce the following noteworthy properties:

Proposition 3.1.4.

(i) Invariance under permutation of indices: $[f(x_{\sigma(0)}), \ldots, f(x_{\sigma(n)})] = [f(x_0), \ldots, f(x_n)]$, $\forall\, \sigma \in \mathcal{S}_{n+1}$;

(ii) Linearity: $[(f + \lambda g)(x_0), \ldots, (f + \lambda g)(x_n)] = [f(x_0), \ldots, f(x_n))] + \lambda\, [g(x_0), \ldots, g(x_n)]$, $\forall\, f, g : \mathbb{K} \to \mathbb{K}$, $\forall\, \lambda \in \mathbb{K}$.

In particular, if $f = \sum_{k=0}^{p} a_k X^k \in \mathbb{K}[X]$, $[f(x_1), \ldots, f(x_n)] = \sum_{k=0}^{p} a_k \left[x_1^k, \ldots, x_n^k\right]$;

(iii) Sequential continuity : if f is the pointwise limit of a sequence $(f_k)_{k \in \mathbb{N}}$, then

$$[f(x_0), \ldots, f(x_n)] = \lim_{k \to \infty} [f_k(x_0), \ldots, f_k(x_n)].$$

Using Newton's interpolation formula and Rolle's theorem, one can prove the following result:

Proposition 3.1.5 (Mean value theorem for divided differences)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be a n times differentiable function, and let $x_0 < \cdots < x_n \in \mathbb{R}$. Then, there exists $\xi \in]x_0, x_n[$ such that:*

$$[f(x_0), \ldots, f(x_n)] = \frac{f^{(n)}(\xi)}{n!}.$$

Another important result on divided difference is the following, which generalizes the Leibniz's rule for derivatives:

Proposition 3.1.6 (Leibniz's formula)**.** *Let $f, g : \mathbb{C} \to \mathbb{C}$ be two applications, and let $x_0, \ldots, w_n \in \mathbb{C}$ be $n + 1$ pairwise distinct points.*
Then, we have:

$$[(f \cdot g)(x_0), \ldots, (f \cdot g)(x_n)] = \sum_{j=0}^{n} [f(x_0), \ldots, f(x_j)] \cdot [g(x_j), \ldots, g(x_n)].$$

We will now focus on the case where f is a polynomial. The following proposition generalizes the well-known identity

$$(a-b)\sum_{j=0}^{k-1} a^{k-j-1}b^j = a^k - b^k,\ \forall\, a,b \in \mathbb{C},\ \forall\, k \in \mathbb{N}.$$

Proposition 3.1.7. *Let $k \in \mathbb{N}$ and let $x_0,\dots,x_n \in \mathbb{C}$ be $n+1$ pairwise distinct points.*

(i) $[x_0^k,\dots,x_n^k] = \sum_{\substack{i_0+\dots+i_n=k-n \\ i_0,\dots,i_n\geq 0}} x_0^{i_0}x_1^{i_1}\dots x_n^{i_n}$ *;*

(ii) $[x_0^k,\dots,x_n^k] = 0$ *provided* $n>k$ *;*

(iii) $[x_0^k,\dots,x_n^k] = \sum_{j=0}^{k-1}\left[x_0^{k-j-1},\dots,x_{n-1}^{k-j-1}\right]x_n^j = \sum_{j=0}^{k-n}\left[x_0^{k-j-1},\dots,x_{n-1}^{k-j-1}\right]x_n^j.$

Proof.

(i) By induction on the number of points:

- For one point x_0, this is obvious.
- Now assume that the formula is true for n pairwise distinct points. Then, we will prove that for $n+1$ pairwise distinct points $x_0,\dots,x_n$, we have:

$$(x_0-x_n)\sum_{i_0+\dots+i_n=k-n} x_0^{i_0}x_1^{i_1}\dots x_n^{i_n} = [x_0^k,\dots,x_{n-1}^k]-[x_1^k,\dots,x_n^k].$$

Indeed,

$$\begin{aligned}
&(x_0-x_n)\sum_{i_0+\dots+i_n=k-n} x_0^{i_0}x_1^{i_1}\dots x_n^{i_n}\\
=&\sum_{\substack{j_0+\dots+j_n=k-(n-1)\\ j_0\geq 1}} x_0^{j_0}x_1^{j_1}\dots x_n^{j_n} - \sum_{\substack{j_0+\dots+j_n=k-(n-1)\\ j_n\geq 1}} x_0^{j_0}x_1^{j_1}\dots x_n^{j_n}\\
=&\sum_{j_0+\dots+j_n=k-(n-1)} x_0^{j_0}x_1^{j_1}\dots x_n^{j_n} - \sum_{j_1+\dots+j_n=k-(n-1)} x_1^{j_1}\dots x_n^{j_n}\\
&-\sum_{j_0+\dots+j_n=k-(n-1)} x_0^{j_0}x_1^{j_1}\dots x_n^{j_n} + \sum_{j_0+\dots+j_{n-1}=k-(n-1)} x_0^{j_0}x_1^{j_1}\dots x_{n-1}^{j_{n-1}}\\
=&\sum_{j_0+\dots+j_{n-1}=k-(n-1)} x_0^{j_0}x_1^{j_1}\dots x_{n-1}^{j_{n-1}} - \sum_{j_1+\dots+j_n=k-(n-1)} x_1^{j_1}\dots x_n^{j_n}\\
=&[x_0^k,\dots,x_{n-1}^k]-[x_1^k,\dots,x_n^k] \text{ by induction hypothesis.}
\end{aligned}$$

(ii) The second item is just a direct consequence of the first one (but it can also be seen as a consequence of the interpolation theorem, noticing that the interpolation polynomial of $x\mapsto x^k$ is X^k).

(iii) The third item is also a consequence of the first one :

$$
\begin{aligned}
\sum_{j=0}^{k-1}\left[x_0^{k-j-1},\dots,x_{n-1}^{k-j-1}\right]x_n^j &= \sum_{j=0}^{k-n}\left[x_0^{k-j-1},\dots,x_{n-1}^{k-j-1}\right]x_n^j\\
&= \sum_{i_n=0}^{k-n}\left(\sum_{i_0+\dots+i_{n-1}=k-n-i_n} x_0^{i_0}x_1^{i_1}\dots x_{n-1}^{i_{n-1}}\right)x_n^{i_n}\\
&= \sum_{i_0+\dots+i_n=k-n} x_0^{i_0}x_1^{i_1}\dots x_n^{i_n}\\
&= \left[x_0^k,\dots,x_n^k\right].
\end{aligned}
$$

□

Similarly, we can prove the following proposition:

Proposition 3.1.8. *Let $k\in\mathbb{N}^*$ and let $x_0,\dots,x_n$ be $n+1$ pairwise distinct points.*

(i) $[x_0^{-k},\dots,x_n^{-k}] = (-1)^n\sum_{\substack{i_0+\dots+i_n=k+n\\ i_0,\dots,i_n\geq 1}} x_0^{-i_0}x_1^{-i_1}\dots x_n^{-i_n}$;

(ii) $\left[x_0^{-k},\dots,x_n^{-k}\right] = -\sum_{j=1}^{k}\left[x_0^{-j},\dots,x_{n-1}^{-j}\right]x_n^{j-k-1}$

Last but not least, in the case where f is an holomorphic function, we have the following noteworthy formula:

Proposition 3.1.9. *Let f be an analytic function in an open neighborhood of the the points $z_0,\dots,z_n\in\mathbb{C}$. Then, we have:*

$$[f(z_0),\dots,f(z_n)] = \frac{1}{2i\pi}\int_\Gamma \frac{f(\xi)}{\prod_{k=0}^n \xi - z_k},$$

where Γ is a finite system of rectifiable Jordan curves oriented in the positive sense surrounding $z_0,\dots,z_n$.

3.2 Non-commutative divided differences

3.2.1 Definition and elementary properties

Definition 3.2.1. Let $n\geq 1$, and let $H_0,\dots,H_n$ be Hilbert spaces. Given bounded linear operators $W_i\in\mathcal{B}(H_i)$, for $i\in[\![0,n]\!]$, and $C_j\in\mathcal{B}\left(H_{j+1},H_j\right)$, for $j\in[\![0,n-1]\!]$, we define the *non-commutative divided differences* as follows:

(i) $\forall\, k\in\mathbb{N}$, $[W_0^k] := W_0^k$;

(ii) $\forall\, k\in\mathbb{N}$, $[W_0^k,\dots,W_n^k]_{(C_0,\dots,C_{n-1})} := \sum_{\substack{i_0+\dots+i_n=k-n\\ i_0,\dots,i_n\geq 0}} W_0^{i_0}C_0W_1^{i_1}\dots C_{n-1}W_n^{i_n}$;

(iii) $\forall\, p=\sum_{k=0}^d a_kX^k\in\mathbb{C}[X]$, $[p(W_0),\dots,p(W_n)]_{(C_0,\dots,C_{n-1})} := \sum_{k=0}^d a_k\left[W_0^k,\dots,W_n^k\right]_{(C_0,\dots,C_{n-1})}$.

Remark 3.2.2. In view of Proposition 3.1.7, if $W_0,\dots,W_n,C_0,\dots,C_{n-1}$ are scalars, we recover the *usual* divided differences:

$$[p(W_0),\dots,p(W_n)]_{(C_0,\dots,C_{n-1})} = C_0\dots C_{n-1}[p(W_0),\dots,p(W_n)] = [p(W_0),\dots,p(W_n)]C_0\dots C_{n-1}.$$

Proposition 3.2.3. *With the above notation, we have:*

(i) $\left[W_0^k,\dots,W_n^k\right]_{(C_0,\dots,C_{n-1})} = 0$ *provided* $n > k$ *;*

(ii) $$\left[W_0^k,\dots,W_n^k\right]_{(C_0,\dots,C_{n-1})} = \sum_{j=0}^{k-1}\left[W_0^{k-j-1},\dots,W_{n-1}^{k-j-1}\right]_{(C_0,\dots,C_{n-2})} C_{n-1}W_n^j$$
$$= \sum_{j=0}^{k-n}\left[W_0^{k-j-1},\dots,W_{n-1}^{k-j-1}\right]_{(C_0,\dots,C_{n-2})} C_{n-1}W_n^j.$$

Proposition 3.2.4. *With the above notation, we have:*

$$\begin{aligned}&W_0\,[p(W_0),\dots,p(W_n)]_{(C_0,\dots,C_{n-1})} - [p(W_0),\dots,p(W_n)]_{(C_0,\dots,C_{n-1})}\,W_n\\ &\quad= [p(W_0),\dots,p(W_{n-1})]_{(C_0,\dots,C_{n-2})}C_{n-1} - C_0[p(W_1),\dots,p(W_n)]_{(C_1,\dots,C_{n-1})}\end{aligned}.$$

In particular,

$$W_0[p(W_0),p(W_1)]_{(C_0)} - [p(W_0),p(W_1)]_{(C_0)}W_1 = p(W_0)C_0 - C_0p(W_1).$$

Remark 3.2.5. If $W_0,\dots,W_n,C_1,\dots,C_n$ are scalars (with $W_0 \neq W_n$ and $C_i \neq 0$, for all $i \in [\![0,n-1]\!]$) then, this identity is equivalent to:

$$[p(W_0),\dots,p(W_n)] = \frac{[p(W_0),\dots,p(W_{n-1})] - [p(W_1),\dots,p(W_n)]}{W_0 - W_n}.$$

Proof of Proposition 3.2.4. First, note that by linearity, it is enough to show this identity for the monomials $p(X) = X^k$, $k \in \mathbb{N}$.

Now, we have:

$$\begin{aligned}
&W_0[W_0^k,\dots W_n^k]_{(C_0,\dots,C_{n-1})} - [W_0^k,\dots W_n^k]_{(C_0,\dots,C_{n-1})}W_n\\
&= \sum_{i_0+\dots+i_n=k-n} W_0^{i_0+1}C_0W_1^{i_1}\dots C_{n-1}W_n^{i_n} - \sum_{i_0+\dots+i_n=k-n} W_0^{i_0}C_0W_1^{i_1}\dots C_{n-1}W_n^{i_n+1}\\
&= \sum_{\substack{i_0+\dots+i_n=k-n+1\\ i_0\geq 1}} W_0^{i_0}C_0W_1^{i_1}\dots C_{n-1}W_n^{i_n} - \sum_{\substack{i_0+\dots+i_n=k-n+1\\ i_n\geq 1}} W_0^{i_0}C_0W_1^{i_1}\dots C_{n-1}W_n^{i_n}\\
&= \sum_{i_0+\dots+i_n=k-n+1} W_0^{i_0}C_0W_1^{i_1}\dots C_{n-1}W_n^{i_n} - \sum_{i_0+\dots+i_n=k-n+1} W_0^{i_0}C_0W_1^{i_1}\dots C_{n-1}W_n^{i_n}\\
&\quad - C_0\sum_{i_1+\dots+i_n=k-n+1} W_1^{i_1}C_1W_2^{i_2}\dots C_{n-1}W_n^{i_n} + \sum_{i_0+\dots+i_n=k-n+1} W_0^{i_0}C_0W_1^{i_1}\dots C_{n-2}W_{n-1}^{i_{n-1}}C_{n-1}\\
&= [W_0^k,\dots,W_{n-1}^k]_{(C_0,\dots,C_{n-2})}C_{n-1} - C_0[W_1^k,\dots,W_n^k]_{(C_1,\dots,C_{n-1})}.
\end{aligned}$$

$\square$

3.2.2 An application to functional calculus for polynomials

Let $n \geq 2$, let $H_1, \ldots, H_n$ be Hilbert spaces, and let

$$\widetilde{T_n} = \begin{pmatrix} W_1 & C_1^{(1)} & \cdots & \cdots & C_1^{(n-1)} \\ 0 & W_2 & C_2^{(1)} & \cdots & C_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & C_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & W_n \end{pmatrix} : \bigoplus_{i=1}^{n} H_i \to \bigoplus_{i=1}^{n} H_i$$

be a bounded linear operator and let $p \in \mathbb{C}[X]$ be a polynomial.

The non-commutative divided differences enable us to compute easily $p\left(\widetilde{T_n}\right)$, by induction on n.

Indeed, for $n = 2$, we have $p\left(\widetilde{T_2}\right) = \begin{pmatrix} p(W_1) & [p(W_1), p(W_2)]_{\left(C_1^{(1)}\right)} \\ 0 & p(W_2) \end{pmatrix}$ and, then, for all $n \geq 3$, we can write $\widetilde{T_n} = \left(\begin{array}{c|c} \widetilde{T_{n-1}} & \widetilde{C} \\ \hline 0 & W_n \end{array}\right)$, with $\widetilde{C} = \left[C_1^{(n-1)} \quad C_2^{(n-2)} \quad \ldots \quad C_{n-1}^{(1)}\right]^{\intercal}$ and, then, we obtain

$$p\left(\widetilde{T_n}\right) = \left(\begin{array}{c|c} p\left(\widetilde{T_{n-1}}\right) & \left[p\left(\widetilde{T_{n-1}}\right), p(W_n)\right]_{\left(\widetilde{C}\right)} \\ \hline 0 & p(W_n) \end{array}\right).$$

More precisely, we obtain the following result:

Theorem 3.2.6. *With the above notation, we have :*

$$p\left(\widetilde{T_n}\right) = \begin{pmatrix} p(W_1) & D_1^{(1)} & \cdots & \cdots & D_1^{(n-1)} \\ 0 & p(W_2) & D_2^{(1)} & \cdots & D_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & D_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & p(W_n) \end{pmatrix},$$

$$\text{where} \quad D_i^{(j)} = \sum_{\substack{1 \leq l \leq j \\ i = k_0 < \cdots < k_l = j+i}} \left[p\left(W_{k_0}\right), \ldots, p\left(W_{k_l}\right)\right]_{\left(C_{k_0}^{(k_1-k_0)}, \ldots, C_{k_{l-1}}^{(k_l-k_{l-1})}\right)}.$$

Proof. First, observe that, by linearity, it is enough to prove this formula in the case where p is a monomial, *i.e.* when $p(X) = X^m$, for some $m \in \mathbb{N}$.

Now, we will prove this formula by induction on n.

- For $n = 2$, it is easy to see (by induction on m) that $\widetilde{T_1}^m = \begin{pmatrix} W_0^m & [W_0^m, W_1^m]_{\left(C_1^{(1)}\right)} \\ 0 & W_1^m \end{pmatrix}$, for every $m \in \mathbb{N}$.

- Now, let $n \geq 2$, and assume that the formula is true for $\widetilde{T_1}$ and $\widetilde{T_{n-1}}$.

  As mentioned above, we write $\widetilde{T_n} = \left(\begin{array}{c|c} \widetilde{T_{n-1}} & \widetilde{C} \\ \hline 0 & W_n \end{array}\right)$, with $\widetilde{C} = \begin{bmatrix} C_1^{(n-1)} & C_2^{(n-2)} & \dots & C_{n-1}^{(1)} \end{bmatrix}^\mathsf{T}$ and, then, we obtain

$$\widetilde{T_n}^m = \left(\begin{array}{c|c} \widetilde{T_{n-1}}^m & \left[\widetilde{T_{n-1}}^m, W_n^m\right]_{(\widetilde{C})} \\ \hline 0 & W_n^m \end{array}\right).$$

  Now, just replace $\widetilde{T_{n-1}}^m$ by its *explicit* expression, *i.e.*

$$\widetilde{T_{n-1}}^m = \begin{pmatrix} W_1^m & D_1^{(1)} & \cdots & \cdots & D_1^{(n-2)} \\ 0 & W_2^m & D_2^{(1)} & \cdots & D_2^{(n-3)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & D_{n-2}^{(1)} \\ 0 & \cdots & \cdots & 0 & W_{n-1}^m \end{pmatrix},$$

$$\text{where} \quad D_i^{(j)} = \sum_{\substack{1 \leq l \leq j \\ i=k_0<\cdots<k_l=j+i}} \left[W_{k_0}^m, \dots, W_{k_l}^m\right]_{\left(C_{k_0}^{(k_1-k_0)}, \dots, C_{k_{l-1}}^{(k_l-k_{l-1})}\right)}$$

  and compute

$$\begin{bmatrix} D_1^{(n-1)} \\ D_2^{(n-2)} \\ \vdots \\ D_{n-1}^{(1)} \end{bmatrix} := \left[\widetilde{T_{n-1}}^m, W_n^m\right]_{(\widetilde{C})} = \sum_{p=0}^{m-1} \widetilde{T_{n-1}}^{m-p-1} \widetilde{C} W_n^p.$$

  Using Proposition 3.2.3, we have, for $i \in [\![1, n-1]\!]$,

$$\begin{aligned}
D_i^{(n-i)} &= \sum_{p=0}^{m-1} \left(W_i^{m-p-1} C_i^{(n-i)} + \sum_{j=1}^{n-i-1} D_i^{(j)} C_{j+i}^{(n-j-i)} \right) W_n^p \\
&= \sum_{p=0}^{m-1} \left(W_i^{m-p-1} C_i^{(n-i)} + \sum_{j=1}^{n-i-1} \sum_{\substack{1 \leq l \leq j \\ i=k_0<\cdots<k_l=j+i}} \left[W_{k_0}^{m-p-1}, \dots, W_{k_l}^{m-p-1}\right]_{\left(C_{k_0}^{(k_1-k_0)}, \dots, C_{k_{l-1}}^{(k_l-k_{l-1})}\right)} C_{j+i}^{(n-j-i)} \right) W_n^p \\
&= [W_i^m, W_n^m]_{\left(C_i^{(n-i)}\right)} + \sum_{j=1}^{n-1} \sum_{\substack{1 \leq l \leq j \\ i=k_0<\cdots<k_l=j+i}} \left[W_{k_0}^m, \dots, W_{k_l}^m, W_n^m\right]_{\left(C_{k_0}^{(k_1-k_0)}, \dots, C_{k_{l-1}}^{(k_l-k_{l-1})}, C_{j+i}^{(n-j-i)}\right)} \\
&= [W_i^m, W_n^m]_{\left(C_i^{(n-i)}\right)} + \sum_{j=1}^{n-1} \sum_{\substack{1 \leq l \leq j \\ i=k_0<\cdots<k_l=j+i \\ k_{l+1}=n}} \left[W_{k_0}^m, \dots, W_{k_l}^m, W_{(k_{l+1})}^m\right]_{\left(C_{k_0}^{(k_1-k_0)}, \dots, C_{k_{l-1}}^{(k_l-k_{l-1})}, C_{k_l}^{(k_{l+1}-k_l)}\right)}
\end{aligned}$$

$$= \sum_{\substack{1\le l\le n \\ i=k_0<\cdots<k_l=n}} \left[W_{k_0}^m,\dots,W_{k_l}^m\right]_{\left(C_{k_0}^{(k_1-k_0)},\dots,C_{k_{l-1}}^{(k_l-k_{l-1})}\right)}.$$

□

Corollary 3.2.7. *Let*

$$\widetilde{T}_n := \begin{pmatrix} W_1 & C_1 & 0 & \cdots & 0 \\ 0 & W_2 & C_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & C_{n-1} \\ 0 & \cdots & \cdots & 0 & W_n \end{pmatrix} : \bigoplus_{i=1}^n H_i \to \bigoplus_{i=1}^n H_i$$

and let $p \in \mathbb{C}[X]$. Then, we have:

$$p\left(\widetilde{T}_n\right) = \begin{pmatrix} p(W_1) & D_1^{(1)} & \cdots & \cdots & D_1^{(n-1)} \\ 0 & p(W_2) & D_2^{(1)} & \cdots & D_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & D_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & p(W_n) \end{pmatrix},$$

where $D_i^{(j)} = [p(W_i),\dots,p(W_{i+j})]_{(C_i,\dots,C_{i+j-1})}$.

A noteworthy application of Theorem 3.2.6 is the following: given an upper-triangular matrix

$$T_n = \begin{pmatrix} \omega_1 & \alpha_1^{(1)} & \cdots & \cdots & \alpha_1^{(n-1)} \\ 0 & \omega_2 & \alpha_2^{(1)} & \cdots & \alpha_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \alpha_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & \omega_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$$

and a polynomial $p \in \mathbb{C}[X]$, we can compute very easily $p(T_n)$, without diagonalizing T_n (nor computing its Dunford decomposition):

Corollary 3.2.8. *With the above notation,*

$$p(T_n) = \begin{pmatrix} p(\omega_1) & \beta_1^{(1)} & \cdots & \cdots & \beta_1^{(n-1)} \\ 0 & p(\omega_2) & \beta_2^{(1)} & \cdots & \beta_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \beta_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & p(\omega_n) \end{pmatrix},$$

$$\textit{where} \quad \beta_i^{(j)} = \sum_{\substack{1\le l\le j \\ i=k_0<\cdots<k_l=j+i}} \left(\prod_{s=0}^{l-1} \alpha_{k_s}^{(k_{s+1}-k_s)}\right)\left[p\left(\omega_{k_0}\right),\dots,p\left(\omega_{k_l}\right)\right].$$

Remark 3.2.9. The *sequential continuity* in Proposition 3.1.4 enables us to extend this result to any function that can be *approximated* by polynomials.

Another noteworthy application of Theorem 3.2.6 is that those *non-commutative divided differences* satisfy the Leibniz's rule too:

Corollary 3.2.10 (Leibniz's formula). *Let $p, q \in \mathbb{C}[X]$ be polynomials, let $n \geq 1$, and let $H_0, \ldots, H_n$ be Hilbert spaces. Given bounded linear operators $W_i \in \mathcal{B}(H_i)$, for $i \in [\![0, n]\!]$, and $C_j \in \mathcal{B}\left(H_{j+1}, H_j\right)$, for $j \in [\![0, n-1]\!]$, we have:*

$$[(p\cdot q)(W_0), \ldots, (p\cdot q)(W_n)]_{(C_0,\ldots,C_{n-1})} = \sum_{j=0}^{n} \left[p(W_0), \ldots, p(W_j)\right]_{(C_0,\ldots,C_{j-1})} \cdot \left[q(W_j), \ldots, q(W_n)\right]_{(C_j,\ldots,C_{n-1})}.$$

Proof. Let

$$\widetilde{T_n} := \begin{pmatrix} W_0 & C_0 & 0 & \cdots & 0 \\ 0 & W_1 & C_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & C_{n-1} \\ 0 & \cdots & \cdots & 0 & W_n \end{pmatrix}.$$

We write $(p\cdot q)\left(\widetilde{T_n}\right) = p\left(\widetilde{T_n}\right)\cdot q\left(\widetilde{T_n}\right)$ and, then, we apply Corollary 3.2.7 to write the explicit formulas for $(p\cdot q)\left(\widetilde{T_n}\right)$, $p\left(\widetilde{T_n}\right)$ and $q\left(\widetilde{T_n}\right)$. We obtain the Leibniz's formula by identifying the coefficient in the top-right hand corner on both side of the equality. □

3.2.3 Extension to rational fractions

Our goal is now to extend the definition of *non-commutative divided differences* to rational fractions. Using the complex partial fraction decomposition, it is enough to give a consistent definition for rational fractions of the form $r(X) = X^{-k}$, with $k \in \mathbb{N}^*$.

Definition 3.2.11. Let $n \geq 1$, and let $H_0, \ldots, H_n$ be Hilbert spaces. Given $n+1$ invertible bounded linear operators $W_i \in \mathcal{B}(H_i)$, for $i \in [\![0, n]\!]$, and n bounded linear operators $C_j \in \mathcal{B}(H_j, H_{j-1})$, for $j \in [\![1, n]\!]$, we extend Definition 3.2.1 as follows:

(i) $\forall\, k \in \mathbb{N}^*$, $\left[W_0^{-k}\right] := W_0^{-k}$;

(ii) $\forall\, k \in \mathbb{N}^*$, $\left[W_0^{-k}, \ldots, W_n^{-k}\right]_{(C_0,\ldots,C_{n-1})} := (-1)^n \sum_{\substack{i_0+\cdots+i_n = k-n \\ i_0,\ldots,i_n \geq 1}} W_0^{-i_0} C_0 W_1^{-i_1} \ldots C_{n-1} W_n^{-i_n}$.

Remark 3.2.12. In view of Proposition 3.1.8, if $W_0, W_1, \ldots, W_n, C_0, \ldots, C_{n-1}$ are scalars, we recover again the *usual* divided differences.

Definition 3.2.13. Let $r = \frac{p}{q} \in \mathbb{C}(X)$, and assume $0 \notin q(\sigma(W_i))$, for all $i \in [\![0, n]\!]$. If the complex partial fraction decomposition of r is $r(X) = e(X) + \sum_{j=1}^{r} \sum_{k_j=1}^{d_j} (X - \alpha_j)^{-k_j}$ (with $e \in \mathbb{C}[X]$), then, we define (with the above notation):

$$[r(W_0),\dots,r(W_n)]_{(C_0,\dots,C_{n-1})} = [p(W_0),\dots,p(W_n)]_{(C_0,\dots,C_{n-1})} + \sum_{j=1}^{r}\sum_{k_j=1}^{d_j}\left[(W_0-\alpha_j \mathrm{Id})^{-k_j},\dots,(W_n-\alpha_j\mathrm{Id})^{-k_j}\right]_{(C_0,\dots,C_{n-1})}.$$

Then, the analog of Proposition 3.2.3 and Proposition 3.2.4 are respectively:

Proposition 3.2.14.

$$\left[W_0^{-k},\dots,W_n^{-k}\right]_{(C_0,\dots,C_{n-1})} = -\sum_{j=1}^{k}\left[W_0^{-j},\dots,W_{n-1}^{-j}\right]_{(C_0,\dots,C_{n-2})} C_{n-1} W_n^{j-k-1}.$$

Proposition 3.2.15.

$$\begin{aligned}&W_0[r(W_0),\dots,r(W_n)]_{(C_0,\dots,C_{n-1})} - [r(W_0),\dots,r(W_n)]_{(C_0,\dots,C_{n-1})}W_n\\ &\quad = [r(W_0),\dots,r(W_{n-1})]_{(C_0,\dots,C_{n-2})}C_{n-1} - C_0[r(W_1),\dots,r(W_n)]_{(C_1,\dots,C_{n-1})}.\end{aligned}$$

In particular,

$$W_0[r(W_0),r(W_1)]_{(C_0)} - [r(W_0),r(W_1)]_{(C_0)}W_1 = r(W_0)C_0 - C_0 r(W_1).$$

Now, we can get an analog of Theorem 3.2.6:

Theorem 3.2.16. *Let $n \geq 2$, let $H_1,\dots,H_n$ be Hilbert spaces, and let*

$$\widetilde{T}_n = \begin{pmatrix} W_1 & C_1^{(1)} & \cdots & \cdots & C_1^{(n-1)} \\ 0 & W_2 & C_2^{(1)} & \cdots & C_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & C_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & W_n \end{pmatrix} : \bigoplus_{i=1}^{n} H_i \to \bigoplus_{i=1}^{n} H_i$$

be a bounded linear operator. Let $r = \frac{p}{q} \in \mathbb{C}(X)$ such that $0 \notin q(\sigma(W_i))$, for all $i \in [\![1,n]\!]$.

Then, we have :

$$r\left(\widetilde{T}_n\right) = \begin{pmatrix} r(W_1) & D_1^{(1)} & \cdots & \cdots & D_1^{(n-1)} \\ 0 & r(W_2) & D_2^{(1)} & \cdots & D_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & D_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & r(W_n) \end{pmatrix},$$

$$\text{where } D_i^{(j)} = \sum_{\substack{1\leq l\leq j \\ i=k_0<\cdots<k_l=j+i}} \left[r\left(W_{k_0}\right),\dots,r\left(W_{k_l}\right)\right]_{\left(C_{k_0}^{(k_1-k_0)},\dots,C_{k_{l-1}}^{(k_l-k_{l-1})}\right)}.$$

Proof. First, considering the partial fraction decomposition of r, observe that as we have already proved this formula for polynomials, we only need to prove this formula for $\left(\widetilde{T_n}\right)^{-k}$, where $k \in \mathbb{N}^*$.

The main idea of this proof will be to combine the arguments developed in the proof of Theorem 3.2.6 with the following identity:

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}CB^{-1} \\ 0 & B^{-1} \end{bmatrix}, \text{ if } A \text{ and } B \text{ are invertible operators} \tag{3.2}$$

As for the proof of Theorem 3.2.6, we proceed by induction on n.

- For $n = 2$, we use successively Theorem 3.2.6 and (3.2) to get

$$\widetilde{T_2}^k = \begin{pmatrix} W_1^k & \left[W_1^k, W_2^k\right]_{\left(C_1^{(1)}\right)} \\ 0 & W_1^k \end{pmatrix},$$

  and, then,

$$\left(\widetilde{T_2}\right)^{-k} = \begin{pmatrix} W_1^{-k} & -W_1^{-k}\left[W_1^k, W_2^k\right]_{\left(C_1^{(1)}\right)} W_2^{-k} \\ 0 & W_2^{-k} \end{pmatrix} = \begin{pmatrix} W_1^{-k} & \left[W_1^{-k}, W_2^{-k}\right]_{\left(C_1^{(1)}\right)} \\ 0 & W_2^{-k} \end{pmatrix}.$$

- The inductive step then works exactly as in the proof of Theorem 3.2.6 (using Proposition 3.2.14 instead of Proposition 3.2.3).

□

Corollary 3.2.17. *Let*

$$\widetilde{T_n} := \begin{pmatrix} W_1 & C_1 & 0 & \cdots & 0 \\ 0 & W_2 & C_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & C_{n-1} \\ 0 & \cdots & \cdots & 0 & W_n \end{pmatrix} : \bigoplus_{i=1}^n H_i \to \bigoplus_{i=1}^n H_i$$

and let $r = \frac{p}{q} \in \mathbb{C}(X)$ such that $0 \notin q(\sigma(W_i))$, for all $i \in [\![1, n]\!]$.

Then, we have:

$$r\left(\widetilde{T_n}\right) = \begin{pmatrix} r(W_1) & D_1^{(1)} & \cdots & \cdots & D_1^{(n-1)} \\ 0 & r(W_2) & D_2^{(1)} & \cdots & D_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & D_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & r(W_n) \end{pmatrix},$$

where $D_i^{(j)} = [r(W_i), \ldots, r(W_{i+j})]_{(C_i, \ldots, C_{i+j-1})}$.

Corollary 3.2.18 (Leibniz's formula)**.** *Let $n \geq 1$, let $H_0, \ldots, H_n$ be Hilbert spaces. Given two rational fractions $r = \frac{p}{q}, s = \frac{p'}{q'} \in \mathbb{C}(X)$, given bounded linear operators $W_i \in \mathcal{B}(H_i)$ such that $0 \notin q(\sigma(W_i)) \cup$*

$q'(\sigma(W_i))$, for $i \in [\![0,n]\!]$, and $C_j \in \mathcal{B}\left(H_{j+1}, H_j\right)$, for $j \in [\![0,n-1]\!]$, we have:

$$[(r\cdot s)(W_0),\dots,(r\cdot s)(W_n)]_{(C_0,\dots,C_{n-1})} = \sum_{j=0}^{n} \left[r(W_0),\dots,r(W_j)\right]_{(C_0,\dots,C_{j-1})} \cdot \left[s(W_j),\dots,s(W_n)\right]_{(C_j,\dots,C_{n-1})}.$$

Corollary 3.2.19. *Let*

$$T_n = \begin{pmatrix} \omega_1 & \alpha_1^{(1)} & \cdots & \cdots & \alpha_1^{(n-1)} \\ 0 & \omega_1 & \alpha_2^{(1)} & \cdots & \alpha_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \alpha_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & \omega_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{C}),$$

with $w_i \neq w_j$, for all $i \neq j$, and let $r = \frac{p}{q} \in \mathbb{C}(X)$ such that $q(\omega_i) \neq 0$, for all $i \in [\![1,n]\!]$.

Then, we have:

$$r(T_n) = \begin{pmatrix} r(\omega_1) & \beta_1^{(1)} & \cdots & \cdots & \beta_1^{(n-1)} \\ 0 & r(\omega_2) & \beta_2^{(1)} & \cdots & \beta_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \beta_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & r(\omega_n) \end{pmatrix},$$

$$\textit{where} \quad \beta_i^{(j)} = \sum_{\substack{1\leq l\leq j \\ i=k_0<\cdots<k_l=j+i}} \left(\prod_{s=0}^{l-1} \alpha_{k_s}^{(k_{s+1}-k_s)}\right)\left[r\left(\omega_{k_0}\right),\dots,r\left(\omega_{k_l}\right)\right].$$

3.2.4 Non-commutative divided differences for analytic functions

In this paragraph, we finally extend the notion of non-commutative divided differences to analytic functions.

In view of Proposition 3.1.9, we can define those non-commutative divided differences as follows:

Definition 3.2.20. Let $n \geq 1$, and let $H_0,\dots,H_n$ be Hilbert spaces. Given bounded linear operators $W_i \in \mathcal{B}(H_i)$, for $i \in [\![0,n]\!]$, and $C_j \in \mathcal{B}(H_j, H_{j-1})$, for $j \in [\![1,n]\!]$, and given a map f which is analytic in the neighbourhood of $\bigcup_{i=0}^{n} \sigma(W_i)$, we define:

(i) $[f(W_0)] := f(W_0) = \int_\Gamma f(\xi)(\xi \mathrm{Id} - W_0)^{-1}$, where Γ is a finite system of rectifiable Jordan curves, oriented in the positive sense, surrounding $\sigma(W_0)$;

(ii) $[f(W_0),\dots,f(W_n)]_{(C_0,\dots,C_{n-1})} := \int_\Gamma f(\xi)(\xi\mathrm{Id}-W_0)^{-1}C_0(\xi\mathrm{Id}-W_1)^{-1}\dots C_{n-1}(\xi\mathrm{Id}-W_n)^{-1}\mathrm{d}\xi$, where Γ is a finite system of rectifiable Jordan curves, oriented in the positive sense, surrounding $\bigcup_{i=0}^{n}\sigma(W_i)$.

As in the cases of polynomials and rational fractions, we have the characteristic identity:

Proposition 3.2.21. *With the above notation, we have:*

$$\begin{aligned} &W_0[f(W_0),\dots,f(W_n)]_{(C_0,\dots,C_{n-1})} - [f(W_0),\dots,f(W_n)]_{(C_0,\dots,C_{n-1})}W_n \\ &\quad = [f(W_0),\dots,f(W_{n-1})]_{(C_0,\dots,C_{n-2})}C_{n-1} - C_0[f(W_1),\dots,f(W_n)]_{(C_1,\dots,C_{n-1})}. \end{aligned}$$

In particular,

$$W_0[f(W_0),f(W_1)]_{(C_0)} - [f(W_0),f(W_1)]_{(C_0)}W_1 = f(W_0)C_0 - C_0 f(W_1).$$

This provides the following explicit functional calculus:

Theorem 3.2.22. *Let $n \geq 2$, let $H_1,\dots,H_n$ be Hilbert spaces, and let*

$$\widetilde{T_n} = \begin{pmatrix} W_1 & C_1^{(1)} & \cdots & \cdots & C_1^{(n-1)} \\ 0 & W_2 & C_2^{(1)} & \cdots & C_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & C_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & W_n \end{pmatrix} : \bigoplus_{i=1}^{n} H_i \to \bigoplus_{i=1}^{n} H_i$$

be a bounded linear operator. Let f be a map which is analytic in the neighborhood of $\bigcup_{i=0}^{n} \sigma(W_i)$.

Then, we have :

$$f\left(\widetilde{T_n}\right) = \begin{pmatrix} f(W_1) & D_1^{(1)} & \cdots & \cdots & D_1^{(n-1)} \\ 0 & f(W_2) & D_2^{(1)} & \cdots & D_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & D_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & f(W_n) \end{pmatrix},$$

$$\text{where} \quad D_i^{(j)} = \sum_{\substack{1\leq l\leq j \\ i=k_0<\cdots<k_l=j+i}} \left[f\left(W_{k_0}\right),\dots,f\left(W_{k_l}\right)\right]_{\left(C_{k_0}^{(k_1-k_0)},\dots,C_{k_{l-1}}^{(k_l-k_{l-1})}\right)}.$$

Remark 3.2.23. We have $\sigma\left(\widetilde{T_n}\right) \subset \bigcup_{i=1}^{n} \sigma(W_i)$. Indeed, for $\lambda \in \mathbb{C}$, we have:

$$\widetilde{T_n} = \begin{pmatrix} W_1 - \lambda \mathrm{Id} & C_1^{(1)} & \cdots & \cdots & C_1^{(n-1)} \\ 0 & W_2 - \lambda \mathrm{Id} & C_2^{(1)} & \cdots & C_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & C_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & W_n - \lambda \mathrm{Id} \end{pmatrix}.$$

If, for all $i \in [\![1,n]\!]$, $W_i - \lambda\mathrm{Id}$ is invertible, then, it follows from Section 3.2.3 (see Theorem 3.2.16) that $\widetilde{T_n} - \lambda\mathrm{Id}$ is invertible too.

Proof of Theorem 3.2.22. Let Γ be a finite system of rectifiable Jordan curves, oriented in the positive sense, surrounding $\bigcup_{i=1}^{n} \sigma(W_i)$. Then, by definition, $f\left(\widetilde{T_n}\right) = \frac{1}{2i\pi}\int_\Gamma f(\xi)\left(\xi\mathrm{Id} - \widetilde{T_n}\right)^{-1} \mathrm{d}\xi$. The results just follows by integrating the formula of $\left(\xi\mathrm{Id} - \widetilde{T_n}\right)^{-1}$ given by Theorem 3.2.16. $\square$

Corollary 3.2.24. *Let*

$$\widetilde{T_n} := \begin{pmatrix} W_1 & C_1 & 0 & \cdots & 0 \\ 0 & W_2 & C_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & C_{n-1} \\ 0 & \cdots & \cdots & 0 & W_n \end{pmatrix}$$

and let f be a function which is holomorphic in the neighborhood of $\bigcup_{i=1}^{n} \sigma(W_i)$.

Then, we have:

$$f\left(\widetilde{T_n}\right) = \begin{pmatrix} f(W_1) & D_1^{(1)} & \cdots & \cdots & D_1^{(n-1)} \\ 0 & f(W_2) & D_2^{(1)} & \cdots & D_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & D_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & f(W_n) \end{pmatrix},$$

where $D_i^{(j)} = [f(W_i), \ldots, f(W_{i+j})]_{(C_i, \ldots, C_{i+j-1})}$.

Corollary 3.2.25 (Leibniz's formula)**.** *Let $n \geq 1$, and let $H_0, \ldots, H_n$ be Hilbert spaces. Given bounded linear operators $W_i \in \mathcal{B}(H_i)$, for $i \in [\![0, n]\!]$, and $C_j \in \mathcal{B}\left(H_{j+1}, H_j\right)$, for $j \in [\![0, n-1]\!]$, and given two functions f, g which are holomorphic in the neighborhood of $\bigcup_{i=0}^{n} \sigma(W_i)$, we have:*

$$[(f \cdot g)(W_0), \ldots, (f \cdot g)(W_n)]_{(C_0, \ldots, C_{n-1})} = \sum_{j=0}^{n-1} \left[f(W_0), \ldots, f(W_j)\right]_{(C_0, \ldots, C_j)} \left[g(W_j), \ldots, g(W_n)\right]_{(C_{j+1}, \ldots, C_{n-1})}.$$

Corollary 3.2.26. *Let*

$$T_n = \begin{pmatrix} \omega_1 & \alpha_1^{(1)} & \cdots & \alpha_1^{(n-1)} \\ 0 & \omega_2 & \alpha_2^{(1)} \cdots & \alpha_2^{(n-2)} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \alpha_{n-1}^{(1)} \\ 0 & \cdots & 0 & \omega_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$$

such that $w_i \neq w_j$, for all $i \neq j$, and let f be a function holomorphic in the neighborhood of $\omega_1, \ldots, \omega_n$.

Then, we have:

$$f(T_n) = \begin{pmatrix} f(\omega_1) & \beta_1^{(1)} & \cdots & \cdots & \beta_1^{(n-1)} \\ 0 & f(\omega_2) & \beta_2^{(1)} & \cdots & \beta_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \beta_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & f(\omega_n) \end{pmatrix},$$

$$\text{where} \quad \beta_i^{(j)} = \sum_{\substack{1 \leq l \leq j \\ i = k_0 < \cdots < k_l = j+i}} \left(\prod_{s=0}^{l-1} \alpha_{k_s}^{(k_{s+1} - k_s)} \right) \left[f\left(\omega_{k_0}\right), \ldots, f\left(\omega_{k_l}\right) \right].$$

3.3 The equation $AX - XB = Y$

The *recursive relations* of Proposition 3.2.4, Proposition 3.2.15 and Proposition 3.2.21 lead us to focus on the equation

$$AX - XB = Y \tag{3.3}$$

where A, X, B, Y are operators (on a Banach space).

This equation has been studied *e.g.* in [9, 51]. As it plays an important role on studying functions of matrices, we give here some details about the solvability of this equation.

We start with the following existence and uniqueness result:

Theorem 3.3.1 (Sylvester-Rosenblum). *If A and B are operators acting on a Banach space such that $\sigma(A) \cap \sigma(B) = \emptyset$, then, the equation $AX - XB = Y$ has a unique solution X for every operator Y.*

Remark 3.3.2. A consequence of this result is that, with the notations introduced in the previous section, if $W_0, \dots W_n$ are operators such that $\sigma(W_i) \cap \sigma(W_j) = \emptyset$, for all $i \neq j$, then, the non-commutative divided differences are well-defined by the recursive relation:

$$[f(W_0)] = f(W_0);$$
$$W_0[f(W_0), \dots, f(W_n)]_{(C_0,\dots,C_{n-1})} - [f(W_0), \dots, f(W_n)]_{(C_0,\dots,C_{n-1})} W_n = [f(W_0), \dots, f(W_{n-1})]_{(C_1,\dots,C_{n-2})} C_{n-1} - C_0[f(W_1), \dots, f(W_n)]_{(C_1,\dots,C_{n-1})}.$$

Proof of Theorem 3.3.1. Let $M_{A,B} : X \mapsto AX - XB$, so that the Sylvester's equation becomes $M_{A,B}(X) = Y$. It is enough to prove that, if $\sigma(A) \cap \sigma(B) = \emptyset$, then, $M_{A,B}$ is invertible. In order to do this, we define $L_A : X \mapsto AX$ and $R_B : X \mapsto XB$, so that we have $M_{A,B} = L_A - R_B$. It is easy to see that $L_A R_B = R_B L_A$, that $\sigma(L_A) \subset \sigma(A)$ and that $\sigma(R_B) \subset \sigma(B)$.
Now, saying that $M_{A,B}$ is invertible is equivalent to say that $0 \notin \sigma(M_{A,B}) = \sigma(L_A - R_B)$. Thus, the theorem is a consequence of the following lemma:

Lemma 3.3.3. *If $\mathcal{A}$ and $\mathcal{B}$ are commuting operators, then, $\sigma(\mathcal{A} - \mathcal{B}) \subset \sigma(\mathcal{A}) - \sigma(\mathcal{B})$.*

Proof of Lemma 3.3.3.

- In the finite dimensional case, it is just a consequence of the simultaneous trigonalization of commuting endomorphisms: since $\mathcal{A}$ and $\mathcal{B}$ commute, there exists a basis in which $\mathcal{A}$ and $\mathcal{B}$ are both upper-triangular. The results follows from the fact that the spectrum of an upper-triangular matrix is the set of the diagonal entries.
- In the general case, we need to use the Gelfand transformation (see Section 1.1.3). Let $\mathcal{C}$ be the maximal commutative algebra containing both $\mathcal{A}$ and $\mathcal{B}$. Observe that for any operator $\mathcal{T} \in \mathcal{C}$, $\sigma(\mathcal{T}) = \sigma_{\mathcal{C}}(\mathcal{T})$. Now, we write:

$$\begin{aligned}
\sigma(\mathcal{A} - \mathcal{B}) &= \sigma_{\mathcal{C}}(\mathcal{A} - \mathcal{B}) \\
&= \left\{\chi(\mathcal{A} - \mathcal{B}) : \chi \in \widehat{\mathcal{C}}\right\} \\
&= \left\{\chi(\mathcal{A}) - \chi(\mathcal{B}) : \chi \in \widehat{\mathcal{C}}\right\} \\
&\subset \left\{\chi(\mathcal{A}) : \chi \in \widehat{\mathcal{C}}\right\} - \left\{\chi'(\mathcal{B}) : \chi' \in \widehat{\mathcal{C}}\right\} \\
&= \sigma_{\mathcal{C}}(\mathcal{A}) - \sigma_{\mathcal{C}}(\mathcal{B}) \\
&= \sigma(\mathcal{A}) - \sigma(\mathcal{B}).
\end{aligned}$$

□

This completes the proof of Theorem 3.3.1. □

Now, let us focus on the expression of this unique solution. We focus on two expressions of the solution:

Theorem 3.3.4. *Let A and B be two operators such that $\sigma(B) \subset D(0,\rho)$ and $\sigma(A) \subset \mathbb{C} \setminus \overline{D(0,\rho)}$, for some $\rho > 0$. Then, the solution of the equation $AX - XB = Y$ is*

$$X = \sum_{n=0}^{\infty} A^{-n-1} Y B^n.$$

Proof. The only thing that need to be proved is that the series is convergent (if it is the case, then, it is obvious that this sum is solution of the Sylvester equation). First of all, observe that as $\sigma(A) \subset \mathbb{C} \setminus \overline{D(0,\rho)}$, it makes sense to consider negative powers of A. Now, let $\rho_1 < \rho < \rho_2$ such that $\sigma(B) \subset D(0,\rho_1)$ and $\sigma(A) \subset \mathbb{C} \setminus \overline{D(0,\rho_2)}$. Then, $\sigma(A^{-1}) \subset D(0,\rho_2^{-1})$. By the spectral radius formula (Proposition 1.1.7), there exists an integer $n_0 \geq 0$ such that, for all $n \geq n_0$, $\|B^n\| < \rho_1^n$ and $\|A^{-n}\| < \rho_2^{-n}$. Thus, we have $\|A^{-n-1} Y B^n\| < \left(\frac{\rho_1}{\rho_2}\right)^n \|A^{-1} Y\|$, and the series converges normally. □

Remark 3.3.5. If a and b are non-zero scalars, with $b < a$, then, $\sum_{n=0}^{\infty} a^{-n-1} y b^n = \frac{y}{a} \sum_{n=0}^{\infty} \left(\frac{b}{a}\right)^n = \frac{y}{a-b}$.

The next theorem gives an expression of the solution whenever $\sigma(A)$ and $\sigma(B)$ are disjoint, without any more special assumption:

Theorem 3.3.6 (Rosenblum)**.** *Let Γ be a finite system of rectifiable Jordan curves, oriented in the positive sense, with total winding number 1 around $\sigma(A)$ and 0 around $\sigma(B)$. Then, the solution of the equation $AX - XB = Y$ can be expressed as*

$$X = \frac{1}{2i\pi} \int_\Gamma (A - \xi \mathrm{Id})^{-1} Y (B - \xi \mathrm{Id})^{-1} \mathrm{d}\xi.$$

Proof. If $AX - XB = Y$, then, for every complex number ξ, we have $(A - \xi\mathrm{Id})X - X(B - \xi\mathrm{Id}) = Y$. If $A - \xi\mathrm{Id}$ and $B - \xi\mathrm{Id}$ are invertible, this gives:

$$X(B - \xi\mathrm{Id})^{-1} - (A - \xi\mathrm{Id})^{-1} X = (A - \xi\mathrm{Id})^{-1} Y (B - \xi\mathrm{Id})^{-1}.$$

The result follows by integrating this equality over Γ, and noting that $\int_\Gamma (B - \xi\mathrm{Id})^{-1} \mathrm{d}\xi = 0$ and $\int_\Gamma (A - \xi\mathrm{Id})^{-1} \mathrm{d}\xi = -2i\pi\mathrm{Id}$. □

Remark 3.3.7. If a and b are two scalars such that $a \neq b$, and if Γ is a loop with winding number 1 around a and 0 around b, then, the residue theorem implies that $\frac{1}{2i\pi} \int_\Gamma \frac{y}{(a-\xi)(b-\xi)} \mathrm{d}\xi = \frac{y}{a-b}$.

A noteworthy application of the study of this equation is the diagonalization of certain matrices with operator coefficients:

Lemma 3.3.8. *Let A and B be two operators such that $\sigma(A) \cap \sigma(B) = \emptyset$. Then, for any operator C, the matrices $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ and $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ are similar.*

More precisely, we have:

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \begin{pmatrix} \mathrm{Id} & -X \\ 0 & \mathrm{Id} \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} \mathrm{Id} & X \\ 0 & \mathrm{Id} \end{pmatrix}$$

where X is the unique solution of the equation $AX - XB = C$.

An easy induction gives the following theorem:

Theorem 3.3.9. *Let $W_1, \ldots, W_n$ be operators such that $\sigma(W_i) \cap \sigma(W_j) = \emptyset$, for all $i \neq j$. Then, for any operators $\left\{C_i^{(j)}\right\}_{\substack{1\le j\le n-1 \\ 1\le i\le n+1-j}}$, the matrices*

$$\widetilde{T_n} := \begin{pmatrix} W_1 & C_1^{(1)} & \cdots & \cdots & C_1^{(n-1)} \\ 0 & W_2 & C_2^{(1)} & \cdots & C_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & C_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & W_n \end{pmatrix}$$

and

$$\widetilde{D_n} := \begin{pmatrix} W_1 & 0 & \cdots & 0 \\ 0 & W_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & W_n \end{pmatrix}$$

are similar.

More precisely, we have:

$$\widetilde{T_n} = \begin{pmatrix} \mathrm{Id} & -X_1^{(1)} & \cdots & \cdots & -X_1^{(n-1)} \\ 0 & \mathrm{Id} & -X_2^{(1)} & \cdots & -X_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & -X_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & \mathrm{Id} \end{pmatrix} \begin{pmatrix} W_1 & 0 & \cdots & 0 \\ 0 & W_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & W_n \end{pmatrix} \begin{pmatrix} \mathrm{Id} & Y_1^{(1)} & \cdots & \cdots & Y_1^{(n-1)} \\ 0 & \mathrm{Id} & Y_2^{(1)} & \cdots & Y_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & Y_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & \mathrm{Id} \end{pmatrix}$$

where, for all $j \in [\![1, n-1]\!]$, for all $i \in [\![1, n+1-j]\!]$, $X_i^{(j)}$ is the unique solution of the equation

$$W_i X_i^{(j)} - X_i^{(j)} W_{i+j} = C_i^{(j)} - \sum_{k=1}^{j-1} C_i^{(k)} X_{k+i}^{(j-k)} \tag{3.4}$$

and

$$Y_i^{(j)} = \sum_{\substack{1\le l\le j \\ i=k_0<\cdots<k_l=j+i}} X_{k_0}^{(k_1-k_0)} \ldots X_{k_{l-1}}^{(k_l-k_{l-1})} \tag{3.5}$$

Remark 3.3.10. In order to use this theorem, we first need to determine all the coefficients $X_i^{(1)}$, then all the coefficients $X_i^{(2)}$,... until the coefficient $X_1^{(n-1)}$.

Corollary 3.3.11. *With the above notation, if f is a map which is holomorphic in the neighborhood of*

$\bigcup_{i=1}^{n} \sigma(W_i)$, *we have:*

$$f\left(\widetilde{T_n}\right) = \begin{pmatrix} \mathrm{Id} & -X_1^{(1)} & \cdots & \cdots & -X_1^{(n-1)} \\ 0 & \mathrm{Id} & -X_2^{(1)} & \cdots & -X_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & -X_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & \mathrm{Id} \end{pmatrix} \begin{pmatrix} f(W_1) & 0 & \cdots & 0 \\ 0 & f(W_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & f(W_n) \end{pmatrix} \begin{pmatrix} \mathrm{Id} & Y_1^{(1)} & \cdots & \cdots & Y_1^{(n-1)} \\ 0 & \mathrm{Id} & Y_2^{(1)} & \cdots & Y_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & Y_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & \mathrm{Id} \end{pmatrix}.$$

Another interesting application of Theorem 3.3.9 is to provide an alternative of Theorem 2.3.18:

Corollary 3.3.12. *Let $\omega_1, \ldots, \omega_n \in \mathbb{D}$, and let*

$$M_n := \begin{pmatrix} \omega_1 & \alpha_1^{(1)} & \alpha_1^{(2)} & \cdots & \alpha_1^{(n-1)} \\ 0 & \omega_2 & \alpha_2^{(1)} & \alpha_2^{(2)} & \vdots \\ \vdots & \ddots & \ddots & \ddots & \alpha_{n-2}^{(2)} \\ \vdots & & \ddots & \ddots & \alpha_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & \omega_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{C}),$$

with

$$\alpha_i^{(j)} = \prod_{k=i+1}^{j+i-1} (-\overline{\omega_k}) \sqrt{1-|w_i|^2} \sqrt{1-|\omega_{i+j}|^2}, \quad \textit{for all } j \in [\![1,n]\!], \; \textit{for all } i \in [\![n-j,n]\!].$$

Then, we have:

$$M_n = \begin{pmatrix} 1 & -x_1^{(1)} & \cdots & \cdots & -x_1^{(n-1)} \\ 0 & 1 & -x_2^{(1)} & \cdots & -x_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & -x_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \begin{pmatrix} \omega_1 & 0 & \cdots & 0 \\ 0 & \omega_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \omega_n \end{pmatrix} \begin{pmatrix} 1 & y_1^{(1)} & \cdots & \cdots & y_1^{(n-1)} \\ 0 & 1 & y_2^{(1)} & \cdots & y_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & y_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}$$

with

$$x_i^{(j)} = \frac{\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+j}|^2}}{\omega_i - \omega_{i+j}} \left(\prod_{k=i+1}^{i+j-1} \frac{1-\overline{\omega_k}\omega_{i+j}}{\omega_{i+j}-\omega_k} \right) \tag{3.6}$$

$$y_i^{(j)} = \frac{\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+j}|^2}}{\omega_i - \omega_{i+j}} \left(\prod_{k=i+1}^{i+j-1} \frac{1-\overline{\omega_k}\omega_{i+j}}{\omega_i-\omega_k} \right) \tag{3.7}$$

for all $1 \leq j \leq n$, $n-j \leq i \leq n$. Moreover,

$$\begin{pmatrix} 1 & -x_1^{(1)} & \cdots & \cdots & -x_1^{(n-1)} \\ 0 & 1 & -x_2^{(1)} & \cdots & -x_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & -x_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & y_1^{(1)} & \cdots & \cdots & y_1^{(n-1)} \\ 0 & 1 & y_2^{(1)} & \cdots & y_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & y_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Proof. • For the coefficients $x_i^{(j)}$, we proceed by induction on j:

– For $j = 1$, the coefficients $x_i^{(1)}$ are (respectively) the solutions of the equations

$$\omega_i x_i^{(1)} - x_i^{(1)}\omega_{i+1} = \alpha_i^{(1)},$$

which gives

$$x_i^{(1)} = \frac{\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+1}|^2}}{\omega_i - \omega_{i+1}}.$$

– Now, let $j \in [\![2, n]\!]$, and assume that

$$x_i^{(m)} = \frac{\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+m}|^2}}{\omega_i - \omega_{i+m}}\left(\prod_{k=i+1}^{m+i-1}\frac{1-\overline{\omega_k}\omega_{i+m}}{\omega_{i+m}-\omega_k}\right), \quad \text{for all } m \in [\![1, j-1]\!], \text{ for all } i \in [\![n-m, n]\!].$$

Then, for all $i \in [\![n-j, n]\!]$, we have:

$$\begin{aligned}
\omega_i x_i^{(j)} - x_i^{(j)}\omega_{i+j} &= \alpha_i^{(j)} - \sum_{l=1}^{j-1}\alpha_i^{(l)} x_{i+l}^{(j-l)} \\
&= \left(\prod_{k=i+1}^{i+j-1} -\overline{\omega_k}\right)\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+j}|^2} \\
&\quad + \sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+j}|^2}\sum_{l=1}^{j-1}\left\{\left(\prod_{k=i+1}^{i+l-1} -\overline{\omega_k}\right)\left(\frac{1-|\omega_{i+l}|^2}{\omega_{i+j}-\omega_{i+l}}\right)\left(\prod_{k=i+l+1}^{i+j-1}\frac{1-\overline{\omega_k}\omega_{i+j}}{\omega_{i+j}-\omega_k}\right)\right\} \\
&= \frac{\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+j}|^2}\cdot N_{i,j}}{\prod_{k=i+1}^{i+j-1}\left(\omega_{i+j}-\omega_k\right)},
\end{aligned}$$

where

$$\begin{aligned}
N_{i,j} &:= \left(\prod_{k=i+1}^{i+j-1} -\overline{\omega_k}\right)\prod_{q=i+1}^{i+j-1}\left(\omega_{i+j}-\omega_q\right) \\
&\quad + \sum_{l=1}^{j-1}(1-|\omega_{i+l}|^2)\left(\prod_{k=i+1}^{i+l-1} -\overline{\omega_k}\right)\prod_{p=i+l+1}^{i+j-1}\left(1-\overline{\omega_p}\omega_{i+j}\right)\prod_{q=i+1}^{i+l-1}\left(\omega_{i+j}-\omega_q\right) \\
&= \left(\prod_{k=i+1}^{i+j-2} -\overline{\omega_k}\right)\prod_{q=i+1}^{i+j-2}\left(\omega_{i+j}-\omega_q\right)\left[-\overline{\omega_{i+j-1}}(\omega_{i+j}-\omega_{i+j-1}) + 1 - |\omega_{i+j-1}|^2\right] \\
&\quad + \sum_{l=1}^{j-2}(1-|\omega_{i+l}|^2)\left(\prod_{k=i+1}^{i+l-1} -\overline{\omega_k}\right)\prod_{p=i+l+1}^{i+j-1}\left(1-\overline{\omega_p}\omega_{i+j}\right)\prod_{q=i+1}^{i+l-1}\left(\omega_{i+j}-\omega_q\right)
\end{aligned}$$

$$= (1-\overline{\omega_{i+j-1}}\omega_{i+j})\left(\prod_{k=i+1}^{i+j-2} -\overline{\omega_k}\right)\prod_{q=i+1}^{i+j-2}\left(\omega_{i+j}-\omega_q\right)$$
$$+\sum_{l=1}^{j-2}(1-|\omega_{i+l}|^2)\left(\prod_{k=i+1}^{i+l-1} -\overline{\omega_k}\right)\prod_{p=i+l+1}^{i+j-1}\left(1-\overline{\omega_p}\omega_{i+j}\right)\prod_{q=i+1}^{i+l-1}\left(\omega_{i+j}-\omega_q\right).$$

Iterating this process, we get

$$N_{i,j} = \prod_{k=i+1}^{i+j-1} 1-\overline{\omega_k}\omega_{i+j}$$

and, thus,

$$x_i^{(j)} = \frac{\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+j}|^2}}{\omega_i-\omega_{i+j}}\left(\prod_{k=i+1}^{j+i-1}\frac{1-\overline{\omega_k}\omega_{i+j}}{\omega_{i+j}-\omega_k}\right).$$

- Now, if we denote by

$$X_n := \begin{pmatrix} 1 & -x_1^{(1)} & \cdots & \cdots & -x_1^{(n-1)} \\ 0 & 1 & -x_2^{(1)} & \cdots & -x_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & -x_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix},$$

observe that $X_n = \begin{bmatrix} X_{n-1} & C_n \\ 0 & 1 \end{bmatrix}$, where $C_n = \begin{bmatrix} -x_1^{(n-1)} & -x_2^{(n-2)} & \dots & -x_{n-1}^{(1)} \end{bmatrix}^\mathsf{T}$. Thus, we have $X_n^{-1} = \begin{bmatrix} X_{n-1}^{-1} & -X_{n-1}^{-1}C_n \\ 0 & 1 \end{bmatrix}$, and we can prove by induction on n that

$$X_n^{-1} = \begin{pmatrix} 1 & y_1^{(1)} & \cdots & \cdots & y_1^{(n-1)} \\ 0 & 1 & y_2^{(1)} & \cdots & y_2^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & y_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}$$

where

$$y_i^{(j)} = \frac{\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+j}|^2}}{\omega_i-\omega_{i+j}}\left(\prod_{k=i+1}^{i+j-1}\frac{1-\overline{\omega_k}\omega_{i+j}}{\omega_i-\omega_k}\right), \text{ for all } j\in[\![1,n]\!], \text{ for all } i\in[\![n-j,n]\!].$$

$\square$

Chapter 4

Contractive matrices

The main goal of this chapter will be to generalize Lemma 1.2.14 – which gives a characterization of contractive 2×2 matrices – to matrices of bigger size.

Observe that in view of Schur's decomposition theorem – which states that every square matrix is unitary equivalent to an upper-triangular matrix – it makes sense to restrict ourselves to that case.

For instance, let

$$T_3 = \begin{pmatrix} \omega_1 & \alpha_1 & \beta \\ 0 & \omega_2 & \alpha_2 \\ 0 & 0 & \omega_3 \end{pmatrix} \in \mathcal{M}_3(\mathbb{C}).$$

Under which condition is T_3 a contraction ?

In essence, following the approach used for 2×2 matrices, one can try to calculate the operator norm of a 3×3 matrix acting on the Euclidean space $\mathbb{C}^3$ using the formula:

$$\|T\|^2 = \|T^*T\| = r(T^*T) = \sup\{|\lambda| : \det(T^*T - \lambda \mathrm{Id}) = 0\}.$$

This computation of the operator norm $\|T\|$, the largest singular value of T, leads to an equation of degree 3 with no trivial solution. The expression of the solutions – involving Cardans's formulas – are quite intricate and, therefore, the criterion derived from this observation holds limited practical interest. We follow here a different approach based on a result about completion of matrices going back to Parrott (see [44, 23], [66, Theorem 12.22] and [6, 17]).

4.1 Parrott's theorem

In this section, we prove a slightly more precise version of Parrott's theorem. We adopt the approach presented by Davis, Kahan, and Weinberger in [17], noticing that the ideas developed in this proof enables in fact to be more precise regarding the selection of solutions with minimal norms.

First of all, let us recall that the defect operator of a contraction T is given by $D_T = (\mathrm{Id} - T^*T)^{1/2}$.

Theorem 4.1.1 (Parrott). *Let H_1, H_2, K_1, K_2 be Hilbert spaces, and assume that the operators* $\begin{bmatrix} A \\ C \end{bmatrix} \in \mathcal{B}(H_1, K_1 \oplus K_2)$ *and* $\begin{bmatrix} C & D \end{bmatrix} \in \mathcal{B}(H_1 \oplus H_2, K_2)$ *are contractions.*

Then,

$$T = \begin{bmatrix} A & B \\ C & D \end{bmatrix} : H_1 \oplus H_2 \to K_1 \oplus K_2$$

is a contraction if and only if there exists a contraction $W \in \mathcal{B}(H_2, K_1)$ such that:

$$B = D_{Z^*} W D_Y - Z C^* Y,$$

where $Z \in \mathcal{B}(H_1, K_1)$ and $Y \in \mathcal{B}(H_2, K_2)$ are contractions such that $D = D_{C^*} Y$ and $A = Z D_C$.

Moreover,

(i) Y and Z can be chosen to be (respectively) Y_0 and Z_0, the solutions of minimal operator norm among all solutions of the operator equations $D = D_{C^*} Y$ and $A = Z D_C$;

(ii) If T is a contraction, there exists a unique contraction W_0 such that:

$$B = D_{Z_0^*} W_0 D_{Y_0} - Z_0 C^* Y_0 \ \text{ and } \ Im\left(D_{Z_0^*}\right)^{\perp} \subset Ker(W_0^*).$$

This operator satisfies:

$$\|W_0\| = \inf\{\|W\| : B = D_{Z_0^*} W D_{Y_0} - Z_0 C^* Y_0\}.$$

We shall call Y_0 and Z_0 the minimal solutions *and we shall refer to W_0 as the* minimal solution *of the equation*

$$B = D_{Z_0^*} W D_{Y_0} - Z_0 C^* Y_0.$$

In order to prove this result, we start with the following lemma. Recall that if S and T are two operators on a Hilbert space H, the notation $S \le T$ means that $T - S$ is a positive operator. If $T = \lambda \mathrm{Id}$, for some $\lambda \in \mathbb{C}$, we may simply write $S \le \lambda$.

Lemma 4.1.2 (Douglas [19]). *Let L, M_1, M_2 be Hilbert spaces. Suppose that $A \in \mathcal{B}(L, M_1)$, $B \in \mathcal{B}(L, M_2)$ and $c \ge 0$. Then, $B^* B \le c^2 A^* A$ if and only if there exists $C \in \mathcal{B}(M_1, M_2)$ such that:*

$$\begin{cases} B = CA \\ \|C\| \le c \end{cases} . \tag{4.1}$$

Moreover, if it is the case, there exists a unique operator C_0 satisfying (4.1) such that $Im(A)^{\perp} \subset Ker(C_0)$. The operator C_0 satisfies:

$$\|C_0\|^2 = \inf\{\|C\|^2 : C \text{ satisfies } (4.1)\} = \inf\{\mu \ge 0 : B^* B \le \mu A^* A\},$$

and will thus be referred as the minimal solution *of the equation $B = CA$.*

Proof.

- Assume first that there exists an operator $C \in \mathcal{B}(M_1, M_2)$ such that $B = CA$ and $\|C\| \leq c$. We have:
$$\|C\| \leq c \iff \langle Cx, Cx\rangle \leq c^2\langle x, x\rangle,\ \forall\, x \in M_1 \iff c^2\mathrm{Id} - C^*C \geq 0.$$
Moreover, the equality $B = CA$ implies that $c^2A^*A - B^*B = A^*(c^2\mathrm{Id} - C^*C)A \geq 0$, *i.e.* $B^*B \leq c^2A^*A$, using the following elementary fact:

  Fact 1. Let H, K be two Hilbert spaces, $T \in \mathcal{B}(H,K)$ and $P \in \mathcal{B}(K)$ be a positive operator. Then, $T^*PT \geq 0$.

- Assume now that $B^*B \leq c^2A^*A$. First, note that:
$$B^*B \leq c^2A^*A \iff \langle Bx, Bx\rangle \leq c^2\,\langle Ax, Ax\rangle\ \forall\, x \in L \iff \|Bx\| \leq c\|Ax\|,\ \forall\, x \in L.$$
Now, let $C : \mathrm{Im}(A) \to M_2$, $Ax \mapsto Bx$.

  C is well-defined, as if $Ax = Ay$, then, $\|Bx - By\| \leq c\|Ax - Ay\| = 0$, and so $Bx = By$.

  Moreover, C is continuous with $\|C\| \leq c$. Now, we extend C by continuity on $\overline{\mathrm{Im}(A)}$ and we set, for instance, $C = 0$ on $\mathrm{Im}(A)^{\perp}$. We still have $B = CA$ and $\|C\| \leq c$, and C is now defined on all $M_1 = \overline{\mathrm{Im}(A)} \oplus \mathrm{Im}(A)^{\perp}$.

  In the previous construction, the only degree of freedom that we have is when we define C on $\mathrm{Im}(A)^{\perp}$. Thus, if we require that $C = 0$ on $\mathrm{Im}(A)^{\perp}$, C is unique. We denote by C_0 this operator. Moreover, this requirement minimizes the norm of C. Hence, for every $\mu \geq 0$ such that $B^*B \leq \mu A^*A$, we have $\|C_0\|^2 \leq \mu$, which means that
$$\|C_0\|^2 \leq \inf\{\mu \geq 0 / B^*B \leq \mu A^*A\}. \tag{4.2}$$
Finally, we have seen in the beginning of the proof that if $\|C_0\|^2 \leq \lambda$, then, the equality $B = C_0A$ implies that $B^*B \leq \lambda A^*A$. Thus, we cannot have a strict inequality in (4.2).

□

From this lemma, we deduce the following result about column matrices.

Proposition 4.1.3. *Let H, K_1, K_2 be Hilbert spaces. Suppose that $A \in \mathcal{B}(H,K_1)$ and $B \in \mathcal{B}(H,K_2)$ are contractions. Then, $\begin{bmatrix} A \\ B \end{bmatrix} : H_1 \to K_1 \oplus K_2$ is a contraction if and only if there exists a contraction $V \in \mathcal{B}(H,K_1)$ such that $A = VD_B$.*

Moreover, if it is the case, there exists a unique contraction V_0 such that $A = V_0D_B$ and $Im(D_B)^{\perp} \subset Ker(V_0)$. This contraction V_0 satisfies
$$\|V_0\| = \inf\{\|V\| \,:\, A = VD_B\},$$
and will thus be referred as the minimal solution of the equation $A = VD_B$.

Proof. The column matrix $\begin{bmatrix} A \\ B \end{bmatrix}$ is a contraction if and only if $A^*A \leq \mathrm{Id} - B^*B = D_B^*D_B$. Using Lemma 4.1.2, we obtain $A = VD_B$ with $\|V\| \leq 1$. □

Corollary 4.1.4. *Let H_1, H_2, K be Hilbert spaces. Suppose that $A \in \mathcal{B}(H_1, K)$ and $B \in \mathcal{B}(H_2, K)$ are contractions. Then, $\begin{bmatrix} A & B \end{bmatrix} : H_1 \oplus H_2 \to K$ is a contraction if and only if there exists a contraction $V \in \mathcal{B}(H_1, K)$ such that $A = D_{B^*} V$.*

Moreover, if it is the case, there exists a unique contraction V_0 such that $A = D_{B^*} V_0$ and $Im(D_{B^*})^\perp \subset Ker(V_0^*)$.

This contraction V_0 satisfies

$$\|V_0\| = \inf\{\|V\| : A = D_{B^*} V\},$$

and will thus be referred as the minimal solution *of the equation $A = D_{B^*} V$.*

Proof. Observe that $\begin{bmatrix} A & B \end{bmatrix}$ is a contraction if and only if $\begin{bmatrix} A & B \end{bmatrix}^* = \begin{bmatrix} A^* \\ B^* \end{bmatrix}$ is a contraction, and then apply Proposition 4.1.3. □

The following technical corollary will play a crucial role in the proof of Parrott's theorem:

Corollary 4.1.5. *Let H, K_1, K_2, A, B be as in Proposition 4.1.3, and let $U \in \mathcal{B}(H)$ be an arbitrary (but fixed) isometry. Then, $\begin{bmatrix} A \\ B \end{bmatrix} : H_1 \to K_1 \oplus K_2$ is a contraction if and only if there exists a contraction $V \in \mathcal{B}(H, K_1)$ such that $A = VUD_B$.*

Moreover, if it is the case, there exists a unique contraction V_0 such that $A = V_0 U D_B$ and $Im(UD_B)^\perp \subset Ker(V_0)$. This contraction V_0 satisfies

$$\|V_0\| = \inf\{\|V\| : A = VUD_B\},$$

and will thus be referred as the minimal solution *of the equation $A = VUD_B$.*

Proof. It is enough to prove the sufficiency part. By Proposition 4.1.3, if $\begin{bmatrix} A \\ B \end{bmatrix}$ is a contraction, there exists a contraction $W \in \mathcal{B}(H, K_1)$ such that $A = WD_B$. Moreover, W can be chosen such that $W = 0$ on $\mathrm{Im}(D_B)^\perp$ (and in this case, the minimal solution W_0 is unique).

Now, let $V = WU^*$. As U is an isometry, it is easy to see that V is a contraction and that $VUD_B = WD_B = A$. Moreover, $V = 0$ on $\mathrm{Im}(UD_B)^\perp$. Indeed, let $x \in \mathrm{Im}(UD_B)^\perp$. For all $x' \in H$, $\langle x, UD_B x' \rangle = 0$, which can be rewritten $\langle U^* x, D_B x' \rangle = 0$. Thus, for $x \in \mathrm{Im}(UD_B)^\perp$, $U^* x \in \mathrm{Im}(D_B)^\perp$ and, then, $Vx = WU^* x = 0$ (by minimality of W). It is moreover easy to see that there exists a unique V such that $A = VUD_B$ and $V = 0$ on $\mathrm{Im}(D_B)^\perp$. □

Now, we are ready for the proof of Theorem 4.1.1:

Proof of Theorem 4.1.1. First of all, the existence of two contractions $Z \in \mathcal{B}(H_1, K_1)$ and $Y \in \mathcal{B}(H_2, K_2)$ such that $D = D_{C^*} Y$ and $A = ZD_C$ comes from Proposition 4.1.3 and Corollary 4.1.4, as $\begin{bmatrix} A \\ C \end{bmatrix}$ and $\begin{bmatrix} C & D \end{bmatrix}$ are contractions. We denote the minimal solutions by Y_0, and respectively Z_0.

Set $\mathbf{A} = \begin{bmatrix} A & B \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} C & D \end{bmatrix}$, so that we have $T = \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}$, with $\|\mathbf{B}\| \leq 1$. Now, using that $TD_T = D_{T^*}T$, we have:

$$\begin{aligned}
\mathbf{Id}_{H_1 \oplus H_2} - \mathbf{B}^*\mathbf{B} &= \begin{bmatrix} \mathrm{Id}_{H_1} - C^*C & -C^*D \\ -D^*C & \mathrm{Id}_{H_2} - D^*D \end{bmatrix} \\
&= \begin{bmatrix} \mathrm{Id}_{H_1} - C^*C & -C^*D_{C^*}Y_0 \\ -Y_0^*D_{C^*}C & \mathrm{Id}_{H_2} - Y_0^*D_{C^*}D_{C^*}Y_0 \end{bmatrix} \\
&= \begin{bmatrix} \mathrm{Id}_{H_1} - C^*C & -D_C C^*Y_0 \\ -Y_0^*CD_C & \mathrm{Id}_{H_2} - Y_0^*Y_0 + Y_0^*CC^*Y_0 \end{bmatrix} \\
&= \mathbf{S}^*\mathbf{S},
\end{aligned}$$

where $\mathbf{S} = \begin{bmatrix} D_C & -C^*Y_0 \\ 0 & D_{Y_0} \end{bmatrix}$.

For every $w \in H_1 \oplus H_2$, we have:

$$\left\langle (\mathbf{Id}_{H_1 \oplus H_2} - \mathbf{B}^*\mathbf{B})w, w \right\rangle = \langle \mathbf{S}^*\mathbf{S}w, w \rangle,$$

which is equivalent to $\|\mathbf{S}w\| = \|D_{\mathbf{B}}w\|$. Thus, there is an isometry $\mathbf{U} \in \mathcal{B}(H_1 \oplus H_2)$ such that $\mathbf{S} = \mathbf{U}D_{\mathbf{B}}$. Indeed, let $U : \mathrm{Im}\,(D_{\mathbf{B}}) \to H_1 \oplus H_2$, $D_{\mathbf{B}}x \mapsto \mathbf{S}x$. We extend $\mathbf{U}$ by continuity to $\overline{\mathrm{Im}\,(D_{\mathbf{B}})}$, and we set $\mathbf{U} = \mathrm{Id}$ on $\mathrm{Im}\,(D_{\mathbf{B}})^{\perp}$.

Suppose that T is a contraction. Then, by Corollary 4.1.5, there exists a contraction

$$\mathbf{V} = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \in \mathcal{B}(H_1 \oplus H_2, K_1)$$

such that:

$$\begin{cases} \mathbf{A} = \mathbf{V}\mathbf{U}D_{\mathbf{B}}, & (4.3) \\ \mathbf{V} = 0 \text{ on } \mathrm{Im}(\mathbf{S})^{\perp}. & (4.4) \end{cases}$$

By Corollary 4.1.4, there exists a contraction $W \in \mathcal{B}(H_2, K_1)$ such that $\mathbf{V} = \begin{bmatrix} V_1 & D_{V_1^*}W \end{bmatrix}$. The operator W can be chosen such that $\mathrm{Im}(D_{V_1^*}) \subset \mathrm{Ker}(W^*)$ (in that case, the minimal solution W_0 is unique).

Then, (4.3) is equivalent to

$$\begin{aligned}
\begin{bmatrix} A & B \end{bmatrix} &= \begin{bmatrix} V_1 & D_{V_1^*}W \end{bmatrix} \begin{bmatrix} D_C & -C^*Y_0 \\ 0 & D_{Y_0} \end{bmatrix} \\
&= \begin{bmatrix} V_1D_C & -V_1C^*Y_0 + D_{V_1^*}WD_{Y_0} \end{bmatrix}. \qquad (4.5)
\end{aligned}$$

In particular, we have $A = V_1D_C$. We now show that $V_1 = Z_0$.

Fact 1. $\mathrm{Im}(D_C)^{\perp} \oplus \{0\} \subset \mathrm{Im}(\mathbf{S})^{\perp}$.

Proof. Let $v \in \mathrm{Im}(D_C)^{\perp} = \mathrm{Ker}(D_C)$. In order to prove that $\begin{bmatrix} v \\ 0 \end{bmatrix} \in \mathrm{Ker}(\mathbf{S}^*) = \mathrm{Im}(\mathbf{S})^{\perp}$, notice that we have

$$\mathbf{S}^* \begin{bmatrix} v \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -Y_0^* C v \end{bmatrix}.$$

As we know that $Y_0^* = 0$ on $\mathrm{Im}(D_{C^*})^\perp = \mathrm{Ker}(D_{C^*})$, it is enough to show $Cv \in \mathrm{Ker}(D_{C^*})$. Using again the identity $CD_C = D_{C^*}C$, we have

$$\|D_{C^*}Cv\|^2 = \langle D_{C^*}Cv, D_{C^*}Cv\rangle = \left\langle Cv, D_{C^*}^2 Cv\right\rangle = \left\langle Cv, CD_C^2 v\right\rangle = 0,$$

which completes the proof of the Fact 1. □

Continuing the proof of Theorem 4.1.1, we can deduce from (4.4) that $V_1 = 0$ on $\mathrm{Im}(D_C)^\perp$ and, thus, $V_1 = Z_0$. Finally, (4.5) is equivalent to

$$B = -Z_0 C^* Y_0 + D_{Z_0^*} W D_{Y_0}.$$

Conversely, if there exists a contraction $W \in \mathcal{B}(H_2, K_1)$ such that $B = D_{Z_0^*} V D_{Y_0} - Z_0 C^* Y_0$, then it is easy to check that $\mathbf{A} = \mathbf{V}'\mathbf{S} = \mathbf{V}'\mathbf{U}\mathbf{D}_\mathbf{B}$, with $\mathbf{V}' = \begin{bmatrix} Z & D_{Z^*}W \end{bmatrix}$. As $\mathbf{V}'$ is a contraction (Corollary 4.1.4), this implies that T is a contraction (Corollary 4.1.5). □

4.2 Contractive 3×3 matrices

In [29, Lemma 2.7], Gupta has used Parrott's theorem on matrix completion to characterize the contractive matrices of the form

$$T = \begin{pmatrix} \omega & \alpha & 0 \\ 0 & \omega & \beta \\ 0 & 0 & \omega \end{pmatrix}, \quad \text{with } \omega \in \mathbb{D} \text{ and } \alpha, \beta \in \mathbb{C}.$$

In this section, we extend this argument to achieve a broader characterization of contractive 3×3 upper-triangular matrices, whose entries can be either scalars or operators.

First of all, the following result provides a criterion for determining whether a 3×3 operator matrix is a contraction:

Theorem 4.2.1. *Let H_1, H_2, H_3 be three Hilbert spaces. Let $W_i \in \mathcal{B}(H_i)$, $1 \le i \le 3$, be three contractions and denote*

$$T = \begin{bmatrix} W_1 & A_1 & B \\ 0 & W_2 & A_2 \\ 0 & 0 & W_3 \end{bmatrix} \in \mathcal{B}(H_1 \oplus H_2 \oplus H_3).$$

Then, T is a contraction if and only if there exist three contractions $V_1 \in \mathcal{B}(H_2, H_1)$, $V_2 \in \mathcal{B}(H_3, H_2)$, $V_3 \in \mathcal{B}(H_3, H_1)$ such that:

$$\begin{cases} A_1 = D_{W_1^*} V_1 D_{W_2}, & (4.6) \\ A_2 = D_{W_2^*} V_2 D_{W_3}, & (4.7) \\ B = \left[D_{W_1^*}(\mathrm{Id} - V_1 V_1^*) D_{W_1^*}\right]^{1/2} V_3 \left[D_{W_3}(\mathrm{Id} - V_2^* V_2) D_{W_3}\right]^{1/2} - D_{W_1^*} V_1 W_2^* V_2 D_{W_3}. & (4.8) \end{cases}$$

Proof. First, if T is a contraction, then $\begin{bmatrix} W_1 & A_1 \\ 0 & W_2 \end{bmatrix}$ and $\begin{bmatrix} W_2 & A_2 \\ 0 & W_3 \end{bmatrix}$ are also contractions, as they are compressions of T. Then, Parrott's theorem (Theorem 4.1.1) implies that (4.6) and (4.7) are satisfied. Moreover, V_1 and V_2 can be chosen such that $\mathrm{Im}\left(D_{W_1^*}\right)^\perp \subset \mathrm{Ker}(V_1^*)$ and $\mathrm{Im}\left(D_{W_2^*}\right)^\perp \subset \mathrm{Ker}(V_2^*)$. With this choice, V_1 and V_2 are (respectively) the solutions of the equations $A_1 = D_{W_1^*} V_1 D_{W_2}$ and $A_2 = D_{W_2^*} V_2 D_{W_3}$ with minimal operator norm. In the following we assume that (4.6) and (4.7) are satisfied, and that V_1 and V_2 are chosen as above.

Now, denote
$$A = \begin{bmatrix} W_1 & A_1 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & W_2 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} A_2 \\ W_3 \end{bmatrix}.$$
By Parrott's theorem, T is a contraction if and only if:
$$B = (\mathrm{Id} - ZZ^*)^{1/2} V_3 (\mathrm{Id} - Y^*Y)^{1/2} - ZC^*Y, \text{ for some contraction } V_3 \in \mathcal{B}(H_3, H_1). \tag{4.9}$$
Here, Y and Z are contractions such that
$$\begin{cases} D = D_{C^*} Y \\ Y^* = 0 & \text{on } \mathrm{Im}\,(D_{C^*})^\perp \end{cases}$$
and
$$\begin{cases} A = Z D_C \\ Z = 0 & \text{on } \mathrm{Im}\,(D_C)^\perp \end{cases}.$$
The existence of such Y and Z is ensured by Proposition 4.1.3 and Corollary 4.1.4, as $\begin{bmatrix} A \\ C \end{bmatrix}$ and $\begin{bmatrix} C & D \end{bmatrix}$ are contractions.

We have $\mathrm{Id} - CC^* = \begin{bmatrix} \mathrm{Id} - W_2 W_2^* & 0 \\ 0 & \mathrm{Id} \end{bmatrix}$ and $\mathrm{Id} - C^*C = \begin{bmatrix} \mathrm{Id} & 0 \\ 0 & \mathrm{Id} - W_2^* W_2 \end{bmatrix}$.

- Assume first that $\|W_2\| < 1$. Then, D_C and D_{C^*} are invertible, and we get
$$Y = D_{C^*}^{-1} D = \begin{bmatrix} D_{W_2^*}^{-1} A_2 \\ W_3 \end{bmatrix} = \begin{bmatrix} V_2 D_{W_3} \\ W_3 \end{bmatrix},$$
$$Z = A D_C^{-1} = \begin{bmatrix} W_1 & A_1 D_{W_2}^{-1} \end{bmatrix} = \begin{bmatrix} W_1 & D_{W_1^*} V_1 \end{bmatrix}.$$

  It follows that $D_{Z^*} = \left[D_{W_1^*}(\mathrm{Id} - V_1 V_1^*) D_{W_1^*}\right]^{1/2}$, $D_Y = \left[D_{W_3}(\mathrm{Id} - V_2^* V_2) D_{W_3}\right]^{1/2}$ and $ZC^*Y = D_{W_1^*} V_1 W_2^* V_2 D_{W_3}$. Therefore, (4.9) is equivalent to (4.8).

- In the general case, D_C and D_{C^*} are not necessarily invertible. Let $Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$ and $Z = \begin{bmatrix} Z_1 & Z_2 \end{bmatrix}$.

  We have:
$$D = D_{C^*} Y \iff \begin{cases} D_{W_2^*} Y_1 & = A_2 \\ Y_2 & = W_3 \end{cases}$$

$$\iff \begin{cases} Y_1^* D_{W_2^*} & = D_{W_3} V_2^* D_{W_2^*} \\ Y_2 & = W_3 \end{cases}.$$

Thus, we have $Y_1^* = D_{W_3} V_2^*$ on $\overline{\operatorname{Im}\left(D_{W_2^*}\right)}$.
Moreover, we have:

$$\begin{aligned} \operatorname{Im}\left(D_{W_2^*}\right)^\perp \oplus \{0\} &= \operatorname{Ker}\left(D_{W_2^*}\right) \oplus \{0\} \\ &\subset \operatorname{Ker}(D_{C^*}) \\ &= \operatorname{Im}(D_{C^*})^\perp. \end{aligned}$$

Thus, $Y^* = 0$ on $\operatorname{Im}(D_{C^*})^\perp$ implies that $Y_1^* = 0 = D_{W_2} V_2^*$ on $\operatorname{Im}\left(D_{W_2^*}\right)^\perp$. In the end, we get $Y_1^* = D_{W_3} V_2^*$ on $\overline{\operatorname{Im}\left(D_{W_2^*}\right)} \oplus \operatorname{Im}\left(D_{W_2^*}\right)^\perp$ and, therefore, $Y = \begin{bmatrix} V_2 D_{W_3} \\ W_3 \end{bmatrix}$.

Furthermore, we have also:

$$\begin{aligned} A = ZD_C &\iff \begin{cases} Z_1 & = W_1 \\ Z_2 D_{W_2} & = A_1 \end{cases} \\ &\iff \begin{cases} Z_1 & = W_1 \\ Z_2 D_{W_2} & = D_{W_1^*} V_1 D_{W_2} \end{cases}, \end{aligned}$$

which implies that $Z_2 = D_{W_1^*} V_1$ on $\overline{\operatorname{Im}\left(D_{W_2}\right)}$.

As $\{0\} \oplus \operatorname{Im}\left(D_{W_2}\right)^\perp \subset \operatorname{Im}(D_C)^\perp$, $Z = 0$ on $\operatorname{Im}(D_C)^\perp$ implies that $Z_2 = 0$ on $\operatorname{Im}\left(D_{W_2}\right)^\perp$.

Claim. $V_1 = 0$ on $\operatorname{Im}\left(D_{W_2}\right)^\perp$.

Proof of the claim. If it were not the case, the operator $\widetilde{V_1}$ defined by

$$\widetilde{V_1} = \begin{cases} V_1 & \text{on } \overline{\operatorname{Im}\left(D_{W_2}\right)} \\ 0 & \text{on } \operatorname{Im}\left(D_{W_2}\right)^\perp \end{cases}$$

would be a solution of the equation $A_1 = D_{W_1^*} \widetilde{V_1} D_{W_2}$ satisfying $\left\|\widetilde{V_1}\right\| < \|V_1\|$, which is impossible. □

Therefore, we have $Z_2 = D_{W_1^*} V_1$ on $\overline{\operatorname{Im}\left(D_{W_2}\right)} \oplus \operatorname{Im}\left(D_{W_2}\right)^\perp$ and, then, we get $Z = \begin{bmatrix} W_1 & D_{W_1^*} V_1 \end{bmatrix}$.
We conclude the proof as in the case where $\|W_2\| < 1$. □

We obtain the following general criterion in the scalar case:

Theorem 4.2.2. *Let $\omega_1, \omega_2, \omega_3 \in \overline{\mathbb{D}}$. Then, $T = \begin{pmatrix} \omega_1 & \alpha_1 & \beta \\ 0 & \omega_2 & \alpha_2 \\ 0 & 0 & \omega_3 \end{pmatrix}$ is a contraction when acting on the*

Hilbert space $\mathbb{C}^3$ *if and only if:*

$$\begin{cases} |\omega_2| < 1, \\ |\alpha_i|^2 \leq (1-|\omega_i|^2)(1-|\omega_{i+1}|^2), \quad i = 1,2, & (4.10) \\ \left|\beta(1-|\omega_2|^2) + \alpha_1\alpha_2\overline{\omega_2}\right|^2 \leq \left[(1-|\omega_1|^2)(1-|\omega_2|^2) - |\alpha_1|^2\right]\left[(1-|\omega_2|^2)(1-|\omega_3|^2) - |\alpha_2|^2\right] & (4.11) \end{cases}$$

$$\text{or } \begin{cases} |\omega_2| = 1, \\ \alpha_i = 0, \quad i = 1,2, & (4.12) \\ |\beta|^2 \leq (1-|\omega_1|^2)(1-|\omega_3|^2). & (4.13) \end{cases}$$

Proof. This result follows from Theorem 4.2.1. We give some details for completeness. As in the proof of Theorem 4.2.1, if T is a contraction, then the two dimensional compressions $\begin{bmatrix} \omega_1 & \alpha_1 \\ 0 & \omega_2 \end{bmatrix}$ and $\begin{bmatrix} \omega_2 & \alpha_2 \\ 0 & \omega_3 \end{bmatrix}$ are also contractions. Thus (4.10) is satisfied, and it will be assumed from now on. Note that if $|\omega_2| = 1$, this implies that $\alpha_1 = \alpha_2 = 0$.

We use similar notation as in the proof of Theorem 4.2.1, with:

$$A = \begin{bmatrix} \omega_1 & \alpha_1 \end{bmatrix}, \quad B = \begin{bmatrix} \beta \end{bmatrix}, \quad C = \begin{bmatrix} 0 & \omega_2 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} \alpha_2 \\ \omega_3 \end{bmatrix}.$$

By Theorem 4.1.1, T is a contraction if and only if:

$$B = (\mathrm{Id} - ZZ^*)^{1/2} V (\mathrm{Id} - Y^*Y)^{1/2} - ZC^*Y, \text{ for some contraction } V, \tag{4.14}$$

where Y and Z are contractions such that

$$\begin{cases} D = D_{C^*} Y \\ Y^* = 0 & \text{on } \mathrm{Im}\,(D_{C^*})^\perp \end{cases}$$

and

$$\begin{cases} A = Z D_C \\ Z = 0 & \text{on } \mathrm{Im}\,(D_C)^\perp \end{cases}.$$

We have

$$\mathrm{Id} - CC^* = \begin{bmatrix} 1-|\omega_2|^2 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$\mathrm{Id} - C^*C = \begin{bmatrix} 1 & 0 \\ 0 & 1-|\omega_2|^2 \end{bmatrix}.$$

- Assume first that $|\omega_2| < 1$. An easy computation shows that

$$Y = (\mathrm{Id} - CC^*)^{-1/2} D = \begin{bmatrix} \frac{\alpha_2}{\sqrt{1-|\omega_2|^2}} \\ \omega_3 \end{bmatrix}$$

and

$$Z = A\,(\mathrm{Id} - C^*C)^{-1/2} = \begin{bmatrix} \omega_1 & \frac{\alpha_1}{\sqrt{1-|\omega_2|^2}} \end{bmatrix}.$$

Thus, T is a contraction if and only if (4.14) is satisfied, that is

$$\beta + \frac{\alpha_1\alpha_2\overline{\omega_2}}{1-|\omega_2|^2} = \left(1-|\omega_1|^2 - \frac{|\alpha_1|^2}{1-|\omega_2|^2}\right)^{1/2} V \left(1-|\omega_3|^2 - \frac{|\alpha_2|^2}{1-|\omega_2|^2}\right)^{1/2}$$

for some contraction V, which is equivalent to (4.11).

- Assume now that $|\omega_2| = 1$.

  Let $Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ and $Z = \begin{bmatrix} z_1 & z_2 \end{bmatrix}$. As $D = (\mathrm{Id} - CC^*)^{1/2}Y$, we get $D^* = Y^*(\mathrm{Id} - CC^*)^{1/2}$. This holds if and only if $y_2 = \omega_3$. Moreover, $Y^* = 0$ on $\mathrm{Im}\,(D_{C^*})^{\perp}$ implies that $y_1 = 0$.

  Similarly, $A = Z(\mathrm{Id} - C^*C)^{1/2}$ holds if and only if $z_1 = \omega_1$, and $Z = 0$ on $\mathrm{Im}\,(D_C)^{\perp}$ implies that $z_2 = 0$.

  Then, we have $ZC^*Y = 0$ and, therefore, T is a contraction if and only if $|\beta|^2 \le (1-|\omega_3|^2)(1-|\omega_1|^2)$. This is equivalent to (4.13).

□

4.3 Contractive 4×4 matrices

The aim of this section is to establish the following criterion for contractivity:

Theorem 4.3.1. *Let* $T = \begin{pmatrix} \omega_1 & \alpha_1 & \beta_1 & \gamma \\ 0 & \omega_2 & \alpha_2 & \beta_2 \\ 0 & 0 & \omega_3 & \alpha_3 \\ 0 & 0 & 0 & \omega_4 \end{pmatrix} \in \mathcal{M}_4(\mathbb{C})$*, with* $\omega_1, \omega_2, \omega_3, \omega_4 \in \overline{\mathbb{D}}$.

Then T *is a contraction when acting on the Hilbert space* $\mathbb{C}^4$ *if and only if:*

$$\begin{cases}
|\omega_2| < 1 \quad ; \qquad |\omega_3| < 1 \\
|\alpha_i|^2 \le (1-|\omega_i|^2)(1-|\omega_{i+1}|^2), \quad i = 1,3 \qquad\qquad (4.15) \\
|\alpha_2|^2 < (1-|\omega_2|^2)(1-|\omega_3|^2) \\
|\beta_i(1-|\omega_{i+1}|^2) + \alpha_i\alpha_{i+1}\overline{\omega_{i+1}}|^2 \le \left[(1-|\omega_i|^2)(1-|\omega_{i+1}|^2) - |\alpha_i|^2\right] \times \\
\qquad \left[(1-|\omega_{i+1}|^2)(1-|\omega_{i+2}|^2) - |\alpha_{i+1}|^2\right], \quad i = 1,2 \qquad\qquad (4.16) \\
\Big|\gamma[(1-|\omega_2|^2)(1-|\omega_3|^2) - |\alpha_2|^2] + \alpha_1\beta_2\overline{\omega_2}(1-|\omega_3|^2) + \alpha_3\beta_1\overline{\omega_3}(1-|\omega_2|^2) \\
\qquad + \beta_1\beta_2\overline{\alpha_2} + \alpha_1\alpha_2\alpha_3\overline{\omega_2\omega_3}\Big|^2 (1-|\omega_2|^2)(1-|\omega_3|^2) \\
\quad \le \Big[\Big((1-|\omega_1|^2)(1-|\omega_2|^2) - |\alpha_1|^2\Big)\Big((1-|\omega_2|^2)(1-|\omega_3|^2) - |\alpha_2|^2\Big) - \\
\qquad \Big|\alpha_1\alpha_2\overline{\omega_2} + \beta_1(1-|\omega_2|^2)\Big|^2\Big] \qquad\qquad (4.17) \\
\quad \times \Big[\Big((1-|\omega_2|^2)(1-|\omega_3|^2) - |\alpha_2|^2\Big)\Big((1-|\omega_3|^2)(1-|\omega_4|^2) - |\alpha_3|^2\Big) \\
\qquad - \Big|\alpha_2\alpha_3\overline{\omega_3} + \beta_2(1-|\omega_3|^2)\Big|^2\Big]
\end{cases}$$

$$
\text{or } \begin{cases}
|\omega_2| < 1 \quad ; \quad |\omega_3| < 1 \\
|\alpha_i|^2 \le (1-|\omega_i|^2)(1-|\omega_{i+1}|^2), \qquad i = 1,3 & (4.18) \\
|\alpha_2|^2 = (1-|\omega_2|^2)(1-|\omega_3|^2) \\
\beta_i = \dfrac{-\alpha_i \alpha_{i+1} \omega_{i+1}}{1-|\omega_{i+1}|^2}, \qquad i = 1,2 & (4.19) \\
\left|\gamma - \dfrac{\overline{\omega_2 \omega_3} \alpha_1 \alpha_2 \alpha_3}{(1-|\omega_2|^2)(1-|\omega_3|^2)}\right|^2 \left(1-|\omega_2|^2\right)\left(1-|\omega_3|^2\right) & (4.20) \\
\quad \le \left[(1-|\omega_1|^2)(1-|\omega_2|^2) - |\alpha_1|^2\right] \times \left[(1-|\omega_3|^2)(1-|\omega_4|^2) - |\alpha_3|^2\right]
\end{cases}
$$

$$
\text{or } \begin{cases}
|\omega_2| = 1 \quad ; \quad |\omega_3| < 1 \\
\alpha_1 = \alpha_2 = 0 \\
|\alpha_3|^2 \le (1-|\omega_3|^2)(1-|\omega_4|^2) \\
|\beta_1|^2 \le (1-|\omega_1|^2)(1-|\omega_3|^2) \\
\left|\beta_2\left(1-|\omega_3|^2\right)\right|^2 \le (1-|\omega_3|^2)(1-|\omega_4|^2) - |\alpha_3|^2 \\
\left|\gamma(1-|\omega_3|^2) + \alpha_3 \beta_1 \overline{\omega_3}\right|^2 \le \left[(1-|\omega_3|^2)(1-|\omega_4|^2) - |\alpha_3|^2\right]\left[(1-|\omega_1|^2)(1-|\omega_3|^2) - |\beta_1|^2\right]
\end{cases}
$$

$$
\text{or } \begin{cases}
|\omega_3| = 1 \quad ; \quad |\omega_2| < 1 \\
|\alpha_1|^2 \le (1-|\omega_1|^2)(1-|\omega_2|^2) \\
\alpha_2 = \alpha_3 = 0 \\
|\beta_1(1-|\omega_2|^2)|^2 \le (1-|\omega_1|^2)(1-|\omega_2|^2) - |\alpha_1|^2 \\
|\beta_2|^2 \le (1-|\omega_2|^2)(1-|\omega_4|^2) \\
\left|\gamma(1-|\omega_2|^2) + \alpha_1 \beta_2 \overline{\omega_2}\right|^2 \le \left[(1-|\omega_1|^2)(1-|\omega_2|^2) - |\alpha_1|^2\right]\left[(1-|\omega_2|^2)(1-|\omega_4|^2) - |\beta_2|^2\right]
\end{cases}
$$

$$
\text{or } \begin{cases}
|\omega_2| = |\omega_3| = 1 \\
\alpha_1 = \alpha_2 = \alpha_3 = 0 \\
|\beta_i|^2 \le (1-|\omega_i|^2)(1-|\omega_{i+2}|^2), \qquad i = 1,2 \\
\left|\gamma\right|^2 \le (1-|\omega_1|^2)(1-|\omega_4|^2)
\end{cases}
$$

We will follow the same approach as for 3×3 matrices, employing Parrott's theorem on matrix completion. However, we will find that the calculations become more intricate.

First of all, we will need the following technical lemma, which provides the diagonalization of a specific matrix M. This will allow us to define powers M^s of M.

Lemma 4.3.2. *Let $\omega, \alpha \in \mathbb{C}$, $\omega \neq 0$, $M = \begin{pmatrix} 1-|\omega|^2 & -\overline{\omega}\alpha \\ -\overline{\alpha}\omega & 1-|\alpha|^2 \end{pmatrix}$ and $s \in \mathbb{R}$.*

(i) We have $M = \frac{1}{|\alpha|^2+|\omega|^2} \begin{pmatrix} -\alpha & \overline{\omega} \\ \omega & \overline{\alpha} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1-|\alpha|^2-|\omega|^2 \end{pmatrix} \begin{pmatrix} -\overline{\alpha} & \overline{\omega} \\ \omega & \alpha \end{pmatrix}$;

(ii) We can define

$$
M^s = \frac{1}{|\alpha|^2+|\omega|^2} \begin{pmatrix} |\alpha|^2 + |\omega|^2\left[1-|\alpha|^2-|\omega|^2\right]^s & -\alpha\overline{\omega}\left(1-\left[1-|\alpha|^2-|\omega|^2\right]^s\right) \\ -\overline{\alpha}\omega\left(1-\left[1-|\alpha|^2-|\omega|^2\right]^s\right) & |\omega|^2 + |\alpha|^2\left[1-|\alpha|^2-|\omega|^2\right]^s \end{pmatrix},
$$

provided $\left[1-|\alpha|^2-|\omega|^2\right]^s$ *is well-defined.*

Proof. We just diagonalize M. Using the characteristic polynomial may lead to quite complicated computations. However, let us recall that the product of the eigenvalue is equal to the determinant, and that the sum of the eigenvalues is equal to the trace. In that case, this gives that the product of the eigenvalues is equal to $1-|\alpha|^2-|\omega|^2$ and that the sum of the eigenvalues is equal to $2-|\omega|^2-|\alpha|^2$. Thus, a *natural* candidate for the spectrum of M is $\sigma := \{1; 1-|\alpha|^2-|\omega|^2\}$. We can then easily check that $\begin{pmatrix}-\alpha\\ \omega\end{pmatrix}$ and $\begin{pmatrix}\overline{\omega}\\ \overline{\alpha}\end{pmatrix}$ are eigenvectors associated (respectively) to 1 and $1-|\alpha|^2-|\omega|^2$, and then we obtain the first item of this lemma. The second item follows as a direct consequence of the first one. □

Now, let us start with the particular case where one of the diagonal element of the matrix, ω_3, is equal to 0:

Lemma 4.3.3. *Let* $T = \begin{pmatrix}\omega_1 & \alpha_1 & \beta_1 & \gamma\\ 0 & \omega_2 & \alpha_2 & \beta_2\\ 0 & 0 & 0 & \alpha_3\\ 0 & 0 & 0 & \omega_4\end{pmatrix} \in \mathcal{M}_4(\mathbb{C})$, *with* $\omega_2 \in \mathbb{D}$ *and* $\omega_1, \omega_4 \in \overline{\mathbb{D}}$.

Then T is a contraction if and only if:

$$\begin{cases}|\alpha_i|^2 \le (1-|\omega_i|^2)(1-|\omega_{i+1}|^2), \quad i=1,3 & (4.21)\\ |\alpha_2|^2 = 1-|\omega_2|^2 & \\ \beta_1 = \dfrac{-\alpha_1\alpha_2\overline{\omega_2}}{1-|\omega_2|^2} & (4.22)\\ \beta_2 = 0 & (4.23)\\ |\gamma|^2(1-|\omega_2|^2) \le \left[(1-|\omega_1|^2)(1-|\omega_2|^2)-|\alpha_1|^2\right]\left[1-|\omega_4|^2-|\alpha_3|^2\right] & (4.24)\end{cases}$$

$$or \begin{cases}|\alpha_i|^2 \le (1-|\omega_i|^2)(1-|\omega_{i+1}|^2), \quad i=1,3 & (4.25)\\ |\alpha_2|^2 < 1-|\omega_2|^2 & \\ |\beta_i(1-|\omega_{i+1}|^2)+\alpha_i\alpha_{i+1}\overline{\omega_{i+1}}|^2 \le \left[(1-|\omega_i|^2)(1-|\omega_{i+1}|^2)-|\alpha_i|^2\right]\times & \\ \qquad \left[(1-|\omega_{i+1}|^2)(1-|\omega_{i+2}|^2)-|\alpha_{i+1}|^2\right], \quad i=1,2 & (4.26)\\ \left|\gamma\left(1-|\omega_2|^2-|\alpha_2|^2\right)+\beta_2\left(\overline{\omega_2}\alpha_1+\overline{\alpha_2}\beta_1\right)\right|^2(1-|\omega_2|^2) & \\ \quad \le \Big[\left((1-|\omega_1|^2)(1-|\omega_2|^2)-|\alpha_1|^2\right)\left(1-|\omega_2|^2-|\alpha_2|^2\right) & \\ \qquad -\left|\beta_1(1-|\omega_2|^2)+\alpha_1\alpha_2\overline{\omega_2}\right|^2\Big] & (4.27)\\ \qquad \times\left[\left(1-|\omega_2|^2-|\alpha_2|^2\right)\left(1-|\omega_4|^2-|\alpha_3|^2\right)-|\beta_2|^2\right] & \end{cases}$$

Proof. Denote $\omega_3 = 0$. First of all, if T is a contraction, then the compressions $\begin{bmatrix}\omega_1 & \alpha_1 & \beta_1\\ 0 & \omega_2 & \alpha_2\\ 0 & 0 & \omega_3\end{bmatrix}$

and $\begin{bmatrix} \omega_2 & \alpha_2 & \beta_2 \\ 0 & \omega_3 & \alpha_3 \\ 0 & 0 & \omega_4 \end{bmatrix}$ are contractions. As it is a necessary condition for T to be a contraction, we will assume in the following that it is the case. Thus, from Theorem 4.2.2, we have:

$$|\alpha_i|^2 \le (1-|\omega_i|^2)(1-|\omega_{i+1}|^2), \quad 1 \le i \le 3 \tag{4.28}$$

and

$$\begin{aligned} |\beta_i(1-|\omega_{i+1}|^2) + \alpha_i\alpha_{i+1}\overline{\omega_{i+1}}|^2 \le &\left[(1-|\omega_i|^2)(1-|\omega_{i+1}|^2) - |\alpha_i|^2\right]\times \\ &\left[(1-|\omega_{i+1}|^2)(1-|\omega_{i+2}|^2) - |\alpha_{i+1}|^2\right], \quad 1 \le i \le 2 \end{aligned} \tag{4.29}$$

Let $A = \begin{bmatrix} \omega_1 & \alpha_1 & \beta_1 \end{bmatrix}$, $B = \begin{bmatrix} \gamma \end{bmatrix}$, $C = \begin{bmatrix} 0 & \omega_2 & \alpha_2 \\ 0 & 0 & \omega_3 \\ 0 & 0 & 0 \end{bmatrix}$ and $D = \begin{bmatrix} \beta_2 \\ \alpha_3 \\ \omega_4 \end{bmatrix}$.

By assumption, $\begin{bmatrix} A \\ C \end{bmatrix}$ and $\begin{bmatrix} C & D \end{bmatrix}$ are contractions.

By Parrott's theorem, T is a contraction if and only if

$$B = (\mathrm{Id} - ZZ^*)^{1/2} V (\mathrm{Id} - Y^*Y)^{1/2} - ZC^*Y, \text{ for some contraction } V, \tag{4.30}$$

where Y and Z are contractions such that $D = (\mathrm{Id} - CC^*)^{1/2} Y$ and $A = Z(\mathrm{Id} - C^*C)^{1/2}$.

Note that the existence of two contractions Y and Z such that $D = (\mathrm{Id} - CC^*)^{1/2} Y$ and $A = Z(\mathrm{Id} - C^*C)^{1/2}$ is ensured by Parrott's theorem for column (respectively row) matrix-operators, as we are assuming that $\begin{bmatrix} A \\ C \end{bmatrix}$ and $\begin{bmatrix} C & D \end{bmatrix}$ are contractions.

We have $\mathrm{Id} - CC^* = \begin{bmatrix} 1-|\omega_2|^2-|\alpha_2|^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $\mathrm{Id} - C^*C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1-|\omega_2|^2 & -\alpha_2\overline{\omega_2} \\ 0 & -\omega_2\overline{\alpha_2} & 1-|\alpha_2|^2 \end{bmatrix}$.

- Assume first that $\omega_2 \neq 0$ and that $|\alpha_2|^2 < 1 - |\omega_2|^2$.

Then, denoting $\Sigma = |\alpha_2|^2 + |\omega_2|^2$, we have

$$(\mathrm{Id} - CC^*)^{-1/2} = \begin{bmatrix} \frac{1}{\sqrt{1-\Sigma}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and, by Lemma 4.3.2,

$$(\mathrm{Id} - C^*C)^{-1/2} = \frac{1}{\Sigma}\begin{bmatrix} \Sigma & 0 & 0 \\ 0 & |\alpha_2|^2 + \frac{|\omega_2|^2}{\sqrt{1-\Sigma}} & -\alpha_2\overline{\omega_2}\left[1 - \frac{1}{\sqrt{1-\Sigma}}\right] \\ 0 & -\overline{\alpha_2}\omega_2\left[1 - \frac{1}{\sqrt{1-\Sigma}}\right] & |\omega_2|^2 + \frac{|\alpha_2|^2}{\sqrt{1-\Sigma}} \end{bmatrix}.$$

Thus, we get $Y = (\mathrm{Id} - CC^*)^{-1/2} D = \begin{bmatrix} \frac{\beta_2}{\sqrt{1-\Sigma}} \\ \alpha_3 \\ \omega_4 \end{bmatrix}$.

Moreover, we get also

$$\begin{aligned} Z &= A\,(\mathrm{Id} - C^*C)^{-1/2} \\ &= \begin{bmatrix} \omega_1 & \frac{\alpha_1\left(|\alpha_2|^2\sqrt{1-\Sigma}+|\omega_2|^2\right)-\beta_1\overline{\alpha_2}\omega_2\left(\sqrt{1-\Sigma}-1\right)}{\Sigma\sqrt{1-\Sigma}} & \frac{-\alpha_2\overline{\omega_2}\alpha_1\left(\sqrt{1-\Sigma}-1\right)+\beta_1\left(|\omega_2|^2\sqrt{1-\Sigma}+|\alpha_2|^2\right)}{\Sigma\sqrt{1-\Sigma}} \end{bmatrix} \\ &:= \frac{1}{\Sigma\sqrt{1-\Sigma}} \begin{bmatrix} z_1' & z_2' & z_3' \end{bmatrix}, \end{aligned}$$

where

$$\begin{aligned} z_1' &= \Sigma\sqrt{1-\Sigma}\,\omega_1; \\ z_2' &= \alpha_1\left(|\alpha_2|^2\sqrt{1-\Sigma}+|\omega_2|^2\right) - \beta_1\overline{\alpha_2}\omega_2\left(\sqrt{1-\Sigma}-1\right); \\ z_3' &= -\alpha_2\overline{\omega_2}\alpha_1\left(\sqrt{1-\Sigma}-1\right) + \beta_1\left(|\omega_2|^2\sqrt{1-\Sigma}+|\alpha_2|^2\right). \end{aligned}$$

Thus, we have:

$$ZC^*Y = \frac{\beta_2}{\Sigma\,(1-\Sigma)}\,(\overline{\omega_2}z_2' + \overline{\alpha_2}z_3') = \frac{\beta_2}{1-\Sigma}\,(\alpha_1\overline{\omega_2} + \beta_1\overline{\alpha_2}).$$

Now, let us consider the Cholesky factorization (see *e.g.* [48, chapter 4]) of $\mathrm{Id} - C^*C$: let S_C be the (unique) upper-triangular matrix with positive diagonal such that

$$\mathrm{Id} - C^*C = S_C^* S_C.$$

It is easy to see that:

$$S_C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{1-|\omega_2|^2} & \frac{-\alpha_2\overline{\omega_2}}{\sqrt{1-|\omega_2|^2}} \\ 0 & 0 & \sqrt{\frac{1-|\omega|^2-|\alpha_2|^2}{1-|\omega_2|^2}} \end{bmatrix}.$$

Moreover, for every $x \in \mathbb{C}^3$, we have $\|S_C\,x\|^2 = \langle(\mathrm{Id} - C^*C)\,x, x\rangle = \|D_C\,x\|^2$. Therefore, there exists an isometry $U_C \in \mathcal{B}\left(\mathbb{C}^3\right)$ such that $D_C = U_C S_C$. Since U_C acts on the finite dimensional space $\mathbb{C}^3$, the operator U_C is even unitary.

Now, let $\widetilde{Z} = ZU_C$. We have

$$A = \widetilde{Z}S_C \tag{4.31}$$

and

$$\mathrm{Id} - ZZ^* = \mathrm{Id} - \widetilde{Z}\widetilde{Z}^*. \tag{4.32}$$

From (4.31) we get:

$$\widetilde{Z} = AS_C^{-1} = \begin{bmatrix} \omega_1 & \frac{\alpha_1}{\sqrt{1-|\omega_2|^2}} & \frac{\alpha_1\alpha 2\overline{\omega_2}+\beta_1(1-|\omega_2|^2)}{\sqrt{1-|\omega_2|^2}\sqrt{1-|\omega_2|-|\alpha_2|^2}} \end{bmatrix}.$$

Finally, by (4.30) and (4.32), T is a contraction if and only if:

$$\begin{aligned}&\left|\gamma\left(1-|\omega_2|^2-|\alpha_2|^2\right)+\beta_2\left(\overline{\omega_2}\alpha_1+\overline{\alpha_2}\beta_1\right)\right|^2(1-|\omega_2|^2)\\&\leq\left[\left((1-|\omega_1|^2)(1-|\omega_2|^2)-|\alpha_1|^2\right)\left(1-|\omega_2|^2-|\alpha_2|^2\right)-\left|\beta_1(1-|\omega_2|^2)+\alpha_1\alpha_2\overline{\omega_2}\right|^2\right]\times\cdot\\&\qquad\qquad\left[\left(1-|\omega_2|^2-|\alpha_2|^2\right)\left(1-|\omega_4|^2-|\alpha_3|^2\right)-|\beta_2|^2\right]\end{aligned}$$

- Assume now that $\omega_2=0$ and $|\alpha_2|<1$.

  We have then $\mathrm{Id}-CC^*=\begin{bmatrix}1-|\alpha_2|^2 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$ and $\mathrm{Id}-C^*C=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1-|\alpha_2|^2\end{bmatrix}$.

  Thus, we obtain

$$Y=(\mathrm{Id}-CC^*)^{-1/2}D=\begin{bmatrix}\frac{\beta_2}{\sqrt{1-|\alpha_2|^2}}\\ \alpha_3\\ \omega_4\end{bmatrix}$$

  and

$$Z=A\,(\mathrm{Id}-C^*C)^{-1/2}=\begin{bmatrix}\omega_1 & \alpha_1 & \frac{\beta_1}{\sqrt{1-|\alpha_2|^2}}\end{bmatrix}$$

  and we are brought back to the previous case (taking $\omega_2=0$ in the expressions of Y and Z). Note that in this case, as $\mathrm{Id}-C^*C$ is diagonal, $S_C=D_C$ and, thus, $\widetilde{Z}=Z$.

- Assume now that $\omega_2\neq 0$ and $\alpha_2=e^{i\theta_2}\sqrt{1-|\omega_2|^2}$, for $\theta_2\in]-\pi,\pi]$.

  Note that in that case, (4.29) implies that $\beta_1=\frac{-\alpha_1\alpha_2\overline{\omega_2}}{(1-|\omega_2|^2)}$ and $\beta_2=0$.

  Moreover, we have also:

$$(\mathrm{Id}-CC^*)^{1/2}=\begin{bmatrix}0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

  and

$$(\mathrm{Id}-C^*C)^{1/2}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1-|\omega_2|^2 & -e^{i\theta_2}\overline{\omega_2}\sqrt{1-|\omega_2|^2}\\ 0 & -\omega_2e^{-i\theta_2}\sqrt{1-|\omega_2|} & |\omega_2|^2\end{bmatrix}.$$

  Let $Y=\begin{bmatrix}y_1\\ y_2\\ y_3\end{bmatrix}$ and $Z=\begin{bmatrix}z_1 & z_2 & z_3\end{bmatrix}$. Then, $D=(\mathrm{Id}-CC^*)^{1/2}\,Y$ if and only if $y_2=\alpha_3$ and $y_3=\omega_4$. Taking for Y the minimal solution of the equation $D=D_{C^*}Y$, we can set $y_1=0$.

  Moreover, $A=Z\,(\mathrm{Id}-C^*C)^{1/2}$ if and only if:

$$\begin{cases} z_1 & & & =\omega_1\\ & (1-|\omega_2|^2)z_2 & -e^{-i\theta_2}\omega_2\sqrt{1-|\omega_2|^2}z_3 & =\alpha_1\\ & -e^{i\theta_2}\overline{\omega_2}\sqrt{1-|\omega_2|^2}z_2 & +|\omega_2|^2z_3 & =\beta_1\end{cases}$$

which is equivalent to:

$$\begin{cases} z_1 & & = \omega_1 \\ & (1-|\omega_2|^2)z_2 \quad -e^{-i\theta_2}\omega_2\sqrt{1-|\omega_2|^2}z_3 & = \alpha_1 \end{cases}$$

as we have, by assumption, $\beta_1 = \frac{-\alpha_1 e^{i\theta_2}\overline{\omega_2}}{\sqrt{1-|\omega_2|^2}}$. Again, we take for Z the minimal solution of the equation $A = ZD_C$, which means that $Z = 0$ on $\mathrm{Im}(D_C)^{\perp}$.

It is easy to see that the rank of D_C is equal to 2, and that $\left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\ \sqrt{1-|\omega_2|^2}\\ -e^{-i\theta_2}\omega_2\end{pmatrix}\right)$ is an orthonormal basis of $\mathrm{Im}(D_C)$. Using Gram-Schmidt's algorithm, we find that $\left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\ \sqrt{1-|\omega_2|^2}\\ -e^{-i\theta_2}\omega_2\end{pmatrix}, \begin{pmatrix}0\\ e^{i\theta_2}\overline{\omega_2}\\ \sqrt{1-|\omega_2|^2}\end{pmatrix}\right)$ is an orthonormal basis of $\mathbb{C}^3$ adapted to the decomposition $\mathbb{C}^3 = \mathrm{Im}(D_C) \oplus \mathrm{Im}(D_C)^{\perp}$. If we require that $D_C = 0$ on $\mathrm{Im}(D_C)^{\perp}$, then, we obtain the following system:

$$\begin{cases} z_1 & & = \omega_1 \\ & (1-|\omega_2|^2)z_2 \quad -e^{-i\theta_2}\omega_2\sqrt{1-|\omega_2|^2}z_3 & = \alpha_1 \\ & e^{i\theta_2}\overline{\omega_2}z_2 \quad +\sqrt{1-|\omega_2|^2}z_3 & = 0 \end{cases},$$

which is equivalent to:

$$\begin{cases} z_1 & = \omega_1 \\ z_2 & = \alpha_1 \\ z_3 & = \frac{-e^{i\theta_2}\overline{\omega_2}\alpha_1}{\sqrt{1-|\omega_2|^2}} \end{cases}.$$

Then we have $ZC^*Y = 0$.

Finally, T is a contraction if and only if

$$|\gamma|^2 \le \left[1-|\omega_1|^2 - \frac{|\alpha_1|^2}{1-|\omega_2|^2}\right]\left[1-|\omega_4|^2-|\alpha_3|^2\right].$$

- Assume now that $\omega_2 = 0$ and that $\alpha_2 = e^{i\theta_2}$, for some $\theta_2 \in]-\pi,\pi]$.

  We have then $\mathrm{Id} - CC^* = \begin{bmatrix}0&0&0\\0&1&0\\0&0&1\end{bmatrix}$ and $\mathrm{Id} - C^*C \begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}$.

  Again, let $Y = \begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$ and $Z = \begin{bmatrix}z_1 & z_2 & z_3\end{bmatrix}$.

  We have
$$D = (\mathrm{Id} - CC^*)^{1/2}\, Y \iff \begin{cases} y_2 = \alpha_3 \\ y_3 = \omega_4 \end{cases}.$$

Taking for Y the minimal solution of the equation $D = D_{C^*}Y$ gives us $y_1 = 0$.

Moreover, we have $A = Z(\mathrm{Id} - C^*C)^{1/2} \iff \begin{cases} z_1 = \omega_1 \\ z_2 = \alpha_1 \end{cases}$.

Taking for Z the minimal solution of the equation $A = ZD_C$ gives us $z_3 = 0$. We are therefore brought back to the previous case (taking $\omega_2 = 0$ in the expressions of Y and Z). ☐

Definition 4.3.4. For $\omega \in \mathbb{D}$, we define the Möbius transformation

$$M_\omega : \mathbb{C}\backslash\left\{\frac{1}{\overline{\omega}}\right\} \ni z \mapsto \frac{\omega - z}{1 - \overline{\omega} z}$$

We recall that M_ω is involutive, holomorphic on $\mathbb{D}$, and that for all $z \in \overline{\mathbb{D}}$, $|M_\omega(z)| \leq 1$, with equality if and only if $|z| = 1$. For more details, see for instance [43, Chapter IX, Section 2].

Definition 4.3.5. Let $\omega \in \mathbb{D}$. For a Hilbert space H, we define

$$\mathcal{B}_\omega(H) := \left\{T \in \mathcal{B}(H) : \frac{1}{\overline{\omega}} \notin \sigma(T)\right\}$$

and

$$M_\omega := \mathcal{B}_\omega(H) \ni T \mapsto (\omega \mathrm{Id} - T)(\mathrm{Id} - \overline{\omega} T)^{-1}.$$

Remark 4.3.6. All contractions belong to $\mathcal{B}_\omega(H)$.

Lemma 4.3.7. *With the above notations:*

(i) M_ω is involutive ;

(ii) T is a contraction if and only if $M_\omega(T)$ is a contraction.

Proof.

(i) Firstly, we note that if $T \in \mathcal{B}_\omega(H)$, then the spectrum $\sigma(M_\omega(T)) = M_\omega(\sigma(T))$ does not contain $\frac{1}{\overline{\omega}}$. This is because the equation $\frac{\omega - z}{1 - \overline{\omega} z} = \frac{1}{\overline{\omega}}$ has no solution. Therefore, $M_\omega(M_\omega(T))$ is well-defined. Subsequently, using properties of the rational functional calculus (Corollary 3.2.19), we can write $M_\omega(M_\omega(T)) = (M_\omega \circ M_\omega)(T) = T$.

(ii) This can be seen as a consequence of the von Neumann inequality. For completeness, we present here a more elementary and direct proof. From the previous item, it is enough to prove that if T is a contraction, then so is $M_\omega(T)$.
Let $x \in H$, and let $y = (\mathrm{Id} - \overline{\omega} T)^{-1} x$. First of all, we have:

$$\begin{aligned} \|x\|^2 &= \|(\mathrm{Id} - \overline{\omega} T) y\|^2 \\ &= \|y\|^2 + |\omega|^2 \|Ty\|^2 - 2\mathrm{Re}(\omega \langle Ty, y \rangle) \end{aligned}$$

Then, we can write:

$$\begin{aligned}\|M_\omega(T)x\|^2 &= \|(\overline{\omega}\mathrm{Id} - T)y\|^2 \\ &= |\omega|^2\|y\|^2 + \|Ty\|^2 - 2\mathrm{Re}(\omega\langle Ty, y\rangle) \\ &= \|x\|^2 - (1-|\omega|^2)\|y\|^2 + (1-|\omega|^2)\|Ty\|^2 \\ &\le \|x\|^2 - (1-|\omega|^2)\|y\|^2 + (1-|\omega|^2)\|y\|^2 = \|x\|^2\end{aligned}$$

In the end, we get $\|M_\omega(T)x\| \le \|x\|$, for all $x \in H$. Thus, $M_\omega(T)$ is a contraction.

□

Proof of Theorem 4.3.1.

- We first consider the case where $|\omega_2| < 1$ and $|\omega_3| < 1$.
  In this case, we use Lemma 4.3.7: T is a contraction if and only if $M_{\omega_3}(T)$ is a contraction. In the following, as long as there is no ambiguity, we will just write M instead of M_{ω_3}. Using the rational functional calculus, we get the following matrix representation for $M(T)$:

$$M(T) = \begin{pmatrix} M(\omega_1) & \alpha_1[M(\omega_1), M(\omega_2)] & \lambda_1 & \mu \\ 0 & M(\omega_2) & \alpha_2[M(\omega_2), M(\omega_3)] & \lambda_2 \\ 0 & 0 & M(\omega_3) & \alpha_3[M(\omega_3), M(\omega_4)] \\ 0 & 0 & 0 & M(\omega_4) \end{pmatrix} \tag{4.33}$$

where
$$\begin{aligned}\lambda_1 &= \alpha_1\alpha_2[M(\omega_1), M(\omega_2), M(\omega_3)] + \beta_1[M(\omega_1), M(\omega_3)] \\ \lambda_2 &= \alpha_2\alpha_3[M(\omega_2), M(\omega_3), M(\omega_4)] + \beta_2[M(\omega_2), M(\omega_4))] \\ \mu &= \alpha_1\alpha_2\alpha_3[M(\omega_1), M(\omega_2), M(\omega_3), M(\omega_4)] + \alpha_1\beta_2[M(\omega_1), M(\omega_2), M(\omega_4)] \\ &\quad + \alpha_3\beta_1[M(\omega_1), M(\omega_3), M(\omega_4)] + \gamma[M(\omega_1), M(\omega_4)].\end{aligned}$$

For $i, j, k, l \in [\![1, 4]\!]$, we have:

$$\begin{aligned}[M(\omega_i), M(\omega_j)] &= \frac{|\omega_3|^2 - 1}{(1-\overline{\omega_3}\omega_i)(1-\overline{\omega_3}\omega_j)} \\ [M(\omega_i), M(\omega_j), M(\omega_k)] &= \frac{\overline{\omega_3}(|\omega_3|^2 - 1)}{(1-\overline{\omega_3}\omega_i)(1-\overline{\omega_3}\omega_j)(1-\overline{\omega_3}\omega_k)} \\ [M(\omega_i), M(\omega_j), M(\omega_k), M(\omega_l)] &= \frac{\overline{\omega_3}^2(|\omega_3|^2 - 1)}{(1-\overline{\omega_3}\omega_i)(1-\overline{\omega_3}\omega_j)(1-\overline{\omega_3}\omega_k)(1-\overline{\omega_3}\omega_l)}\end{aligned}$$

Hence, (4.33) can be rewritten:

$$M(T) = \begin{pmatrix} \frac{\omega_3-\omega_1}{1-\overline{\omega_3}\omega_1} & \frac{\alpha_1(|\omega_3|^2-1)}{(1-\overline{\omega_3}\omega_1)(1-\overline{\omega_3}\omega_2)} & \frac{-\alpha_1\alpha_2\overline{\omega_3}-\beta_1(1-\overline{\omega_3}\omega_2)}{(1-\overline{\omega_3}\omega_1)(1-\overline{\omega_3}\omega_2)} & \mu \\ 0 & \frac{\omega_3-\omega_2}{1-\overline{\omega_3}\omega_2} & \frac{-\alpha_2}{1-\overline{\omega_3}\omega_2} & \frac{-\alpha_2\alpha_3\overline{\omega_3}+\beta_2(|\omega_3|^2-1)}{(1-\overline{\omega_3}\omega_2)(1-\overline{\omega_3}\omega_4)} \\ 0 & 0 & 0 & \frac{-\alpha_3}{1-\overline{\omega_3}\omega_4} \\ 0 & 0 & 0 & \frac{\omega_3-\omega_4}{1-\overline{\omega_3}\omega_4} \end{pmatrix}$$

$$
\begin{aligned}
\text{where} \quad \mu &= \frac{\gamma(|\omega_3|^2-1)}{(1-\overline{\omega_3}\omega_1)(1-\overline{\omega_3}\omega_4)} + \frac{\alpha_1\beta_2\overline{\omega_3}(|\omega_3|^2-1)}{(1-\overline{\omega_3}\omega_1)(1-\overline{\omega_3}\omega_2)(1-\overline{\omega_3}\omega_4)} \\
&\quad - \frac{\alpha_3\beta_1\overline{\omega_3}}{(1-\overline{\omega_3}\omega_1)(1-\overline{\omega_3}\omega_4)} - \frac{\alpha_1\alpha_2\alpha_3\overline{\omega_3}^2}{(1-\overline{\omega_3}\omega_1)(1-\overline{\omega_3}\omega_2)(1-\overline{\omega_3}\omega_4)} \\
&= \frac{(|\omega_3|^2-1)[\gamma(1-\overline{\omega_3}\omega_2)+\alpha_1\beta_2\overline{\omega_3}] - \overline{\omega_3}[\alpha_3\beta_1(1-\overline{\omega_3}\omega_2)+\alpha_1\alpha_2\alpha_3\overline{\omega_3}]}{(1-\overline{\omega_3}\omega_1)(1-\overline{\omega_3}\omega_2)(1-\overline{\omega_3}\omega_4)}
\end{aligned}
$$

Now, we apply Lemma 4.3.3 to $M(T)$, recalling that $|M(w_2)| = 1$ if and only if $|\omega_2| = 1$ and using the key identity:

$$
\forall\, u,v \in \mathbb{C},\ (1-|u|^2)(1-|v|^2) = |1-\overline{u}v|^2 - |u-v|^2 \tag{4.34}
$$

Using this identity, it is easy to see that

$$
\left|\frac{-\alpha_2}{1-\overline{\omega_3}\omega_2}\right|^2 \le 1 - \left|\frac{\omega_3-\omega_2}{1-\overline{\omega_3}\omega_2}\right|^2
$$

if and only if

$$
|\alpha_2|^2 \le 1-|\omega_2|^2,
$$

and that equality holds in one the two inequalities if and only if equality holds in the other one.

Thus, we can distinguish two subcases:

- If $|\alpha_2|^2 < 1-|\omega_2|^2$, it is easy to see (using Equation (4.34)) that (4.25) and (4.26) applied to $M(T)$ are (respectively) equivalent to (4.15) and (4.16). Let us give some details to show that (4.27) applied to $M(T)$ is equivalent to (4.17):

  (4.27) applied to $M(T)$ is equivalent to

$$
|E|^2\left(1-\left|\frac{\omega_3-\omega_2}{1-\overline{\omega_3}\omega_2}\right|^2\right) \le F\cdot G \tag{4.35}
$$

  where

$$
\begin{aligned}
E = {} & \frac{(|\omega_3|^2-1)[\gamma(1-\overline{\omega_3}\omega_2)+\alpha_1\beta_2\overline{\omega_3}] - \overline{\omega_3}[\alpha_3\beta_1(1-\overline{\omega_3}\omega_2)+\alpha_1\alpha_2\alpha_3\overline{\omega_3}]}{(1-\overline{\omega_3}\omega_1)(1-\overline{\omega_3}\omega_2)(1-\overline{\omega_3}\omega_4)} \\
& \times\left(1-\left|\frac{\omega_3-\omega_2}{1-\overline{\omega_3}\omega_2}\right|^2 - \left|\frac{\alpha_2}{1-\overline{\omega_3}\omega_2}\right|^2\right) - \frac{\alpha_2\alpha_3\overline{\omega_3}+\beta_2(1-|\omega_3|^2)}{(1-\overline{\omega_3}\omega_2)(1-\overline{\omega_3}\omega_4)} \\
& \times\left(\frac{\alpha_1(\overline{\omega_3}-\overline{\omega_2})(|\omega_3|^2-1)}{|1-\overline{\omega_3}\omega_2|^2(1-\overline{\omega_3}\omega_1)} + \frac{|\alpha_2|^2\alpha_1\overline{\omega_3}+\beta_1\overline{\alpha_2}(1-\overline{\omega_3}\omega_2)}{|1-\overline{\omega_3}\omega_2|^2(1-\overline{\omega_3}\omega_1)}\right)
\end{aligned}
$$

$$F=\left[1-\left|\frac{\omega_3-\omega_2}{1-\overline{\omega_3}\omega_2}\right|^2-\left|\frac{\alpha_2}{1-\overline{\omega_3}\omega_2}\right|^2\right]\left[\left(1-\left|\frac{\omega_3-\omega_1}{1-\overline{\omega_3}\omega_1}\right|^2\right)\left(1-\left|\frac{\omega_3-\omega_2}{1-\overline{\omega_3}\omega_2}\right|^2\right)\right.$$
$$\left.-\left|\frac{\alpha_1(|\omega_3|^2-1)}{(1-\overline{\omega_3}\omega_1)(1-\overline{\omega_3}\omega_2)}\right|^2\right]-\left|\frac{-\alpha_1\alpha_2\overline{\omega_3}-\beta_1(1-\overline{\omega_3}\omega_2)}{(1-\overline{\omega_3}\omega_1)(1-\overline{\omega_3}\omega_2)}\left(1-\left|\frac{\omega_3-\omega_2}{1-\overline{\omega_3}\omega_2}\right|^2\right)\right.$$
$$\left.+\frac{\alpha_1\alpha_2(1-|\omega_3|^2)(\overline{\omega_3}-\overline{\omega_2})}{(1-\overline{\omega_3}\omega_1)(1-\overline{\omega_3}\omega_2)|1-\overline{\omega_3}\omega_2|^2}\right|^2$$

and

$$G=\left(1-\left|\frac{\omega_3-\omega_2}{1-\overline{\omega_3}\omega_2}\right|^2-\left|\frac{\alpha_2}{1-\overline{\omega_3}\omega_2}\right|^2\right)\left(1-\left|\frac{\omega_3-w_4}{1-\overline{\omega_3}\omega_4}\right|^2-\left|\frac{\alpha_3}{1-\overline{\omega_3}\omega_4}\right|^2\right)$$
$$-\left|\frac{\alpha_2\alpha_3\overline{\omega_3}+\beta_2(1-|\omega_3|^2)}{(1-\overline{\omega_3}\omega_2)(1-\overline{\omega_3}\omega_4)}\right|^2$$

On the one hand, we have:

$$(1-\overline{\omega_3}\omega_1)(1-\overline{\omega_3}\omega_2)(1-\overline{\omega_3}\omega_4)|1-\overline{\omega_3}\omega_2|^2\cdot E$$
$$=\left(|\omega_3|^2-1\right)(1-\overline{\omega_3}\omega_2)\Big[\gamma\left((1-|\omega_2|^2)(1-|\omega_3|^2)-|\alpha_2|^2\right)+\alpha_1\beta_2\overline{\omega_2}(1-|\omega_3|^2)$$
$$+\alpha_3\beta_1\overline{\omega_3}(1-|\omega_2|^2)+\beta_1\beta_2\overline{\alpha_2}+\alpha_1\alpha_2\alpha_3\overline{\omega_2\omega_3}\Big]$$

On the other hand, we have:

$$\left|(1-\overline{\omega_3}\omega_2)^2(1-\overline{\omega_3}\omega_1)\right|^2\cdot F$$
$$=\left(1-|\omega_3|^2\right)^2\Big[\left((1-|\omega_1|^2)(1-|\omega_2|^2)-|\alpha_1|^2\right)\left((1-|\omega_2|^2)(1-|\omega_3|^2)-|\alpha_2|^2\right)$$
$$-\left|\alpha_1\alpha_2\overline{\omega_2}+\beta_1\left(1-|\omega_2|^2\right)\right|^2\Big]$$

and

$$|(1-\overline{\omega_3}w_2)(1-\overline{\omega_3}\omega_4)|^2\cdot G$$
$$=\left((1-|\omega_2|^2)(1-|\omega_3|^2)-|\alpha_2|^2\right)\left((1-|\omega_3|^2)(1-|\omega_4|^2)-|\alpha_3|^2\right)$$
$$-\left|\alpha_2\alpha_3\overline{\omega_3}+\beta_2\left(1-|\omega_3|^2\right)\right|^2$$

Therefore, (4.35) is equivalent to (4.17).

– If $|\alpha_2|^2=1-|\omega_2|^2$, we check that (4.21), (4.22), (4.23) and (4.24) applied to $M(T)$ are equivalent to (4.18), (4.19) and (4.20).

- We are now considering the case where $|\omega_3|=1$.

  First of all, if T is a contraction, then the compressions $\begin{bmatrix}\omega_1 & \alpha_1 & \beta_1\\ 0 & \omega_2 & \alpha_2\\ 0 & 0 & \omega_3\end{bmatrix}$ and $\begin{bmatrix}\omega_2 & \alpha_2 & \beta_2\\ 0 & \omega_3 & \alpha_3\\ 0 & 0 & \omega_4\end{bmatrix}$ are contractions. As it is a necessary condition for T to be a contraction, we will assume in the following that it is the case.

let $A = \begin{bmatrix} \omega_1 & \alpha_1 & \beta_1 \end{bmatrix}$, $B = \begin{bmatrix} \gamma \end{bmatrix}$, $C = \begin{bmatrix} 0 & \omega_2 & \alpha_2 \\ 0 & 0 & \omega_3 \\ 0 & 0 & 0 \end{bmatrix}$ and $D = \begin{bmatrix} \beta_2 \\ \alpha_3 \\ \omega_4 \end{bmatrix}$.

By assumption, $\begin{bmatrix} A \\ C \end{bmatrix}$ and $\begin{bmatrix} C & D \end{bmatrix}$ are contractions.

By Parrott's theorem, T is a contraction if and only if:

$$B = (\mathrm{Id} - ZZ^*)^{1/2} V (\mathrm{Id} - Y^*Y)^{1/2} - ZC^*Y, \text{ for some contraction } V, \tag{4.36}$$

where Y and Z are contractions such that $D = (\mathrm{Id} - CC^*)^{1/2} Y$ and $A = Z(\mathrm{Id} - C^*C)^{1/2}$.

We have: $(\mathrm{Id} - CC^*)^{1/2} = \begin{bmatrix} \sqrt{1 - |\omega_2|^2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; $(\mathrm{Id} - C^*C)^{1/2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{1 - |\omega_2|^2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$

– If $|\omega_2| < 1$:

 Applying the criterion for 3×3 matrices, the assumption that the compressions $\begin{bmatrix} \omega_1 & \alpha_1 & \beta_1 \\ 0 & \omega_2 & \alpha_2 \\ 0 & 0 & \omega_3 \end{bmatrix}$ and $\begin{bmatrix} \omega_2 & \alpha_2 & \beta_2 \\ 0 & \omega_3 & \alpha_3 \\ 0 & 0 & \omega_4 \end{bmatrix}$ are contractions means that $|\alpha_1|^2 \leq (1 - |\omega_1|^2)(1 - |\omega_2|^2)$, $\alpha_2 = \alpha_3 = 0$, $|\beta_1(1 - |\omega_2|^2)|^2 \leq (1 - |\omega_1|^2)(1 - |\omega_2|^2) - |\alpha_1|^2$ and $|\beta_2|^2 \leq (1 - |\omega_2|^2)(1 - |\omega_4|^2)$.

 Moreover, the equations $D = (\mathrm{Id} - CC^*)^{1/2} Y$ and $A = Z(\mathrm{Id} - C^*C)^{1/2}$ have both infinitely many solutions. Taking the minimal solutions for Y and Z, we obtain $Y = \begin{bmatrix} \frac{\beta_2}{\sqrt{1 - |\omega_2|^2}} \\ 0 \\ \omega_4 \end{bmatrix}$ and $Z = \begin{bmatrix} \omega_1 & \frac{\alpha_1}{\sqrt{1 - |\omega_2|^2}} & 0 \end{bmatrix}$

 Hence, we have $ZC^*Y = \frac{\alpha_1 \beta_2 \overline{\omega_2}}{1 - |\omega_2|^2}$ and, in the end, T is a contraction if and only if:

$$\left| \gamma + \frac{\alpha_1 \beta_2 \overline{\omega_2}}{1 - |\omega_2|^2} \right|^2 \leq \left[1 - |\omega_4|^2 - \frac{|\beta_2|^2}{1 - |\omega_2|^2} \right] \left[1 - |\omega_1|^2 - \frac{|\alpha_1|^2}{1 - |\omega_2|^2} \right],$$

 which is equivalent to

$$\left| \gamma(1 - |\omega_2|^2) + \alpha_1 \beta_2 \overline{\omega_2} \right|^2 \leq \left[(1 - |\omega_4|^2)(1 - |\omega_2|^2) - |\beta_2|^2 \right] \left[(1 - |\omega_1|^2)(1 - |\omega_2|^2) - |\alpha_1|^2 \right].$$

– If $|\omega_2| = 1$:

 Applying the criterion for 3×3 matrices, the assumption that the compressions $\begin{bmatrix} \omega_1 & \alpha_1 & \beta_1 \\ 0 & \omega_2 & \alpha_2 \\ 0 & 0 & \omega_3 \end{bmatrix}$ and $\begin{bmatrix} \omega_2 & \alpha_2 & \beta_2 \\ 0 & \omega_3 & \alpha_3 \\ 0 & 0 & \omega_4 \end{bmatrix}$ are contractions means that $\alpha_1 = \alpha_2 = \alpha_3 = 0$, $|\beta_1|^2 \leq (1 - |\omega_1|^2)(1 - |\omega_3|^2)$

and $|\beta_2|^2 \leq (1-|\omega_2|^2)(1-|\omega_4|^2)$. Moreover, taking the minimal solutions for Y and Z, we obtain $Y = \begin{bmatrix} 0 \\ 0 \\ \omega_4 \end{bmatrix}$ and $Z = \begin{bmatrix} \omega_1 & 0 & 0 \end{bmatrix}$ Hence, we have $ZC^*Y = 0$ and, in the end, T is a contraction if and only if:

$$\left|\gamma\right|^2 \leq (1-|\omega_1|^2)(1-|\omega_4|^2)$$

- The case where $|\omega_2| = 1$ and $|\omega_3| < 1$ is similar to the case where $|\omega_3| = 1$ and $|\omega_2| < 1$.

□

4.4 Extremal contractive matrices and model space operators

4.4.1 A noteworthy observation: link with model matrices

Recall from Section 2.3.6 that, if Θ_n is a finite Blaschke product with distinct zeros $\omega_1, \ldots, \omega_n \in \mathbb{D}$, the matrix representation M_{Θ_n} of the compressed shift S_{Θ_n} in the Takenaka-Malmquist-Walsh basis is given by:

$$\left[M_{\Theta_n}\right]_{i,j} = \begin{cases} \omega_j & \text{if } i{=}j \\ \prod_{k=i+1}^{j-1}(-\overline{\omega_k})\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_j|^2} & \text{if } i{<}j \\ 0 & \text{if } i{>}j \end{cases} \tag{4.37}$$

For example, for $n = 2$ and $n = 3$, we obtain the following matrices:

$$M_{\Theta_2} := \begin{pmatrix} \omega_1 & \sqrt{1-|\omega_1|^2}\sqrt{1-|\omega_2|^2} \\ 0 & \omega_2 \end{pmatrix}$$

$$M_{\Theta_3} := \begin{pmatrix} \omega_1 & \sqrt{1-|\omega_1|^2}\sqrt{(1-|\omega_2|^2} & -\overline{\omega_2}\sqrt{(1-|\omega_1|^2}\sqrt{(1-|\omega_3|^2)} \\ 0 & \omega_2 & \sqrt{1-|\omega_2|^2}\sqrt{(1-|\omega_3|^2} \\ 0 & 0 & \omega_3 \end{pmatrix}.$$

Those two matrices can be seen as extremal matrices for the criteria of contractivity for (respectively) 2×2 matrices and 3×3 matrices.

Indeed, let us recall that a $2{\times}2$ matrix of the form $T_2 = \begin{pmatrix} \omega_1 & \alpha \\ 0 & \omega_2 \end{pmatrix}$ (with $\omega_1, \omega_2 \in \mathbb{D}$) is a contraction if and only if $|\alpha| \leq \sqrt{1-|\omega_1|^2}\sqrt{1-|\omega_2|^2}$, and that a 3×3 matrix of the form $T_3 = \begin{pmatrix} \omega_1 & \alpha_1 & \beta \\ 0 & \omega_2 & \alpha_2 \\ 0 & 0 & \omega_3 \end{pmatrix}$ (with $\omega_1, \omega_2, \omega_3 \in \mathbb{D}$) is a contraction if and only if

$$\begin{cases} |\alpha_i| \leq \sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+1}|^2},\ i = 1, 2 \\ \left|\beta(1-|\omega_2|^2) + \alpha_1\alpha_2\overline{\omega_2}\right| \leq \sqrt{(1-|\omega_1|^2)(1-|\omega_2|^2) - |\alpha_1|^2} \cdot \sqrt{(1-|\omega_2|^2)(1-|\omega_3|^2) - |\alpha_2|^2} \end{cases}.$$

If we set $\alpha_1 = \sqrt{1-|\omega_1|^2}\sqrt{1-|\omega_2|^2}$ and $\alpha_2 = \sqrt{1-|\omega_2|^2}\sqrt{1-|\omega_3|^2}$, then, T_3 is a contraction if and only if $T_3 = M_{\Theta_3}$.

Similarly, observe that, in Theorem 4.3.1, if we set $\alpha_i = \sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+1}|^2}$, for $i = 1, 2, 3$, then, $T = T_4$ is a contraction if and only if:

$$\begin{cases} \beta_i = -\overline{\omega_{i+1}}\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+2}|^2}, & i = 1,2 \\ \gamma = \overline{\omega_2\omega_3}\sqrt{1-|\omega_1|^2}\sqrt{1-|\omega_4|^2} \end{cases},$$

which is equivalent to say that $T_4 = M_{\Theta_4}$.

The aim of the next section is to generalize those observations.

4.4.2 Extremal contractive matrices of arbitrary size

We will now prove the following result:

Theorem 4.4.1. *Let $n \in \mathbb{N}^*$, let $\omega_1, \ldots, \omega_n \in \mathbb{D}$ and let*

$$T_n = \begin{pmatrix} \omega_1 & \alpha_1^{(1)} & \alpha_1^{(2)} & \cdots & \alpha_1^{(n-1)} \\ 0 & \omega_2 & \alpha_2^{(1)} & \alpha_2^{(2)} & \vdots \\ \vdots & \ddots & \ddots & \ddots & \alpha_{n-2}^{(2)} \\ \vdots & & \ddots & \ddots & \alpha_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & \omega_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{C}).$$

Assume that

$$\alpha_i^{(1)} = \sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+1}|^2}, \text{ for all } 1 \le i \le n-1.$$

Then, T_n is a contraction if and only if $T_n = M_{\Theta_n}$, where M_{Θ_n} is defined as in (4.37).

In other words, T_n is a contraction if and only if

$$\alpha_i^{(j)} = \prod_{k=i+1}^{j+i-1} (-\overline{\omega_k})\sqrt{1-|w_i|^2}\sqrt{1-|\omega_{i+j}|^2}, \text{ for all } 1 \le i \le n,\ 1 \le j \le n-i.$$

Proof. First of all, note that if $T_n = \left(t_{i,j}\right)_{1\le i,j\le n}$, then we have $\alpha_i^{(j)} = t_{i,i+j}$. Now, we will proceed by induction on the size of the matrix T_n.

- For $n = 2$, there is nothing to prove.
- Let $n \ge 3$, and assume that the result is true for a matrix of size $n-1$.

  If T_n is a contraction, then, the two compressions

$$S_n^{(1)} = \begin{bmatrix} \omega_1 & \alpha_1^{(1)} & \alpha_1^{(2)} & \cdots & \alpha_1^{(n-2)} \\ 0 & \omega_2 & \alpha_2^{(1)} & \cdots & \alpha_2^{(n-3)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \alpha_{n-2}^{(1)} \\ 0 & \cdots & \cdots & 0 & \omega_{n-1} \end{bmatrix}$$

and

$$S_n^{(2)} = \begin{bmatrix} \omega_2 & \alpha_2^{(1)} & \alpha_2^{(2)} & \cdots & \alpha_2^{(n-2)} \\ 0 & \omega_3 & \alpha_3^{(1)} & \cdots & \alpha_3^{(n-3)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \alpha_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & \omega_n \end{bmatrix}$$

are contractions too. Using the induction hypothesis, that happens if and only if

$$\alpha_i^{(j)} = \prod_{k=i+1}^{j+i-1} (-\overline{\omega_k})\sqrt{1-|w_i|^2}\sqrt{1-|\omega_{i+j}|^2}$$

for all i, j, except maybe for $\alpha_1^{(n-1)}$ (who does not appear in any of those two compressions). Our aim is now to prove that there exists a unique choice for the coefficient $\alpha_1^{(n-1)}$ such that T_n is a contraction. As we know that for

$$a_1^{(n-1)} = \prod_{k=2}^{n-1} (-\overline{\omega_k})\sqrt{1-|w_1|^2}\sqrt{1-|\omega_n|^2}.$$

$T_n = M_{\Theta_n}$ is a contraction, that will be enough to prove the theorem.

Let $A = \begin{bmatrix} \omega_1 & \alpha_1^{(1)} & \alpha_1^{(2)} \cdots \alpha_1^{(n-2)} \end{bmatrix}$, $B = \begin{bmatrix} \alpha_1^{(n-1)} \end{bmatrix}$, $C = \begin{bmatrix} 0 & \omega_2 & \alpha_2^{(1)} & \cdots & \alpha_2^{(n-3)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \alpha_{n-2}^{(1)} \\ \vdots & & & \ddots & \omega_{n-1} \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix}$ and $D = \begin{bmatrix} \alpha_2^{(n-2)} \\ \vdots \\ \alpha_{n-1}^{(1)} \\ \omega_n \end{bmatrix}$.

As T_n is a contraction, then, $\begin{bmatrix} A \\ C \end{bmatrix}$ and $\begin{bmatrix} C & D \end{bmatrix}$ are contractions too. By Parrott's theorem, T is a contraction if and only if there exists a contraction V such that:

$$B = (\mathrm{Id} - ZZ^*)^{1/2} V (\mathrm{Id} - Y^*Y)^{1/2} - ZC^*Y, \tag{4.38}$$

where Y and Z are contractions such that $D = (\mathrm{Id} - CC^*)^{1/2} Y$ and $A = Z(\mathrm{Id} - C^*C)^{1/2}$.

Note that (4.38) can be rewritten:

$$\left|\alpha_1^{(n-1)} + ZC^*Y\right|^2 \leq (1 - ZZ^*) \times (1 - Y^*Y) \tag{4.39}$$

Fact 1. If $\omega_2 = 0$, then $C^*C = \begin{array}{c} \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \ddots & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} \begin{array}{l} \updownarrow 2 \\ \\ \updownarrow n-3 \\ \\ \end{array} \\ \underset{2}{\longleftrightarrow} \qquad \underset{n-3}{\longleftrightarrow} \end{array}$

Proof. Denote $C = (c_{i,j})_{1\le i,j\le n}$, with $c_{i,j} = \begin{cases} 0 & \text{if } j \le i \\ \omega_j & \text{if } j = i+1 \\ \alpha_{i+1}^{(j-i-1)} & \text{if } j \ge i+2 \end{cases}$.

We have $[C^*C]_{i,j} = \sum_{k=1}^{n-1} \overline{c_{k,i}} c_{k,j} = \sum_{k=1}^{m(i,j)} \overline{c_{k,i}} c_{k,j}$, where $m(i,j) = \min\{i-1, j-1\}$.

We can easily notice that $[C^*C]_{1,j} = 0$, for all $j \in [\![1, n-1]\!]$ and $[C^*C]_{i,1} = 0$, for all $i \in [\![1, n-1]\!]$ (which means that the first row and the first column of C^*C are zero).

Moreover, it is easy to see that $[C^*C]_{2,2} = |\omega_2|^2 = 0$ and that $[C^*C]_{2,j} = \overline{\omega_2}\alpha_2^{(j-2)} = 0$, for all $j \in [\![3, n-1]\!]$.

Furthermore, for $i \in [\![3, n-1]\!]$, we get:

$$\begin{aligned}
[C^*C]_{i,i} &= \sum_{k=1}^{i-2} \left|\alpha_{k+1}^{(i-k-1)}\right|^2 + |\omega_i|^2 \\
&= (1-|\omega_i|^2) \sum_{k=1}^{i-2} \left(\prod_{s=k+2}^{i-1} |\omega_s|^2 \right) (1-|\omega_{k+1}|^2) + |\omega_i|^2 \\
&= (1-|\omega_i|^2) \left(\sum_{k=1}^{i-2} \prod_{s=k+2}^{i-1} |\omega_s|^2 - \sum_{k=1}^{i-2} \prod_{s=k+1}^{i-1} |\omega_s|^2 \right) + |\omega_i|^2.
\end{aligned}$$

Now, we make a change of indices in order to make a telescopic sum appear:

$$\begin{aligned}
[C^*C]_{i,i} &= (1-|\omega_i|^2) \left(\sum_{k=3}^{i} \prod_{s=k}^{i-1} |\omega_s|^2 - \sum_{k=2}^{i-1} \prod_{s=k}^{i-1} |\omega_s|^2 \right) + |\omega_i|^2 \\
&= (1-|\omega_i|^2) \left(1 - \prod_{s=2}^{i-1} |\omega_s|^2 \right) + |\omega_i|^2 \\
&= 1,
\end{aligned}$$

as the first term of the product $\prod_{s=2}^{i-1} |\omega_s|^2$ is equal to 0.

Similarly, for $i \in [\![3, n-1]\!]$ and $j > i$, we get:

$$\begin{aligned}
[C^*C]_{i,j} &= \sum_{k=1}^{i-2} \overline{\alpha_{k+1}^{(i-k-1)}} \alpha_{k+1}^{(j-k-1)} + \overline{\omega_i} \alpha_i^{(j-i)} \\
&= \sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_j|^2} \left[\sum_{k=1}^{i-1} \left((1-|\omega_{k+1}|^2) \prod_{s=k+2}^{i-1} (-\omega_s) \prod_{s=k+2}^{j-1} (-\overline{\omega_s}) \right) - \prod_{s=i}^{j-1} (-\overline{\omega_s}) \right] \\
&= \sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_j|^2} \left[\sum_{k=1}^{i-1} \left((1-|\omega_{k+1}|^2) \prod_{s=k+2}^{i-1} |\omega_s|^2 \prod_{s=i}^{j-1} (-\overline{\omega_s}) \right) - \prod_{s=i}^{j-1} (-\overline{\omega_s}) \right] \\
&= \sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_j|^2} \prod_{s=i}^{j-1} (-\overline{\omega_s}) \left[\sum_{k=3}^{i+1} \prod_{s=k}^{i-1} |\omega_s|^2 - \sum_{k=2}^{i} \prod_{s=k}^{i-1} |\omega_s|^2 - 1 \right]
\end{aligned}$$

$$= \sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_j|^2}\prod_{s=i}^{j-1}(-\overline{\omega_s})\left[1-\prod_{s=2}^{i-1}|\omega_s|^2-1\right]$$
$$= 0.$$

Finally, we conclude using the self-adjointness of C^*C. □

Fact 2. Suppose $\omega_2 = 0$. Then $\|T_n\| \le 1$ if and only if $\alpha_1^{(n-1)} = 0$.

Proof. If $\omega_2 = 0$, we have

$$\mathrm{Id}-C^*C = \begin{array}{cl} \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & & \vdots \\ \vdots & \ddots & 0 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 0 \end{bmatrix} & \begin{array}{l} \updownarrow 2 \\ \updownarrow n-3 \end{array} \\ \underset{2}{\longleftrightarrow} \qquad \underset{n-3}{\longleftrightarrow} & \end{array}.$$

Denote $Z = \begin{bmatrix} z_1 & \dots & z_{n-1} \end{bmatrix}$. The identity $A = Z(\mathrm{Id}-C^*C)^{1/2}$ is equivalent to $\begin{cases} z_1 = \omega_1 \\ z_2 = \sqrt{1-|\omega_1|^2} \end{cases}$.
If we further require that $Z = 0$ on $\mathrm{Im}(D_C)^{\perp}$ (which is a necessary condition for Z to remain a contraction), we obtain:

$$z_3 = \cdots = z_{n-1} = 0$$

Finally, we have:

$$ZZ^* = 1$$

and

$$ZC^* = \begin{bmatrix} 0 & \dots & 0 \end{bmatrix}.$$

Therefore (4.39) holds if and only if:

$$\alpha_1^{(n-1)} = -ZC^*Y = 0.$$

The proof of Fact 2 is complete. □

The general case can be reduced to the case where $\omega_2 = 0$ by using the fact that T_n is a contraction if and only if $M(T_n)$ is a contraction, where $M = M_{\omega_2} : z \mapsto \frac{\omega_2 - z}{1-\overline{\omega_2}z}$ (Lemma 4.3.7). Using the rational functional calculus, we can write:

$$M(T_n) = \begin{pmatrix} M(\omega_1) & \beta_1^{(1)} & * & \cdots & \cdots & * \\ 0 & M(\omega_2) & \beta_2^{(2)} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & * \\ \vdots & & \ddots & \ddots & \ddots & \beta_{n-1}^{(1)} \\ 0 & \cdots & \cdots & \cdots & 0 & M(\omega_n) \end{pmatrix},$$

where $\beta_i^{(1)} = \alpha_i^{(1)}[M(\omega_i), M(\omega_{i+1})] = \frac{\alpha_i^{(1)}(|\omega_2|^2-1)}{(1-\overline{\omega_2}\omega_i)(1-\overline{\omega_2}\omega_{i+1})}$.

Assume that

$$\alpha_i^{(1)} = \sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_{i+1}|^2},$$

for all $i \in [\![1, n-1]\!]$. Then, for all $i \in [\![1, n-1]\!]$, we have:

$$\beta_i^{(1)} = e^{i\phi_i}\sqrt{1-|M(\omega_i)|^2}\sqrt{1-|M(\omega_{i+1})|^2}, \quad \text{for some } \phi_i \in [0, 2\pi[.$$

Set $\theta_i = \phi_1 + \cdots + \phi_i$ and $U = \operatorname{diag}(e^{i\theta_1}, \dots, e^{i\theta_{n-1}}, 1)$. Then U is unitary and

$$U^* M(T_n) U = \begin{pmatrix} M(\omega_1) & \gamma_1^{(1)} & * & \cdots & * \\ 0 & M(\omega_2) & \gamma_2^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & * \\ \vdots & & \ddots & \ddots & \gamma_{n-1}^{(1)} \\ 0 & \cdots & \cdots & 0 & M(\omega_n) \end{pmatrix},$$

where $\gamma_i^{(1)} = \sqrt{1-|M(\omega_i)|^2}\sqrt{1-|M(\omega_{i+1})|^2}$.

The operator $M(T_n)$ is a contraction if and only if $U^*M(T_n)U$ is a contraction and, from the previous reasoning we know that if U^*T_nU is a contraction, then, all its coefficients are uniquely determined. If all the coefficients of U^*T_nU are uniquely determined, then, it is also the case for $M(T_n)$ and for $T_n = M \circ M(T_n)$. $\square$

Chapter 5

Schwarz-Pick type inequalities from an operator theoretical point of view

In this chapter, we use several versions of von Neumann's inequality to prove various Schwarz-Pick type inequalities.

5.1 Three points Schwarz-Pick type inequalities

5.1.1 An operator theoretical proof of Beardon-Minda's inequality

First of all, let us recall the following notation (see *e.g.* [11, Chapter 6] for more details):

Definition 5.1.1. Let $z, w \in \mathbb{D}$ and $f \in \mathcal{H}(\mathbb{D}, \mathbb{D})$. We define:

(i) The complex pseudo-hyperbolic distance $(z, w) := \frac{z-w}{1-\overline{w}z}$;

(ii) The pseudo-hyperblic distance $\rho(z, w) := |(z, w)|$;

(iii) The hyperbolic distance $\mathrm{d}(z, w) = \tanh^{-1}(\rho(z, w))$;

(iv) The hyperbolic divided difference $f^*(z, w) := \frac{(f(z), f(w))}{(z, w)}$.

We provide now an operator-theoretic proof of the following result established by Beardon and Minda [8]:

Theorem 5.1.2 (Beardon-Minda)**.** *Let $f \in \mathcal{H}(\mathbb{D}, \mathbb{D})$ and let ω_1, ω_2 and ω_3 be pairwise distinct points in $\mathbb{D}$. Then,*

$$\mathrm{d}\left(f^*(\omega_1, \omega_2), f^*(\omega_3, \omega_2)\right) \leq \mathrm{d}(\omega_1, \omega_3). \tag{5.1}$$

The proof in [8] requires an assumption that f is not a conformal automorphism of the unit disk. Such an assumption is unnecessary in the subsequent proof.

Proof of Theorem 5.1.2. Let us first notice that Beardon-Minda's inequality (5.1) is equivalent to:

$$\left|\frac{f^*(\omega_1,\omega_2)-f^*(\omega_3,\omega_2)}{1-\overline{f^*(\omega_3,\omega_2)}f^*(\omega_1,\omega_2)}\right| \le \left|\frac{\omega_1-\omega_3}{1-\overline{\omega_3}\omega_1}\right|, \tag{5.2}$$

and that, for $z,\omega \in \mathbb{D}$, we have:

$$f^*(z,\omega) = \frac{f(z)-f(\omega)}{z-\omega}\cdot\frac{1-\overline{\omega}z}{1-\overline{f(\omega)}f(z)}. \tag{5.3}$$

For $u,v \in \mathbb{C}$, we write

$$S_{u,v} := (1-|u|^2)(1-|v|^2) = |1-\overline{u}v|^2 - |u-v|^2. \tag{5.4}$$

Now, let $\omega_1,\omega_2,\omega_3 \in \mathbb{D}$, with $\omega_i \neq \omega_j$ $(i \neq j)$, and consider

$$T = \begin{pmatrix} \omega_1 & \alpha_1 & \beta \\ 0 & \omega_2 & \alpha_2 \\ 0 & 0 & \omega_3 \end{pmatrix},$$

with

$$\alpha_i = \sqrt{1-|\omega_i|^2}\sqrt{(1-|\omega_{i+1}|^2}, \qquad i=1,2,$$

and

$$\beta = \frac{-\overline{\omega_2}\alpha_1\alpha_2}{1-|\omega_2|^2} = -\overline{\omega_2}\sqrt{1-|\omega_1|^2}\sqrt{1-|\omega_3|^2}.$$

Notice that T is the model matrix of size 3 with diagonal entries $\omega_1,\omega_2,\omega_3$.

Our strategy will be the same as for the proof of the Schwarz-Pick lemma (Theorem 1.2.13).

Firstly, by Theorem 4.2.2, T is a contraction. Thus, applying the von Neumann inequality, we get that the operator $f(T)$ remains a contraction, for every function $f \in \mathcal{A}(\overline{\mathbb{D}})$ such that $\|f\|_\infty \le 1$. Then, using the explicit functional calculus of Corollary 3.2.26, the matrix representation of $f(T)$ can be expressed in terms of first order and second order divided differences as follows:

$$f(T) = \begin{pmatrix} f(\omega_1) & \alpha_1[f(\omega_1),f(\omega_2)] & \beta[f(\omega_1),f(\omega_3)]+\alpha_1\alpha_2[f(\omega_1),f(\omega_2),f(\omega_3)] \\ 0 & f(\omega_2) & \alpha_2[f(\omega_2),f(\omega_3)] \\ 0 & 0 & f(\omega_3) \end{pmatrix}.$$

Assume that $f(\omega_i) \neq f(\omega_j)$ whenever $i \neq j$ (otherwise, there is nothing to prove).

In order to simplify the computations, we introduce the following notation:

$$\widetilde{\alpha_i} = \alpha_i[f(\omega_i),f(\omega_{i+1})], \qquad i=1,2,$$
$$\widetilde{\beta} = \beta[f(\omega_1),f(\omega_3)]+\alpha_1\alpha_2[f(\omega_1),f(\omega_2),f(\omega_3)].$$

Using Theorem 4.2.2, the fact that $f(T)$ is a contraction provides the following inequality:

$$\left|\widetilde{\beta}\left(1-|f(\omega_2)|^2\right)+\widetilde{\alpha_1}\widetilde{\alpha_2}\overline{f(\omega_2)}\right|^2 \le \left[S_{f(\omega_1),f(\omega_2)}-|\widetilde{\alpha_1}|^2\right]\times\left[S_{f(\omega_2),f(\omega_3)}-|\widetilde{\alpha_2}|^2\right].$$

Now, our aim will be to prove that this inequality is equivalent to (5.2).

If we multiply each side of this inequality by $|\omega_1 - \omega_3|^2$, we get:

$$S_{\omega_1,\omega_3}\left|\left(1-|f(\omega_2)|^2\right)A+B\right|^2 \leq |\omega_1-\omega_3|^2 C_1 C_3, \tag{5.5}$$

where

$$\begin{aligned}
A &:= -\overline{\omega_2}(f(\omega_1)-f(\omega_3)) + \left(1-|\omega_2|^2\right)\left(\frac{f(\omega_1)-f(\omega_2)}{\omega_1-\omega_2} - \frac{f(\omega_2)-f(\omega_3)}{\omega_2-\omega_3}\right),\\
B &:= \overline{f(\omega_2)}(1-|\omega_2|^2)(\omega_1-\omega_3)\cdot\frac{(f(\omega_1)-f(\omega_2))(f(\omega_2)-f(\omega_3))}{(\omega_1-\omega_2)(\omega_2-\omega_3)},\\
C_i &:= S_{f(\omega_i),f(\omega_2)} - S_{\omega_i,\omega_2}\left|\frac{f(\omega_i)-f(\omega_2)}{\omega_i-\omega_2}\right|^2, \qquad i=1,3.
\end{aligned}$$

We want to prove that (5.5) is equivalent to (5.2). The calculations are somewhat laborious; the key idea is to use (5.3) to make hyperbolic divided differences appear each time we see an expression of the form $f(z)-f(\omega)$. We provide additional details to assist the reader.

On the one hand, we have:

$$\begin{aligned}
C_i &= S_{f(\omega_i),f(\omega_2)} - S_{\omega_i,\omega_2}|f^*(\omega_i,\omega_2)|^2 \times \left|\frac{1-\overline{f(\omega_2)}f(\omega_i)}{1-\overline{\omega_2}\omega_i}\right|^2\\
&= \left|1-\overline{f(\omega_2)}f(\omega_i)\right|^2 - |f^*(\omega_i,\omega_2)|^2 \times |\omega_i-\omega_2|^2 \times \left|\frac{1-\overline{f(\omega_2)}f(\omega_i)}{1-\overline{\omega_2}\omega_i}\right|^2\\
&\quad - |f^*(\omega_i,\omega_2)|^2 \times \left|1-\overline{f(\omega_2)}f(\omega_i)\right|^2 + |f^*(\omega_i,\omega_2)|^2 \times |\omega_i-\omega_2|^2 \times \left|\frac{1-\overline{f(\omega_2)}f(\omega_i)}{1-\overline{\omega_2}\omega_i}\right|^2\\
&= \left|1-\overline{f(\omega_2)}f(\omega_i)\right|^2\left(1-|f^*(\omega_i,\omega_2)|^2\right).
\end{aligned}$$

Thus, we have:

$$C_1C_3 = \left|1-\overline{f(\omega_2)}f(\omega_1)\right|^2 \times \left|1-\overline{f(\omega_2)}f(\omega_3)\right|^2 \times S_{f^*(\omega_1,\omega_2),f^*(\omega_3,\omega_2)}.$$

Now, let us deal with the first member of the inequality. We have:

$$\begin{aligned}
A &= -\overline{\omega_2}f^*(\omega_1,\omega_2)\left(\frac{1-\overline{f(\omega_2)}f(\omega_1)}{1-\overline{\omega_2}\omega_1}\right)(\omega_1-\omega_2) + \overline{\omega_2}f^*(\omega_3,\omega_2)\left(\frac{1-\overline{f(\omega_2)}f(\omega_3)}{1-\overline{\omega_2}\omega_3}\right)(\omega_3-\omega_2)\\
&\quad + \left(1-|\omega_2|^2\right)\left[f^*(\omega_1,\omega_2)\left(\frac{1-\overline{f(\omega_2)}f(\omega_1)}{1-\overline{\omega_2}\omega_1}\right) - f^*(\omega_3,\omega_2)\left(\frac{1-\overline{f(\omega_2)}f(\omega_3)}{1-\overline{\omega_2}\omega_3}\right)\right]\\
&= f^*(\omega_1,\omega_2)\left(1-\overline{f(\omega_2)}f(\omega_1)\right) - f^*(\omega_3,\omega_2)\left(1-\overline{f(\omega_2)}f(\omega_3)\right)\\
&= (f^*(\omega_1,\omega_2)-f^*(\omega_3,\omega_2))\left(1-\overline{f(\omega_2)}f(\omega_1)\right)\left(1-\overline{f(\omega_2)}f(\omega_3)\right)\\
&\quad + \overline{f(\omega_2)}\left[f^*(\omega_1,\omega_2)f(\omega_3)\left(1-\overline{f(\omega_2)}f(\omega_1)\right) - f^*(\omega_3,\omega_2)f(\omega_1)\left(1-\overline{f(\omega_2)}f(\omega_3)\right)\right]\\
&= (f^*(\omega_1,\omega_2)-f^*(\omega_3,\omega_2))\left(1-\overline{f(\omega_2)}f(\omega_1)\right)\left(1-\overline{f(\omega_2)}f(\omega_3)\right) + \overline{f(\omega_2)}D,
\end{aligned}$$

where $D := f^*(\omega_1,\omega_2)f(\omega_3)\big(1-\overline{f(\omega_2)}f(\omega_1)\big) - f^*(\omega_3,\omega_2)f(\omega_1)\big(1-\overline{f(\omega_2)}f(\omega_3)\big)$.

This term can be written as follows, where we make appear the differences $f(\omega_3)-f(\omega_2)$ and $f(\omega_1)-f(\omega_2)$:

$$\begin{aligned} D = {} & f^*(\omega_1,\omega_2)(f(\omega_3)-f(\omega_2))\big(1-\overline{f(\omega_2)}f(\omega_1)\big) \\ & - f^*(\omega_3,\omega_2)(f(\omega_1)-f(\omega_2))\big(1-\overline{f(\omega_2)}f(\omega_3)\big) + f(\omega_2)f^*(\omega_1,\omega_2)\big(1-\overline{f(\omega_2)}f(\omega_1)\big) \\ & \qquad - f(\omega_2)f^*(\omega_3,\omega_2)\big(1-\overline{f(\omega_2)}f(\omega_3)\big). \end{aligned}$$

We obtain

$$\begin{aligned} D &= f^*(\omega_1,\omega_2)f^*(\omega_3,\omega_2)\left(\frac{1-\overline{f(\omega_2)}f(\omega_3)}{1-\overline{\omega_2}\omega_3}\right)(\omega_3-\omega_2)\big(1-\overline{f(\omega_2)}f(\omega_1)\big) \\ &\quad - f^*(\omega_3,\omega_2)f^*(\omega_1,\omega_2)\left(\frac{1-\overline{f(\omega_2)}f(\omega_1)}{1-\overline{\omega_2}\omega_1}\right)(\omega_1-\omega_2)\big(1-\overline{f(\omega_2)}f(\omega_3)\big) + f(\omega_2)A \\ &= f^*(\omega_1,\omega_2)f^*(\omega_3,\omega_2)\big(1-\overline{f(\omega_2)}f(\omega_1)\big)\big(1-\overline{f(\omega_2)}f(\omega_3)\big)\cdot\frac{(1-|\omega_2|^2)(\omega_3-\omega_1)}{(1-\overline{\omega_2}\omega_1)(1-\overline{\omega_2}\omega_3)} + f(\omega_2)A \\ &= -(1-|\omega_2|^2)(\omega_1-\omega_3)\cdot\frac{(f(\omega_1)-f(\omega_2))(f(\omega_2)-f(\omega_3))}{(\omega_1-\omega_2)(\omega_2-\omega_3)} + f(\omega_2)A. \end{aligned}$$

Hence, we get

$$A = (f^*(\omega_1,\omega_2)-f^*(\omega_3,\omega_2))\big(1-\overline{f(\omega_2)}f(\omega_1)\big)\big(1-\overline{f(\omega_2)}f(\omega_3)\big) - B + |f(\omega_2)|^2A.$$

Therefore

$$(1-|f(\omega_2)|^2)A + B = (f^*(\omega_1,\omega_2)-f^*(\omega_3,\omega_2))\big(1-\overline{f(\omega_2)}f(\omega_1)\big)\big(1-\overline{f(\omega_2)}f(\omega_3)\big).$$

Combining all of these elements, the inequality represented by (5.5) transforms into

$$S_{\omega_1,\omega_3}|f^*(\omega_1,\omega_2)-f^*(\omega_3,\omega_2)|^2 \le |\omega_1-\omega_3|^2 \times S_{f^*(\omega_1,\omega_2),f^*(\omega_3,\omega_2)},$$

which is equivalent to (5.2).

Beardon-Minda's inequality is thus proved for $f\in\mathcal{A}(\overline{\mathbb{D}})$. Now, for $f\in\mathcal{H}(\mathbb{D})$, we have $f_r : z\mapsto f(rz)\in\mathcal{A}(\overline{\mathbb{D}})$, for every $r\in]0,1[$. Based on the preceding information, it can be concluded that Beardon-Minda's inequality is satisfied by the functions f_r, for all $r\in]0,1[$, so it is also by f, by letting $r\to 1^-$. $\square$

Remark 5.1.3. The calculations in this proof can be somewhat simplified by assuming that $f(\omega_2)=0$ and composing with a Möbius transformation at the end. However, this approach leads to a loss of symmetry in the formulas.

Remark 5.1.4. In [8], it is further proved that if f does not represent an automorphism of the unit disk, equality holds in Theorem 5.1.2 if and only if f is a Blaschke product of degree no greater than 2. This inference can also be derived through operator theory considerations. We will come back to this point in Section 5.2.

5.1.2 A Beardon-Minda type lemma for derivatives

We now investigate the case where $\omega_1 = \omega_2 = \omega_3 =: \omega$. For a holomorphic function f we use the notation

$$\Gamma(z, f) = \frac{(1-|z|^2)|f'(z)|}{1-|f(z)|^2}.$$

The Schwarz-Pick inequality for derivatives (1.2) can then be expressed as $|\Gamma(z,f)| \leq 1$.

We give now an operator theoretical proof of the following result, proved by Yamashita in [65, Theorem 2].

Theorem 5.1.5. *Let $f \in \mathcal{H}(\mathbb{D},\mathbb{D})$ and let $\Gamma(z,f) = \frac{(1-|z|^2)|f'(z)|}{1-|f(z)|^2}$. Then, for every $\omega \in \mathbb{D}$,*

$$\left|\frac{\partial \Gamma(\omega, f)}{\partial \omega}\right| \leq \frac{1-|\Gamma(\omega,f)|^2}{1-|\omega|^2}. \tag{5.6}$$

Moreover, equality holds if and only if f is a Blaschke product of degree ≤ 2.

Proof. Let $\omega \in \mathbb{D}$, and let $T = \begin{pmatrix} \omega & \alpha & \beta \\ 0 & \omega & \alpha \\ 0 & 0 & \omega \end{pmatrix} \in \mathcal{M}_3(\mathbb{C})$, with $\alpha = 1-|\omega|^2$ and $\beta = -\overline{\omega}(1-|\omega|^2)$. By Theorem 4.2.2, T is a contraction. Moreover, we can easily check that for f in the disk algebra we have

$$f(T) = \begin{pmatrix} f(\omega) & \alpha f'(\omega) & \frac{1}{2}\alpha^2 f''(\omega) + \beta f'(\omega) \\ 0 & f(\omega) & \alpha f'(\omega) \\ 0 & 0 & f(\omega) \end{pmatrix}.$$

In this representation, the divided differences have been replaced in this limit case by first and second-order derivatives. By von Neumann's inequality, $f(T)$ is a contraction. Using Theorem 4.2.2 we obtain:

$$\left|\left(\frac{1}{2}\alpha^2 f''(\omega) + \beta f'(\omega)\right)\left(1-|f(\omega)|^2\right) + \alpha^2 f'(\omega)^2 \overline{f(\omega)}\right| \leq \left(1-|f(\omega)|^2\right)^2 - \left|\alpha f'(\omega)\right|^2,$$

which is equivalent to (5.6). A proof of the equality case can be obtained using model spaces, as discussed in the preceding subsection. □

The inequality (5.6) can be rephrased in terms of *Peschl's invariant derivatives*. Let $f \in \mathcal{H}(\mathbb{D},\mathbb{D})$, let $\omega \in \mathbb{D}$, and consider the mapping

$$g : z \in \mathbb{D} \mapsto \frac{f\left(\frac{z+\omega}{1+\overline{\omega}z}\right) - f(\omega)}{1-\overline{f(\omega)} f\left(\frac{z+\omega}{1+\overline{\omega}z}\right)} \in \mathbb{C}. \tag{5.7}$$

Then g is analytic on $\mathbb{D}$ and $g(0) = 0$. We have $g(z) = \sum_{n=1}^{\infty} \frac{D_n f(\omega)}{n!} z^n$, with $D_n f(z_0) := g^{(n)}(0)$. The quantities $D_n f(\omega)$ are called *Peschl's invariant derivatives* (see *e.g.* [37]).

The first two values of Peschl's invariant derivatives are explicitely computed as:

$$D_1 f(\omega) = \frac{(1-|\omega|^2)f'(\omega)}{1-|f(\omega)|^2},$$
$$D_2 f(\omega) = \frac{(1-|\omega|^2)^2}{1-|f(\omega)|^2}\left[f''(\omega) - \frac{2\overline{\omega}f'(\omega)}{1-|\omega|^2} + \frac{2\overline{f(\omega)}f'(\omega)^2}{1-|f(\omega)|^2}\right].$$

With these notations, the Schwarz-Pick inequality for derivatives (1.2) can be restated as $|D_1 f(\omega)| \leq 1$, while (5.6) can be written as $|D_2 f(\omega)| \leq 2(1-|D_1 f(\omega)|^2)$.

We refer to [12, Proposition 3.4] for a different proof of (5.6).

5.2 Generalization to n points

5.2.1 The Nevanlinna-Pick problem

Let $\omega_1, \dots, \omega_n \in \mathbb{D}$ be n initial data, and let $\lambda_1, \dots, \lambda_n \in \mathbb{D}$ be n target data. The *Nevanlinna-Pick problem* is to determine wether there exists (or not) an holomorphic function $f : \mathbb{D} \to \mathbb{D}$ that interpolates the data, *i.e.* such that

$$f(\omega_i) = \lambda_i, \quad \text{for all } i \in [\![1, n]\!] \tag{5.8}$$

In [49], Pick proved the following theorem:

Theorem 5.2.1 (Pick)**.** *There exists an holomorphic function $f : \mathbb{D} \to \mathbb{D}$ satisfying the interpolation condition* (5.8) *if and only if the* Pick matrix

$$P_n := \left(\frac{1-\overline{\lambda_j}\lambda_i}{1-\overline{\omega_j}\omega_i}\right)_{1\leq i,j\leq n}$$

is positive semi-definite.

Moreover, in the regular case *where P_n is positive definite, the problem has infinitely many solutions, including Blaschke products of degree n. In the* singular case *where P_n has rank $m < n$, the problem has a unique solution, which is a Blaschke product of degree m. Problems which have no solution are said to be* void.

Independently, Nevanlinna ([42]) and Schur ([57]) developed recursive procedures which allow to construct all solution. The interested reader can refer to [2] for a more complete overview about this problem.

In the case $n = 2$, the positive semi-definiteness of the Pick matrix P_2 reduces to the condition:

$$\frac{(1-|\lambda_1|^2)(1-|\lambda_2|^2)}{(1-|\omega_1|^2)(1-|\omega_2|^2)} \geq \frac{|1-\overline{\lambda_2}\lambda_1|^2}{|1-\overline{\omega_2}\omega_1|^2},$$

which is equivalent to the Schwarz-Pick lemma, using the identity (1.3).

In the case $n = 3$, the Pick condition seems too complicated for practical uses. However, it is proved in [8, Theorem 7.1] that this condition is equivalent to the Beardon-Minda inequality:

$$\mathrm{d}\left(\frac{(\lambda_1,\lambda_2)}{(\omega_1,\omega_2)},\frac{(\lambda_3,\lambda_2)}{(\omega_3,\omega_2)}\right)\leq \mathrm{d}(\omega_1,\omega_3).$$

5.2.2 Higher order hyperbolic divided differences

In [7], Baribeau, Rivard, and Wegert generalized the work of Beardon and Minda. First of all, they iterated the notion of hyperbolic divided difference as follows:

Definition 5.2.2. Let $f\in\mathcal{H}(\mathbb{D},\mathbb{D})$ be an holomorphic function, and let $\omega_1,\dots,\omega_k\in\mathbb{D}$ be pairwise distinct points. We set:

$$\Delta^0 f(z)=f(z),\quad z\in\mathbb{D}$$
$$\Delta^j f(z;\omega_1,\dots,\omega_j)=\frac{(\Delta^{j-1}f(z;\omega_1,\dots,\omega_{j-1}),\Delta^{j-1}f(\omega_j;\omega_1,\dots,\omega_{j-1}))}{(z,\omega_j)},\quad j=1,\dots k$$

Then, they obtained the following generalization of Beardon and Minda's work:

Theorem 5.2.3 (Baribeau-Rivard-Wegert)**.** *Let $f\in\mathcal{H}(\mathbb{D},\mathbb{D})$ be an holomorphic function which is not a Blaschke product of degree $n\leq k$, and let $\omega_1,\dots,\omega_k\in\mathbb{D}$ be pairwise distinct points. Then, for all $u,v\in\mathbb{D}$, we have:*

$$\rho\left(\Delta^k f(u;\omega_1,\dots,\omega_k),\Delta^k f(v;\omega_1,\dots,\omega_k)\right)\leq\rho(u,v).$$

Moreover, if equality holds for a pair of distinct points u and v, then, f is a Blaschke product of degree (exactly) $k+1$, and equality holds for all $u,v\in\mathbb{D}$.

Moreover,the link with the Nevanlinna-Pick's problem is also established. In particular, as for Schwarz-Pick's lemma and Beardon-Minda's theorem, Baribeau-Rivard-Wegert's result is equivalent to Pick's criterion. To see this, we first introduce the following definition:

Definition 5.2.4. Let $\omega_1,\dots,\omega_n\in\mathbb{D}$ be n (pairwise distinct) initial data, and let $\lambda_1,\dots,\lambda_n\in\mathbb{D}$ be n target values. Let $\Delta_j^0:=\lambda_j$ for $j=1,\dots,n+1$. Assume that, for some k with $1\leq k\leq n$ and $j=k+1,\dots,n+1$, the divided differences Δ_j^{k-1} and Δ_k^{k-1} are given. Then, we set:

$$\Delta_j^k:=\begin{cases}\frac{\left[\Delta_j^{k-1},\Delta_k^{k-1}\right]}{\left[z_j,z_k\right]} & \text{if } \left|[\Delta_j^{k-1},\Delta_k^{k-1}]\right|\leq\left|[\omega_j,\omega_k]\right|,\\ +\infty & \text{otherwise}\end{cases}$$

This definition is compatible with the definition of $\Delta^k f$ for functions $f\in\mathcal{H}(\mathbb{D},\mathbb{D})$:

Lemma 5.2.5 (Baribeau-Rivard-Wegert)**.** *Let $\omega_1,\dots,\omega_n\in\mathbb{D}$ be pairwise distinct points, and let $f\in\mathcal{H}(\mathbb{D},\mathbb{D})$. If we set $\lambda_j:=f(\omega_j)$, for all $j\in[\![1,n]\!]$, then, we have:*

$$\Delta^k f(\omega_j;\omega_1,\dots,\omega_k)=\Delta_j^k,\quad 0\leq k\leq n-1,\ k+1\leq j\leq n.$$

The complete list of the hyperbolic divided differences can be conveniently arranged in a triangular table (see Table 5.1).

Then, we have the following useful criteria for the solvability of the Nevanlinna-Pick problem:

Points of $\mathbb{D}$		1	2	3	$\ldots$	$n-1$	n	$n+1$
ω_1	$\lambda_1 = \Delta_1^0$							
		Δ_2^1						
ω_2	$\lambda_2 = \Delta_2^0$		Δ_3^2					
		Δ_3^1		Δ_4^3				
					$\ddots$			
ω_3	$\lambda_3 = \Delta_3^0$		Δ_4^2					
		Δ_4^1		Δ_5^3		Δ_n^{n-1}		
ω_4	$\lambda_4 = \Delta_4^0$		Δ_5^2		$\vdots$		Δ_{n+1}^n	
		Δ_5^1		$\vdots$		Δ_{n+1}^{n-1}		
			$\vdots$					
ω_5	$\lambda_5 = \Delta_5^0$				$\iddots$			
$\vdots$	$\vdots$	$\vdots$		Δ_{n+1}^3				
ω_n	$\lambda_n = \Delta_n^0$		Δ_{n+1}^2					
		Δ_{n+1}^1						
ω_{n+1}	$\lambda_{n+1} = \Delta_{n+1}^0$							

Table 5.1: Table of hyperbolic divided differences

Theorem 5.2.6 (Baribeau-Rivard-Wegert)**.** *With the above notation, the Nevanlinna-Pick problem* (5.8) *is regular, singular or void if and only if*

$$\left|\Delta_n^{n-1}\right| < 1, \quad \left|\Delta_n^{n-1}\right| = 1 \quad \text{or} \quad \Delta_n^{n-1} = +\infty$$

respectively.

Remark 5.2.7. The proof of this result in [7, Theorem 4.1] furnishes an algorithm to construct all solutions, which coincides with the classical approaches of Nevanlinna and Schur. This algorithm can be seen as an hyperbolic version of the Newton scheme for polynomial interpolation.

5.2.3 The link with model matrices

We recall (see Section 2.3.6 and Section 4.4) that, if Θ_n is a finite Blaschke product with distinct zeros $\omega_1, \ldots, \omega_n \in \mathbb{D}$, the matrix representation M_{Θ_n} of the compressed shift S_{Θ_n} in the Takenaka-Malmquist-Walsh basis is given by:

$$\left[M_{\Theta_n}\right]_{i,j} = \begin{cases} \omega_j & \text{if } i=j \\ \prod_{k=i+1}^{j-1}(-\overline{\omega_k})\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_j|^2} & \text{if } i<j \\ 0 & \text{if } i>j \end{cases} \tag{5.9}$$

We have seen (see Section 1.2.3 and Section 5.1.1) that we can recover Schwarz-Pick's lemma and Beardon-Minda's inequality by applying the von Neumann inequality to the model matrices M_{Θ_2} and M_{Θ_3} respectively (which are themselves equivalent to the semi-positive definiteness of the Pick matrices P_2 and P_3).

For any $n \in \mathbb{N}^*$, M_{Θ_n} is a contraction. Thus, we can always apply von Neumann's inequality to it, an we get

$$\|f(M_{\Theta_n})\| \leq 1, \quad \text{for all holomorphic map} f : \mathbb{D} \to \mathbb{D} \tag{5.10}$$

In [10, Theorem 3.1], it is moreover proved that we have $\|f(M_{\Theta_n})\| = 1$ if and only if f is a finite Blaschke product of degree $\leq n-1$ (which coincides with the equality case in Beardon-Minda's inequality and Baribeau-Rivard-Wegert's generalization).

It is therefore natural to wonder wether we can recover Baribeau-Rivard-Wegert's result from (5.10). In order to obtain Schwarz-Pick's lemma and Beardon-Minda's inequality from (5.10) if the cases $n = 2$ and $n = 3$, we used criteria to determine whenever a 2×2 (respectively a 3×3) upper-triangular matrix is a contraction. Nevertheless, for 4×4 upper-triangular matrices, the criterion we obtained in Theorem 4.3.1 is quite complicated, and seems therefore to be of limited practical use. We can however establish a connection by a more theoretical (and a more roundabout) way, adapting the argument of Rovnyak in [53].

In order to do this, we consider the diagonalization of M_{Θ_n} (see Theorem 2.3.18): if we denote by $v_1,\dots,v_n$ the columns of the matrix P defined by

$$[P]_{i,j} = \begin{cases} 1 & \text{if } i=j \\ \frac{\sqrt{1-|\omega_i|^2}\sqrt{1-|\omega_j|^2}}{\omega_j-\omega_i}\prod_{k=i+1}^{j-1}\left(\frac{1-\overline{\omega_k}\omega_j}{\omega_j-\omega_k}\right) & \text{if } i<j \\ 0 & \text{if } i>j \end{cases}$$

then, $(v_1,\dots,v_n)$ is a basis of eigenvectors of M_{Θ_n}.

Moreover, for all $j\in[\![1,n]\!]$, we get:

$$\left\langle v_j, v_j\right\rangle = \frac{\prod_{l=1}^{j-1}|\omega_j-\omega_l|^2 + \sum_{i=1}^{j-1}(1-|\omega_i|^2)(1-|\omega_j|^2)\left(\prod_{l=1}^{i-1}|\omega_j-\omega_l|^2\right)\left(\prod_{k=i+1}^{j-1}|1-\overline{\omega_k}\omega_j|^2\right)}{\prod_{l=1}^{j-1}|\omega_j-\omega_l|^2}$$

Denoting by N_j the numerator of this fraction, we can write:

$$\begin{aligned} N_j &= \left(\prod_{l=1}^{j-2}|\omega_j-\omega_l|^2\right)\left(|\omega_j-\omega_{j-1}|^2 + (1-|\omega_{j-1}|^2)(1-|\omega_j|^2)\right) \\ &\quad + \sum_{i=1}^{j-2}(1-|\omega_i|^2)(1-|\omega_j|^2)\left(\prod_{l=1}^{i-1}|\omega_j-\omega_l|^2\right)\left(\prod_{k=i+1}^{j-1}|1-\overline{\omega_k}\omega_j|^2\right) \\ &= |1-\overline{\omega_{j-1}}\omega_j|^2\left(\prod_{l=1}^{j-2}|\omega_j-\omega_l|^2\right) + \sum_{i=1}^{j-2}(1-|\omega_i|^2)(1-|\omega_j|^2)\left(\prod_{l=1}^{i-1}|\omega_j-\omega_l|^2\right)\left(\prod_{k=i+1}^{j-1}|1-\overline{\omega_k}\omega_j|^2\right) \end{aligned}$$

using the key identity

$$|1-\overline{u}v|^2 = (1-|u|^2)(1-|v|^2) + |u-v|^2, \quad \text{for all } u,v\in\mathbb{C}.$$

Iterating this process, we obtain:

$$
\begin{aligned}
N_j &= |1-\overline{\omega_{j-1}}\omega_j|^2\left(\prod_{l=1}^{j-3}|\omega_j-\omega_l|^2\right)\left(|\omega_j-\omega_{j-2}|^2+(1-|\omega_{j-2}|^2)(1-|\omega_j|^2)\right)\\
&\qquad+\sum_{i=1}^{j-3}(1-|\omega_i|^2)(1-|\omega_j|^2)\left(\prod_{l=1}^{i-1}|\omega_j-\omega_l|^2\right)\left(\prod_{k=i+1}^{j-1}|1-\overline{\omega_k}\omega_j|^2\right)\\
&= |1-\overline{\omega_{j-1}}\omega_j|^2\cdot|1-\overline{\omega_{j-2}}\omega_j|^2\left(\prod_{l=1}^{j-3}|\omega_j-\omega_l|^2\right)|1-\overline{\omega_{j-1}}\omega_j|^2\\
&\qquad+\sum_{i=1}^{j-3}(1-|\omega_i|^2)(1-|\omega_j|^2)\left(\prod_{l=1}^{i-1}|\omega_j-\omega_l|^2\right)\left(\prod_{k=i+1}^{j-1}|1-\overline{\omega_k}\omega_j|^2\right)\\
&= \cdots\\
&= \prod_{k=1}^{j-1}|1-\overline{\omega_k}\omega_j|^2.
\end{aligned}
$$

Finally, we get

$$
\left\langle v_j, v_j\right\rangle = \prod_{l=1}^{j-1}\frac{|1-\overline{\omega_l}\omega_j|^2}{|\omega_j-\omega_l|^2}.
$$

Similarly, for $1 \le j < k \le n$, we have:

$$
\begin{aligned}
\left\langle v_j, v_k\right\rangle &= \sum_{i=1}^{j-1}(1-|\omega_i|^2)\left(\frac{\sqrt{1-|\omega_j|^2}}{\omega_j-\omega_i}\prod_{l=i+1}^{j-1}\frac{1-\overline{\omega_l}\omega_j}{\omega_j-\omega_l}\right)\cdot\left(\frac{\sqrt{1-|\omega_k|^2}}{\overline{\omega_k}-\overline{\omega_i}}\prod_{p=i+1}^{k-1}\frac{1-\omega_p\overline{\omega_k}}{\overline{\omega_k}-\overline{\omega_p}}\right)\\
&\qquad+\frac{\sqrt{1-|\omega_j|^2}\sqrt{1-|\omega_k|^2}}{\overline{\omega_k}-\overline{\omega_j}}\prod_{l=j+1}^{k-1}\frac{1-\omega_l\overline{\omega_k}}{\overline{\omega_k}-\overline{\omega_l}}\\
&= \frac{\sqrt{1-|\omega_j|^2}\sqrt{1-|\omega_k|^2}\cdot N_{j,k}}{\left(\prod_{l=1}^{j-1}\omega_j-\omega_l\right)\cdot\left(\prod_{p=1}^{k-1}\overline{\omega_k}-\overline{\omega_p}\right)\cdot\left(1-\overline{\omega_k}\omega_j\right)},
\end{aligned}
$$

where

$$
\begin{aligned}
N_{j,k} := (1-\overline{\omega_k}\omega_j)\sum_{i=1}^{j-1}(1-|\omega_i|^2)&\left(\prod_{l=i+1}^{j-1}1-\overline{\omega_l}\omega_j\right)\left(\prod_{p=i+1}^{k-1}1-\overline{\omega_k}\omega_p\right)\left(\prod_{q=1}^{i-1}\omega_j-\omega_q\right)\left(\prod_{r=1}^{i-1}\overline{\omega_k}-\overline{\omega_r}\right)\\
&+\left(\prod_{p=j}^{k-1}1-\overline{\omega_k}\omega_p\right)\left(\prod_{q=1}^{j-1}\omega_j-\omega_q\right)\left(\prod_{r=1}^{j-1}\overline{\omega_k}-\overline{\omega_l}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= (1-\overline{\omega_k}\omega_j)\sum_{i=1}^{j-2}(1-|\omega_i|^2)\left(\prod_{l=i+1}^{j-1} 1-\overline{\omega_l}\omega_j\right)\left(\prod_{p=i+1}^{k-1} 1-\overline{\omega_k}\omega_p\right)\left(\prod_{q=1}^{i-1}\omega_j-\omega_q\right)\left(\prod_{r=1}^{i-1}\overline{\omega_k}-\overline{\omega_r}\right)\\
&\quad+\left(\prod_{p=j}^{k-1} 1-\overline{\omega_k}\omega_p\right)\left(\prod_{q=1}^{j-2}\omega_j-\omega_q\right)\left(\prod_{r=1}^{j-2}\overline{\omega_k}-\overline{\omega_l}\right)\left[(1-|\omega_j|^2)(1-\overline{\omega_k}\omega_j)+(\omega_j-\omega_{j-1})(\overline{\omega_k}-\overline{\omega_{j-1}})\right]\\
&= (1-\overline{\omega_k}\omega_j)\sum_{i=1}^{j-2}(1-|\omega_i|^2)\left(\prod_{l=i+1}^{j-1} 1-\overline{\omega_l}\omega_j\right)\left(\prod_{p=i+1}^{k-1} 1-\overline{\omega_k}\omega_p\right)\left(\prod_{q=1}^{i-1}\omega_j-\omega_q\right)\left(\prod_{r=1}^{i-1}\overline{\omega_k}-\overline{\omega_r}\right) \quad ,\\
&\qquad\qquad+(1-\overline{\omega_{j-1}}\omega_j)\left(\prod_{p=j-1}^{k-1} 1-\overline{\omega_k}\omega_p\right)\left(\prod_{q=1}^{j-2}\omega_j-\omega_q\right)\left(\prod_{r=1}^{j-2}\overline{\omega_k}-\overline{\omega_l}\right)
\end{aligned}
$$

using the identity

$$(1-|u|^2)(1-\overline{w}v)+(v-u)(\overline{w}-\overline{u}) = (1-\overline{w}u)(1-\overline{u}v), \quad \text{for all } u,v,w\in\mathbb{C} \tag{5.11}$$

Iterating this process, we obtain:

$$N_{j,k} = \left(\prod_{p=1}^{k-1} 1-\overline{\omega_k}\omega_p\right)\left(\prod_{l=1}^{j-1} 1-\overline{\omega_l}\omega_j\right),$$

and, finally,

$$\left\langle v_j, v_k\right\rangle = \frac{\sqrt{1-|\omega_j|^2}\sqrt{1-|\omega_k|^2}}{1-\overline{\omega_k}\omega_j}\left(\prod_{l=1}^{j-1}\frac{1-\overline{\omega_l}\omega_j}{\omega_j-\omega_l}\right)\left(\prod_{p=1}^{k-1}\frac{1-\overline{\omega_k}\omega_p}{\overline{\omega_k}-\overline{\omega_p}}\right).$$

Now, for $j\in[\![1,n]\!]$, let

$$\widetilde{v_j} = \frac{1}{\sqrt{1-|\omega_j|^2}}\left(\prod_{j=1}^{j-1}\frac{\omega_j-\omega_l}{1-\overline{\omega_l}\omega_j}\right)v_j.$$

By construction, $(\widetilde{v_1},\ldots,\widetilde{v_n})$ is a basis of $\mathcal{K}_{\Theta_n}$ satisfying

$$S_{\Theta_n}\left(\widetilde{v_j}\right) = \omega_j\widetilde{v_j}, \quad \text{for all } j\in[\![1,n]\!] \tag{5.12}$$

and

$$\left\langle\widetilde{v_j},\widetilde{v_k}\right\rangle = \frac{1}{1-\overline{\omega_k}\omega_j}, \quad \text{for all } j,k\in[\![1,n]\!] \tag{5.13}$$

Now, if $f\in\mathcal{H}(\mathbb{D},\mathbb{D})$ is an holomorphic map mapping $\mathbb{D}$ to $\mathbb{D}$, we have:

$$f\left(S_{\Theta_n}\right)\left\{\sum_{j=1}^{n} c_j\widetilde{v_j}\right\} = \sum_{j=1}^{n} c_j f(\omega_j)\widetilde{v_j}, \quad \text{for any scalars } c_1,\ldots,c_n\in\mathbb{C}.$$

As S_{Θ_n} is a contraction, by von Neumann's inequality, we have:

$$\left\| \sum_{j=1}^{n} c_j f(\omega_j) \widetilde{v_j} \right\|^2 \leq \left\| \sum_{j=1}^{n} c_j \widetilde{v_j} \right\|^2 .$$

By (5.13), this is equivalent to:

$$\sum_{k=1}^{n} \sum_{j=1}^{n} \frac{1 - \overline{f(\omega_k)} f(\omega_j)}{1 - \overline{\omega_k} \omega_j} c_j \overline{c_k} \geq 0,$$

which is exactly the semi-positive definiteness of the Pick matrix (which is itself equivalent to Baribeau-Rivard-Wegert's result).

5.3 Higher order Schwarz-Pick inequalities

Let $f \in \mathcal{H}(\mathbb{D},\mathbb{D})$ be an analytic function of $\mathbb{D}$ into itself with $f(z) = \sum_{n=0}^{\infty} a_n z^n$. It has been proved by F.W. Wiener that for each $k \geq 1$ we have

$$|a_k| \leq 1 - |a_0|^2. \tag{5.14}$$

We refer for instance to [47] for an operator theoretical proof of this inequality and for applications to Bohr's phenomenon. For $k = 1$, the inequality (5.14) gives $|f'(0)| \leq 1 - |f(0)|^2$. Applying this inequality to $F(z) = f((\omega+z)/1+\overline{\omega}z)$, for a fixed $\omega \in \mathbb{D}$, we obtain the Schwarz-Pick inequality (1.2). For an arbitrary k, a similar reasoning has been used by Ruscheweyh [56] to obtain the following sharp higher-order inequality for an analytic function $f \in \mathcal{H}(\mathbb{D},\mathbb{D})$, $z \in \mathbb{D}$ and $k \geq 1$:

$$|f^{(k)}(z)| \leq \frac{k!(1 - |f(z)|^2)}{(1 - |z|)^k (1 + |z|)}. \tag{5.15}$$

We prove in this section some results related to (5.14).

Theorem 5.3.1. *Let f be an analytic function of $\mathbb{D}$ into $\overline{\mathbb{D}}$ with $f(z) = \sum_{n=0}^{\infty} a_n z^n$ for $z \in \mathbb{D}$. Then, for each $n \geq 1$ and each $k \geq 1$ we have*

$$\left| a_{n+k}(1 - |a_0^2|) + a_n a_k \overline{a}_0 \right|^2 \leq \left[(1 - |a_0|^2)^2 - |a_n|^2 \right] \cdot \left[(1 - |a_0|^2)^2 - |a_k|^2 \right]. \tag{5.16}$$

Proof. As $\|f\|_\infty \leq 1$, the multiplication operator M_f given by $M_f(g) = fg$ acts contractively on the Hardy space $H^2(\mathbb{D})$. Recall that $\{z^n : n \geq 0\}$ is an orthonormal basis of $H^2(\mathbb{D})$. The compression $T = P_K M_{f|K}$ of M_f to the 3-dimensional Euclidean space $K = \text{span}(1, z^n, z^{n+k})$ is also a contraction. The matrix of T is given by

$$T = \begin{pmatrix} a_0 & a_n & a_{n+k} \\ 0 & a_0 & a_k \\ 0 & 0 & a_0 \end{pmatrix}.$$

Then (5.16) is a consequence of Theorem 4.2.2. $\square$

When $a_0 = 0$ we obtain the following consequence.

Corollary 5.3.2. *Let f be a analytic function of $\mathbb{D}$ into $\overline{\mathbb{D}}$ with $f(0) = 0$ and $f(z) = \sum_{n=1}^{\infty} a_n z^n$ for $z \in \mathbb{D}$. Then*

$$|a_{n+k}| \le \sqrt{1-|a_n|^2} \cdot \sqrt{1-|a_k|^2}. \tag{5.17}$$

For $n = k = 1$, and $a_0 = 0$, we obtain the inequality $|a_2| \le 1 - |a_1|^2$. Applying this inequality to (5.7) we obtain Yamashita's inequality $|D_2 f(\omega)| \le 2(1 - |D_1 f(\omega)|^2)$.

The following consequence is an improvement of Wiener's inequality (5.14).

Corollary 5.3.3. *Let f be a analytic function of $\mathbb{D}$ into $\overline{\mathbb{D}}$ with $f(z) = \sum_{n=0}^{\infty} a_n z^n$ for $z \in \mathbb{D}$. Then*

$$1 - |a_0|^2 - |a_n| \ge \frac{\left|a_{2n}(1 - |a_0^2|) + a_n^2 \overline{a}_0\right|}{2(1 - |a_0|^2)}.$$

Proof. Applying (5.16) for $k = n$ we obtain

$$\left|a_{2n}(1 - |a_0^2|) + a_n^2 \overline{a}_0\right| \le \left[(1 - |a_0|^2)^2 - |a_n|^2\right].$$

Therefore

$$1 - |a_0|^2 - |a_n| \ge \frac{\left|a_{2n}(1 - |a_0^2|) + a_n^2 \overline{a}_0\right|}{1 - |a_0|^2 + |a_n|} \ge \frac{\left|a_{2n}(1 - |a_0^2|) + a_n^2 \overline{a}_0\right|}{2(1 - |a_0|^2)}.$$

The proof is complete. □

5.4 Schwarz-Pick inequalities for the polydisc

5.4.1 Using von Neumann's inequality for tuples of 2×2 matrices

Using the von Neumann inequality for tuples of 2×2 matrices (see Theorem 1.2.11) gives an alternative proof of the following known ([54, Lemma 7.5.6]) Schwarz-Pick inequality for the polydisc:

Theorem 5.4.1.

(i) Let $f \in \mathcal{H}(\mathbb{D}^n, \mathbb{D})$ and let $a = (a_1, \dots, a_n)$, $b = (b_1, \dots, b_n) \in \mathbb{D}^n$. Then:

$$\left|\frac{f(a_1, \cdots, a_n) - f(b_1, \cdots, b_n)}{1 - \overline{f(a_1, \cdots, a_n)} f(b_1, \cdots, b_n)}\right| \le \max_{1 \le i \le n} \left|\frac{a_i - b_i}{1 - \overline{a_i} b_i}\right|. \tag{5.18}$$

(ii) Let $f \in \mathcal{H}(\mathbb{D}^n, \mathbb{D})$ and let $a = (a_1, \dots, a_n) \in \mathbb{D}^n$. Then:

$$\sum_{i=1}^{n} (1 - |a_i|^2) \left|\frac{\partial f(a)}{\partial z_i}\right| \le 1 - |f(a)|^2. \tag{5.19}$$

Proof.

(i) We first observe that the result is obvious whenever $a = b$ or $f(a) = f(b)$. Therefore, in the following, we assume $a \ne b$ and $f(a) \ne f(b)$.

For $1 \le i \le n$, let

$$T_i = \begin{pmatrix} a_i & d(a_i - b_i) \\ 0 & b_i \end{pmatrix},$$

with

$$d = \min_{1 \le i \le n} \sqrt{\frac{(1-|a_i|^2)(1-|b_i|^2)}{|a_i - b_i|^2}}.$$

Here, whenever $a_i = b_i$, we make the convention that $\sqrt{\frac{(1-|a_i|^2)(1-|b_i|^2)}{|a_i-b_i|^2}} = +\infty$. As we assume that $a \neq b$, this cannot happen for all the indices i.

It can be easily verified that the matrices T_i are mutually commuting and that $\|T_i\| \le 1$. By induction it can be shown that for all $i \in [\![1,n]\!]$, for all $k_i \in \mathbb{N}$,

$$T_i^{k_i} = \begin{pmatrix} a_i^{k_i} & d\left(a_i^{k_i} - b_i^{k_i}\right) \\ 0 & b_i^{k_i} \end{pmatrix}.$$

Let $p \in \mathbb{C}[X_1, \dots, X_n]$ be a polynomial with $\|p\|_\infty < 1$. We have

$$p(T_1, \cdots, T_n) = \begin{pmatrix} p(a) & d(p(a) - p(b)) \\ 0 & p(b) \end{pmatrix}.$$

Drury's result (Theorem 1.2.11) implies that $\|p(T_1, \dots, T_n)\| \le 1$. As in the one variable case, a computation gives the Schwarz-Pick inequality (5.18) for p. By an approximation argument, (5.18) holds also for functions in the polydisc algebra. Now, if $f \in \mathcal{H}(\mathbb{D}^n, \mathbb{D})$, consider the family of functions $(f_r)_{0<r<1}$ defined by $f_r(z_1, \cdots, z_n) = f(rz_1, \dots, rz_n)$. For all $r \in]0,1[$, f_r is in the polydisc algebra and, thus, f_r satisfies (5.18). Then, let $r \to 1^-$ to conclude the proof.

(ii) The proof follows the same method as that of Theorem 5.4.1. Let $a = (a_2, \dots, a_n) \in \mathbb{D}^n$ and let $p \in \mathbb{C}[X_1, \dots, X_n]$ be a polynomial with $\|p\|_\infty < 1$. For $1 \le k \le n$, let $T_k = \begin{pmatrix} a_k & \gamma_k \\ 0 & a_k \end{pmatrix}$, where $\gamma_k = e^{i\theta_k}(1 - |a_k|^2)$, for some $\theta_k \in [0, 2\pi[$ to be chosen later on. For all $k \in [\![1,n]\!]$, $\|T_k\| \le 1$, and, for all $k, l \in [\![1,n]\!]$, $T_k T_l = T_l T_k$. We have:

$$p(T_1, \cdots, T_n) = \begin{pmatrix} p(a) & \sum_{k=1}^n \gamma_k \frac{\partial p(a)}{\partial z_k} \\ 0 & p(a) \end{pmatrix}.$$

Again, by Theorem 1.2.11, we get $\|p(T_1, \dots, T_n)\| \le 1$. Therefore:

$$\left| \sum_{k=1}^n \gamma_k \frac{\partial p(a)}{\partial z_k} \right| \le 1 - |p(a)|^2. \tag{5.20}$$

Now, let $t_k = \frac{\partial p(a)}{\partial z_k}$, $0 \le k \le n$. We write $t_k = |t_k| e^{i\operatorname{Arg}(t_k)}$ and we set $\theta_k = -\operatorname{Arg}(t_k)$. With this choice we obtain $\gamma_k t_k = (1 - |a_k|^2)|t_k|$. Replacing in (5.20) we get

$$\sum_{i=1}^n (1 - |a_i|^2) \left| \frac{\partial p(a)}{\partial z_i} \right| \le 1 - |p(a)|^2.$$

We conclude by using an approximation argument.

□

Remark 5.4.2. The study of the case of equality in the Schwarz-Pick inequalities for the polydisc is an interesting problem. Knese [38] studied the equality case in (5.19) using operator-theoretical methods (transfer functions) and described which functions play the role of automorphisms of the disk in this context: they turn out to be rational inner functions in the Schur-Agler class of the polydisc with an *added* symmetry constraint.

5.4.2 Peschl's invariant derivatives in several variables

The inequalities from Section 5.1.2 can be extended to analytic functions of several variables.

Let $n \in \mathbb{N}^*$, let $f \in \mathcal{H}(\mathbb{D}^n, \mathbb{D})$, and fix a vector $\omega = (\omega_1, \cdots, \omega_n)$ in $\mathbb{D}^n$. Similarly as in the one variable case, we define

$$g : z = (z_1, \cdots, z_n) \in \mathbb{D}^n \mapsto \frac{f\left(\frac{z_1+\omega_1}{1+\overline{\omega_1}z_1}, \cdots, \frac{z_n+\omega_n}{1+\overline{\omega_n}z_n}\right) - f(\omega_1, \cdots, \omega_n)}{1 - \overline{f(\omega_1, \cdots, \omega_n)} f\left(\frac{z_1+\omega_1}{1+\overline{\omega_1}z_1}, \cdots, \frac{z_n+\omega_n}{1+\overline{\omega_n}z_n}\right)} \in \mathbb{C}$$

and, then, write

$$g(z_1, \cdots, z_n) = \sum_{j_1, \cdots, j_n = 0}^{\infty} \frac{\partial^{j_1+\cdots+j_n} g(0, \cdots, 0)}{\partial^{j_1} z_1 \cdots \partial^{j_n} z_n} z_1^{j_1} \cdots z_n^{j_n} = \sum_{j_1, \cdots, j_n = 0}^{\infty} a_{j_1, \cdots, j_n} z_1^{j_1} \cdots z_n^{j_n}.$$

For $k \in [\![1, n]\!]$, let $D_k f(w) = \partial^k g(0, \ldots, 0) = \sum_{j_1+\cdots+j_n=k} a_{j_1, \cdots, j_n}$. A straightforward computation gives:

$$\begin{aligned}
D_1 f(\omega) &= \sum_{j=1}^{n} \frac{1-|\omega_j|^2}{1-|f(\omega)|^2} \cdot \frac{\partial f}{\partial z_j}(\omega), \\
D_2 f(\omega) &= \sum_{j=1}^{n} \frac{\partial^2 g(0, \cdots, 0)}{\partial^2 z_j} + 2 \sum_{1 \le j < k \le n} \frac{\partial^2 f(0, \cdots, 0)}{\partial z_j \partial z_k} \\
&= \sum_{j=1}^{n} \frac{(1-|\omega_j|^2)^2}{1-|f(\omega)|^2} \left(\frac{\partial^2 f(w)}{\partial^2 z_j} + \frac{2\overline{f(w)}}{1-|f(w)|^2} - \frac{2\overline{\omega_j}}{1-|\omega_j|^2} \cdot \frac{\partial f(\omega)}{\partial z_j} \right) \\
&\quad + 2 \sum_{1 \le j < k \le n} \frac{(1-|z_j|^2)(1-|z_k|^2)}{1-|f(\omega)|^2} \left(\frac{\partial f(\omega)}{\partial z_j \partial z_k} + \frac{2\overline{f(\omega)}}{1-|f(\omega)|^2} \cdot \frac{\partial f(\omega)}{\partial z_j} \cdot \frac{\partial f(\omega)}{\partial z_k} \right).
\end{aligned}$$

With the same method of proof as before, we can obtain the following result:

Theorem 5.4.3. *For $n \in \mathbb{N}^*$ let $w = (\omega_1, \ldots, \omega_n) \in \mathbb{D}^n$ and consider $f \in \mathcal{H}(\mathbb{D}^n, \mathbb{D})$. Then, we have:*

$$|D_2 f(\omega)| \le 2(1 - |D_1 f(\omega)|^2). \tag{5.21}$$

Proof. For $1 \leq k \leq n$, let

$$T_k = \begin{pmatrix} \omega_k & \alpha_k & \beta_k \\ 0 & \omega_k & \alpha_k \\ 0 & 0 & \omega_k \end{pmatrix} \in \mathcal{M}_3(\mathbb{C}),$$

with $\alpha_k = 1 - |\omega_k|^2$ and $\beta_k = -\overline{\omega_k}(1 - |\omega_k|^2)$. By Theorem 4.2.2, T_k is a contraction, for all $k \in [\![1, n]\!]$. Moreover, for all $1 \leq k, j \leq n$, $T_j T_k = T_k T_j$. Therefore, by Knese's result (Theorem 1.2.12), $p(T_1, \ldots, T_n)$ is a contraction, for every $p \in \mathbb{C}[X_1, \ldots, X_n]$ with $\|p\|_\infty < 1$. Moreover, it is easy to check that

$$p(T_1, \ldots, T_n) = \begin{pmatrix} p(\omega) & \gamma_1 & \gamma_2 \\ 0 & p(\omega) & \gamma_1 \\ 0 & 0 & p(\omega) \end{pmatrix},$$

with

$$\gamma_1 = \sum_{j=1}^{n} \alpha_j \frac{\partial p(\omega)}{\partial z_j},$$

$$\gamma_2 = \frac{1}{2} \sum_{j=1}^{n} \alpha_j^2 \frac{\partial^2 p(w)}{\partial^2 z_j} + \sum_{1 \leq j < k \leq n} \alpha_j \alpha_k \frac{\partial^2 p(w)}{\partial z_j \partial z_k} + \sum_{j=1}^{n} \beta_j \frac{\partial p(w)}{\partial z_j}.$$

By Theorem 4.2.2, we obtain :

$$\left| \gamma_2 \left(1 - |p(\omega)|^2\right) + \gamma_1^2 \overline{p(\omega)} \right| \leq (1 - |p(\omega)|^2)^2 - |\gamma_1|^2, \tag{5.22}$$

which is equivalent to (5.21) for polynomials. The inequality extends to all functions $f \in \mathcal{H}(\mathbb{D}^n, \mathbb{D})$. □

5.4.3 Distinguished varieties and Schwarz-Pick inequalities

In the bidisc case, the refined version of Ando's inequality by Agler and McCarthy (see [1] or [3, Section 8.6]) results in corresponding enhancements of Schwarz-Pick type inequalities.

We start by recalling the notion of distinguished variety introduced in [1]:

Definition 5.4.4. A *distinguished variety* is a set of the form $V \cap \overline{\mathbb{D}}^2$, where V is an algebraic set in $\mathbb{C}^2$ (so there is a polynomial $q \in \mathbb{C}[X_1, X_2]$ such that $V = \{(z, w) \in \mathbb{D}^2 : q(z, w) = 0\}$) with the property that

$$\overline{V} \cap \partial(\mathbb{D}^2) = \overline{V} \cap \mathbb{T}^2.$$

Therefore a distinguished variety is the trace on $\mathbb{D}^2$ of a one-dimensional complex algebraic variety V in $\mathbb{C}^2$ such that V intersects $\mathbb{D}^2$ and exits the bidisc through its distinguished boundary, $\mathbb{T}^2$, without intersecting any other part of its topological boundary.

We have the following representation for distinguished varieties:

Theorem 5.4.5 (Agler and McCarthy, [1])**.** *A distinguished variety has the following determinantal representation*

$$V \cap \mathbb{D}^2 = \left\{(z, w) \in \mathbb{D}^2 : \ det\ (\Psi(z) - w\mathrm{Id}) = 0\right\} \tag{5.23}$$

for some matrix-valued rational function Ψ on the unit disc that is unitary on the unit circle.

Agler and McCarthy also proved the following sharpening of Ando's inequality:

Theorem 5.4.6 (Agler and McCarthy, [1]). *For any pair of commuting contractive matrices (T_1, T_2) without unimodular eigenvalues, there is a distinguished variety $V \cap \mathbb{D}^2$ such that the von-Neumann inequality holds on $V \cap \mathbb{D}^2$ for any polynomial p in $\mathbb{C}[X_1, X_2]$, i.e.*

$$\|p(T_1, T_2)\| \leq \sup\{|p(z_1, z_2)| : (z_1, z_2) \in V \cap \mathbb{D}^2\}. \tag{5.24}$$

Theorem 5.4.7.

(i) Let (a_1, a_2) and (b_1, b_2) be two points in the bidisc $\mathbb{D}^2$. Then, there is a distinguished variety $V \cap \mathbb{D}^2$ such that the Schwarz-Pick inequality

$$\left|\frac{f(a_1, a_2) - f(b_1, b_2)}{1 - \overline{f(a_1, a_2)} f(b_1, b_2)}\right| \leq \max\left\{\left|\frac{a_1 - b_1}{1 - \overline{a_1} b_1}\right|, \left|\frac{a_2 - b_2}{1 - \overline{a_2} b_2}\right|\right\} \tag{5.25}$$

holds for any function f which is holomorphic on the bidisc $\mathbb{D}^2$ and continuous on $\overline{\mathbb{D}}^2$ with

$$\sup\{|f(z_1, z_2)| : (z_1, z_2) \in V \cap \mathbb{D}^2\} \leq 1$$

(ii) Let (a_1, a_2) and (b_1, b_2) be two points in the bidisc $\mathbb{D}^2$. Then, there is a distinguished variety $V \cap \mathbb{D}^2$ such that the Schwarz-Pick inequality (5.25) holds for any function f which is holomorphic in the bidisc $\mathbb{D}^2$ and for which there is a sequence of positive real number (r_n) convergent to 1 with $r_n < 1$ such that

$$\sup\{|f(r_n z_1, r_n z_2)| : n \geq 1, (z_1, z_2) \in V \cap \mathbb{D}^2\} \leq 1$$

Proof. Consider the matrices

$$T_1 = \begin{pmatrix} a_1 & d(a_1 - b_1) \\ 0 & b_1 \end{pmatrix}, \quad T_2 = \begin{pmatrix} a_2 & d(a_2 - b_2) \\ 0 & b_2 \end{pmatrix},$$

with

$$d = \min\left\{\sqrt{\frac{(1 - |a_1|^2)(1 - |b_1|^2)}{|a_1 - b_1|^2}}, \sqrt{\frac{(1 - |a_2|^2)(1 - |b_2|^2)}{|a_2 - b_2|^2}}\right\},$$

(using the same conventions as in the proof of Theorem 5.4.1–(i)). Following [1], we can also assume that T_1 and T_2 are jointly diagonalizable (this is the first case in the proof of [1, Theorem 3.1]). It follows from the result proved in [1] (see also [3, p.211] for details and unexplained terminology) that there is a distinguished variety V such that $T = (T_1, T_2)$ can be extended to a pair of commuting unitaries $U = (U_1, U_2)$ with spectrum $\sigma(U) = \overline{V} \cap \partial(\mathbb{D}^2) = \overline{V} \cap \mathbb{T}^2$. As f is in the bidisc algebra, $f(T)$ and $f(U)$ are well-defined and $f(T)$ is a restriction of $f(U)$ to $\mathbb{C}^2 \times \mathbb{C}^2$. We obtain, as in [1], that

$$\|f(T_1, T_2)\| \leq \sup\{|f(z_1, f_2)| : (z_1, z_2) \in V \cap \mathbb{D}^2\}. \tag{5.26}$$

Therefore $f(T_1, T_2)$ is a contraction and the proof of Theorem 5.4.1–(i), implies that inequality (5.25) holds true. The second part follows from the first one applied to the functions $f(r_n z_1, r_n z_2)$ and then making $n \to \infty$. □

The following result follows in a similar manner from the Agler and McCarthy result and the proof of Theorem 5.4.3.

Theorem 5.4.8. *Let $w = (\omega_1, \omega_2) \in \mathbb{D}^2$. Then there exists a distinguished variety $V \cap \mathbb{D}^2$ such that*

$$|D_2 f(\omega)| \leq 2(1 - |D_1 f(\omega)|^2) \tag{5.27}$$

for every $f \in \mathcal{A}(\overline{\mathbb{D}}^2)$ with

$$\sup\{|f(z_1, z_2)| \leq 1 : (z_1, z_2) \in V \cap \mathbb{D}^2\}.$$

Remark 5.4.9. Some Nevanlinna–Pick interpolation problems on distinguished varieties in the bidisc have been studied in [34].

5.5 Operator versions of Schwarz-Pick type inequalities

We move now to operator versions of the Schwarz-Pick and Beardon-Minda inequalities. The first operator generalizations for the Schwarz lemma and the Schwarz-Pick inequality have been proved by Fan in [22]:

Theorem 5.5.1 (Fan). *Let H be an Hilbert space and let $A \in \mathcal{B}(H)$ be a proper contraction, i.e. $\|A\| < 1$. Let $f, g, h \in \mathbb{D}$ be such that $f = gh$ and $|h(z)| \leq 1$ for all $z \in \mathbb{D}$.*

Then, we have:

(i) $g(A)g(A)^* \geq f(A)^* f(A)$;

(ii) $\|g(A)\| \geq \|f(A)\|$.

Strict inequality holds in (i) if and only if $g(A)^* g(A) > 0$ and h is not a constant function of absolute value 1. Equality in (ii) holds if and only if either $g(A) = 0$ or h is a constant function of absolute value 1.

Corollary 5.5.2. *Let A be a proper contraction on a Hilbert space H and let $f \in \mathcal{H}(\mathbb{D})$ be such that $|f(z)| < 1$, for all $z \in \mathbb{D}$ and $f(0) = f'(0) = \cdots = f^{(n-1)}(0) = 0$ for some $n \in \mathbb{N}^*$.*

Then, we have:

(i) $A^{n*}A^n \geq f(A)^* f(A)$;

(ii) $\|A^n\| \geq \|f(A)\|$.

Strict inequality holds in (i) if and only if $A^{n*}A^n > 0$ and if f is not of the form $f(z) = \eta z^n$ for some constant η with $|\eta| = 1$. Equality occurs in (ii) if and only if either $A^n = 0$ or $f(z) = \eta z^n$ with $|\eta| = 1$.

Theorem 5.5.3 (Fan). *Let A be a proper contraction on a Hilbert space, let $f \in \mathcal{H}(\mathbb{D})$ such that $|f(z)| < 1$, for all $z \in \mathbb{D}$, and let $z_0 \in \mathbb{D}$.*

Then, we have:

$$\begin{aligned}
&(\mathrm{Id} - z_0 A^*)^{-1}(A^* - \overline{z_0}\mathrm{Id})(A - z_0\mathrm{Id})(\mathrm{Id} - \overline{z_0}A)^{-1} \\
&\quad \geq (\mathrm{Id} - f(z_0) f(A)^*)^{-1}\left(f(A)^* - \overline{f(z_0)}\mathrm{Id}\right)(f(A) - f(z_0)\mathrm{Id})\left(\mathrm{Id} - \overline{f(z_0)} f(A)\right)^{-1},
\end{aligned} \tag{5.28}$$

which implies that:

$$\left\|(A-z_0\mathrm{Id})(\mathrm{Id}-\overline{z_0}A)^{-1}\right\| \geq \left\|(f(A)-f(z_0)\mathrm{Id})\left(\mathrm{Id}-\overline{f(z_0)}f(A)\right)^{-1}\right\|. \tag{5.29}$$

Strict inequality occurs in (5.28) if and only if $(A^*-\overline{z_0}\mathrm{Id})(A-z_0\mathrm{Id})>0$ *and* f *is not of the form* $f(z)=\varepsilon\frac{z-z_1}{1-\overline{z_1}z}$, *with* $|\varepsilon|=1$ *and* $|z_1|<1$. *Equality occurs in (5.5.3) if and only if either* $A=z_0\mathrm{Id}$ *or* f *is of the form* $f(z)=\varepsilon\frac{z-z_1}{1-\overline{z_1}z}$, *with* $|\varepsilon|=1$ *and* $|z_1|<1$.

Variations of the last theorem have also been discussed in [4].

More recently, Jocić proved in [33, Theorem 3.5] the following operator version of the Schwarz-Pick lemma:

Theorem 5.5.4 (Jocić)**.** *Let* $f\in\mathcal{H}(\mathbb{D},\mathbb{D})$ *be an holomorphic map on the open unit disk and let* $A,B\in\mathcal{B}(H)$ *be two proper contractions on a Hilbert space* H.

Then, we have:

$$\left\|D_{f(A)^*}^{-1}[f(A),f(B)]_{(X)}D_{f(B)}^{-1}\right\| \leq \left\|D_{A^*}^{-1}XD_B^{-1}\right\|.$$

5.5.1 An operator version of the Schwarz-Pick inequality

The following result is a counterpart of [33, Theorem 3.5]. When specialized to scalars, it reduces to the Schwarz-Pick inequality for two distinct points.

Theorem 5.5.5. *Let* H_1,H_2 *be two Hilbert spaces. Consider three contractions* $W_1\in\mathcal{B}(H_1)$, $W_2\in\mathcal{B}(H_2)$ *and* $V\in\mathcal{B}(H_2,H_1)$. *Assume that* $\sigma(W_1)\cap\sigma(W_2)=\emptyset$, *and that* $f\in\mathcal{H}(\mathbb{D},\mathbb{D})$ *is holomorphic on an open neighborhood of* $\sigma(W_1)\cup\sigma(W_2)$. *We denote by* $X=X_{W_1,W_2,V}$ *the unique solution of Sylvester's equation (see Section 3.3):*

$$W_1X-XW_2=D_{W_1^*}VD_{W_2}. \tag{5.30}$$

Then, there exists a contraction $Y\in\mathcal{B}(H_2,H_1)$ *such that:*

$$f(W_1)X-Xf(W_2)=D_{f(W_1)^*}YD_{f(W_2)}.$$

Proof. Let $T=\begin{bmatrix}W_1 & D_{W_1^*}VD_{W_2}\\ 0 & W_2\end{bmatrix}$. Denote $C=D_{W_1^*}VD_{W_2}$. By Parrott's theorem, T is a contraction. Moreover, using (5.30), we have:

$$T=\begin{bmatrix}W_1 & C\\ 0 & W_2\end{bmatrix}=\begin{bmatrix}\mathrm{Id} & -X\\ 0 & \mathrm{Id}\end{bmatrix}\begin{bmatrix}W_1 & 0\\ 0 & W_2\end{bmatrix}\begin{bmatrix}\mathrm{Id} & X\\ 0 & \mathrm{Id}\end{bmatrix}$$

and, thus,

$$f(T)=\begin{bmatrix}\mathrm{Id} & -X\\ 0 & \mathrm{Id}\end{bmatrix}\begin{bmatrix}f(W_1) & 0\\ 0 & f(W_2)\end{bmatrix}\begin{bmatrix}\mathrm{Id} & X\\ 0 & \mathrm{Id}\end{bmatrix}=\begin{bmatrix}f(W_1) & f(W_1)X-Xf(W_2)\\ 0 & f(W_2)\end{bmatrix}.$$

As $\|f\|_\infty\leq 1$, we have $\|f(T)\|\leq 1$ by von Neumann's inequality. Thus, by Parrott's theorem, there exists a contraction $Y\in\mathcal{B}(H_2,H_1)$ such that $f(W_1)X-Xf(W_2)=D_{f(W_1)^*}YD_{f(W_2)}$. □

5.5.2 An operator version of the Beardon-Minda inequality

Using the analogue proof framework as employed in Theorem 5.5.5, we can deduce the following outcome for 3×3 operator matrices:

Theorem 5.5.6. *Let H_1, H_2, H_3 be three Hilbert spaces. Consider three contractions $W_1 \in \mathcal{B}(H_1)$, $W_2 \in \mathcal{B}(H_2)$ and $W_3 \in \mathcal{B}(H_3)$. Let $V_1 \in \mathcal{B}(H_2, H_1)$, $V_2 \in \mathcal{B}(H_3, H_2)$, and $V_3 \in \mathcal{B}(H_3, H_1)$ be contractions. Assume that $\sigma(W_i) \cap \sigma(W_j) = \emptyset$, for all $1 \le i < j \le 3$, and that $f \in \mathcal{H}(\mathbb{D}, \mathbb{D})$ is holomorphic on an open neighborhood of $\sigma(W_1) \cup \sigma(W_2) \cup \sigma(W_3)$. Let X_1, X_2, X_3 be respectively the unique solution of Sylvester's equations:*

$$W_1 X_1 - X_1 W_2 = D_{W_1^*} V_1 D_{W_2}, \tag{5.31}$$

$$W_2 X_2 - X_2 W_3 = D_{W_2^*} V_2 D_{W_3} \text{ and} \tag{5.32}$$

$$W_1 X_3 - X_3 W_3 = B - W_1 X_1 X_2 + X_1 W_2 X_2, \tag{5.33}$$

where:

$$B = \left[D_{W_1^*}(\mathrm{Id} - V_1 V_1^*) D_{W_1^*} \right]^{1/2} V_3 \left[D_{W_3}(\mathrm{Id} - V_2^* V_2) D_{W_3} \right]^{1/2} - D_{W_1^*} V_1 W_2^* V_2 D_{W_3}.$$

Then, there exist three contractions $Y_1 \in \mathcal{B}(H_2, H_1), Y_2 \in \mathcal{B}(H_3, H_2), Y_3 \in \mathcal{B}(H_3, H_1)$ such that:

$$\begin{cases} f(W_1)X_1 - X_1 f(W_2) = D_{f(W_1)^*} Y_1 D_{f(W_2)}, & (5.34) \\ f(W_2)X_2 - X_2 f(W_3) = D_{f(W_2)^*} Y_2 D_{f(W_3)}, & (5.35) \\ f(W_1)X_3 - X_3 f(W_3) = X_1 f(W_2) X_2 - f(W_1) X_1 X_2 & \\ + \left[D_{f(W_1)^*}(\mathrm{Id} - Y_1 Y_1^*) D_{f(W_1)^*} \right]^{1/2} Y_3 \left[D_{f(W_3)}(\mathrm{Id} - Y_2^* Y_2) D_{f(W_3)} \right]^{1/2} & (5.36) \\ \qquad - D_{f(W_1)^*} Y_1 f(W_2)^* Y_2 D_{f(W_3)}. & \end{cases}$$

Proof. Denote $A_1 = D_{W_1^*} V_1 D_{W_2}$ and $A_2 = D_{W_2^*} V_2 D_{W_3}$. Then, according to Theorem 4.2.1, the operator

$$T = \begin{bmatrix} W_1 & A_1 & B \\ 0 & W_2 & A_2 \\ 0 & 0 & W_3 \end{bmatrix}$$

is a contraction. With this notation, X_1, X_2, X_3 are respectively the unique solutions of Sylvester's equations $W_1 X_1 - X_1 W_2 = A_1$, $W_2 X_2 - X_2 W_3 = A_2$ and $W_1 X_3 - X_3 W_3 = B - W_3 X_1 X_2 + X_1 W_2 X_2$.

In view of Theorem 3.3.9 (or, in a more elementary approach, using the computation rules in the Heisenberg group), we can write:

$$T = \begin{bmatrix} \mathrm{Id} & -X_1 & -X_3 \\ 0 & \mathrm{Id} & -X_2 \\ 0 & 0 & \mathrm{Id} \end{bmatrix} \begin{bmatrix} W_1 & 0 & 0 \\ 0 & W_2 & 0 \\ 0 & 0 & W_3 \end{bmatrix} \begin{bmatrix} \mathrm{Id} & X_1 & X_3 + X_1 X_2 \\ 0 & \mathrm{Id} & X_2 \\ 0 & 0 & \mathrm{Id} \end{bmatrix}.$$

This diagonalization allows one to write the 3×3 operator matrix of $f(T)$, which is a contraction by von Neumann's inequality:

$$f(T) = \begin{bmatrix} \mathrm{Id} & -X_1 & X_1 X_2 - X_3 \\ 0 & \mathrm{Id} & -X_2 \\ 0 & 0 & \mathrm{Id} \end{bmatrix} \begin{bmatrix} f(W_1) & 0 & 0 \\ 0 & f(W_2) & 0 \\ 0 & 0 & f(W_3) \end{bmatrix} \begin{bmatrix} \mathrm{Id} & X_1 & X_3 \\ 0 & \mathrm{Id} & X_2 \\ 0 & 0 & \mathrm{Id} \end{bmatrix}.$$

Thus the matrix of $f(T)$ is given by:

$$\begin{bmatrix} f(W_1) & f(W_1)X_1 - X_1 f(W_2) & f(W_1)X_3 - X_3 f(W_3) - X_1 f(W_2)X_2 + f(W_1)X_1X_2 \\ 0 & f(W_2) & f(W_2)X_2 - X_2 f(W_3) \\ 0 & 0 & f(W_3) \end{bmatrix}.$$

We apply again Theorem 4.2.1. □

In the scalar case, the condition (5.36) is equivalent to the Beardon-Minda inequality.

Chapter 6

Other spectral sets

6.1 The unit ball of $\mathbb{C}^n$

In all this section, we denote by $\mathbb{B}_n$ the open unit ball of $\mathbb{C}^n$, *i.e.* $\mathbb{B}_n = \left\{(z_1,\ldots,z_n) \in \mathbb{C}^n : \sum_{i=1}^n |z_i|^2 < 1\right\}$. The set of holomorphic functions on $\mathbb{B}_n$ is supposed to be equipped with the topology of uniform convergence on every compact subset of $\mathbb{B}_n$.

In [30], Hartz, Richter, and Shalit proved an analogous of von Neumann's inequality for *row contractive* tuples of matrices. We start by recalling the definition of a row contractive tuple of operators:

Definition 6.1.1. Let H be an Hilbert space, and let $T = (T_1,\ldots,T_n)$ be a n-tuple of bounded linear operators on H.
We say that T is a *row contraction* if $\|T\|^2 = \|\sum_{i=1}^n T_i T_i^*\| \leq 1$.

Now, we can state the von Neumann type inequality for the unit ball (see [30, Theorem 1.1 and Corollary 3.4]):

Theorem 6.1.2 (Hartz-Richter-Shalit)**.** *Let $d \in \mathbb{N}^*$. There exists a constant C_d such that, for all $n \in \mathbb{N}^*$, for every commuting row contraction $T = (T_1,\ldots,T_n)$ on a Hilbert space of dimension d, and for every polynomial $p \in \mathbb{C}[X_1,\ldots,X_n]$, we have:*

$$\|p(T)\| \leq C_d \sup_{z\in\mathbb{B}_n} |p(z)|.$$

Moreover, we can choose $C_2 = 1$.

Furthermore, they also gave an elementary construction of an holomorphic functional calculus for the unit ball (see [30, Appendix A]), which is a particular case of the Arens-Calderon functional calculus.

Theorem 6.1.3. *Let $T = (T_1,\ldots,T_n)$ be an n-tuple of bounded linear operators on an Hilbert space H, and let $\mathbb{B}$ be an open ball containing the joint spectrum of T. Then, there exists a unique continuous homomorphism $\phi : \mathcal{H}(\mathbb{B},\mathbb{C}) \to \mathcal{B}(H)$ such that $\phi(p) = p(T)$, for every polynomial $p \in \mathbb{C}[X_1,\ldots,X_n]$.*

It turns out that applying Theorem 6.1.2 to a suitably chosen row contractive n-tuple of 2×2 matrices, we can obtain the following Schwarz-Pick type lemma for the unit ball, which is particular case of *Rudin's Schwarz lemma for the unit ball* (take $m=1$ in [55, Theorem 8.1.4]):

Theorem 6.1.4. *Let $f:\mathbb{B}_n\to\mathbb{D}$ be an holomorphic map, and let $\omega, z\in\mathbb{B}_n$.*

Then,

$$\frac{\left|1-\overline{f(\omega)}f(z)\right|^2}{(1-|f(z)|^2)(1-|f(\omega)|^2)}\leq\frac{|1-\langle z,\omega\rangle|^2}{(1-\|z\|^2)(1-\|\omega\|^2)}.$$

In order to show this, we will need the following lemma, which can be seen as a corollary of Lemma 1.2.14:

Lemma 6.1.5. *Let $T=\begin{pmatrix} a & c\\ \overline{c} & b\end{pmatrix}\in\mathcal{M}_2(\mathbb{C})$, with $a,b\in]0,1]$ and $\det(T)=ab-|c|^2\geq 0$.*

Then, $\|T\|\leq 1$ if and only if $|c|^2\leq(1-a)(1-b)$.

Proof. In order to use Lemma 1.2.14, we will write down the Cholesky decomposition of T: Let $S_T=\begin{pmatrix} s_1 & s_2\\ 0 & s_4\end{pmatrix}$ be the Cholesky factorization of T, *i.e.* S_T is an upper-triangular matrix with non-negative diagonal such that $T=S_T^*S_T$.

An easy computation gives:

$$s_1=\sqrt{a}\quad;\quad s_2=\frac{c}{\overline{s_1}}=\frac{c}{\sqrt{a}}\quad;\quad s_4=\sqrt{b-|s_2|^2}=\sqrt{b-\frac{|c|^2}{a}}.$$

Then, $\|T\|\leq 1$ if and only if $\|S_T\|\leq 1$. By Lemma 1.2.14, this happens if and only if

$$\frac{|c|^2}{a}\leq(1-a)\left(1-b+\frac{|c|^2}{a}\right)$$

which is equivalent to

$$|c|^2\leq(1-a)(1-b).$$

□

We will also make use of the following identity, which is a generalization of the key identity (1.3) that has been used several times in this manuscript:

Lemma 6.1.6. *For all ω, z in $\mathbb{C}^n$, we have:*

$$|1-\langle z,\omega\rangle|^2-\left(1-\|z\|^2\right)\left(1-\|\omega\|^2\right)=\|\omega-z\|^2+|\langle\omega,z\rangle|^2-\|\omega\|^2\cdot\|z\|^2 \tag{6.1}$$

Remark 6.1.7. If ω and z are linearly dependent, we have $|\langle\omega,z\rangle|^2=\|\omega\|^2\cdot\|z\|^2$ (this is the equality case in the Cauchy-Schwarz inequality) and, thus, the identity becomes

$$|1-\langle z,\omega\rangle|^2-\left(1-\|z\|^2\right)\left(1-\|\omega\|^2\right)=\|\omega-z\|^2.$$

In particular, for $n=1$, we recover the key identity (1.3).

Now, we are ready to prove Theorem 6.1.4:

Proof of Theorem 6.1.4. Let $\omega = (\omega_1, \dots, \omega_n)$ and $z = (z_1, \dots, z_n)$.

In the trivial cases where $\omega = z$ or $f(\omega) = f(z)$, there is nothing to prove. Therefore, in the following, we will assume that this is not the case.

For every $i \in [\![1, n]\!]$, we consider the matrices:

$$T_i = \begin{pmatrix} \omega_i & \alpha_i \\ 0 & z_i \end{pmatrix},$$

where $\alpha_i = d(\omega_i - z_i)$, for some scalar d to be determined later.

We have $T_i T_j = T_j T_i$, for all i, j, and

$$\sum_{i=1}^{n} T_i T_i^* = \begin{pmatrix} \|\omega\|^2 + \|\alpha\|^2 & \langle \alpha, z \rangle \\ \langle z, \alpha \rangle & \|z\|^2 \end{pmatrix}, \qquad \text{with } \alpha = d(\omega - z).$$

By Lemma 6.1.5, $\|\sum_{i=1}^n T_i T_i^*\| \leq 1$ if and only if:

$$|\langle \alpha, z \rangle|^2 \leq \left(1 - \|\omega\|^2 - \|\alpha\|^2\right)\left(1 - \|z\|^2\right),$$

which is equivalent to:

$$|d|^2 \cdot \left[\left| \langle \omega, z \rangle - \|z\|^2 \right|^2 + \|\omega - z\|^2 \left(1 - \|z\|^2\right) \right] \leq \left(1 - \|\omega\|^2\right)\left(1 - \|z\|^2\right)$$

or, after simplification, to:

$$|d|^2 \cdot \left[\|\omega - z\|^2 + |\langle \omega, z \rangle|^2 - \|\omega\|^2 \cdot \|z\|^2 \right] \leq \left(1 - \|\omega\|^2\right)\left(1 - \|z\|^2\right).$$

Using (6.1), we obtain that $\|\sum_{i=1}^n T_i T_i^*\| \leq 1$ if and only if:

$$|d|^2 \leq \frac{\left(1 - \|\omega\|^2\right)\left(1 - \|z\|^2\right)}{|1 - \langle \omega, z \rangle|^2 - (1 - \|z\|^2)(1 - \|\omega\|^2)}.$$

In the following, we will take $d = \sqrt{\frac{(1 - \|\omega\|^2)(1 - \|z\|^2)}{|1 - \langle \omega, z \rangle|^2 - (1 - \|z\|^2)(1 - \|\omega\|^2)}}$.

For $f : \mathbb{B}_n \to \mathbb{D}$, we have:

$$f(T_1, \dots, T_n) = \begin{pmatrix} f(\omega) & d\,[f(\omega) - f(z)] \\ 0 & f(z) \end{pmatrix}$$

and, by Theorem 6.1.2, $\|f(T_1, \dots, T_n)\| \leq 1$. By Lemma 1.2.14, this is equivalent to:

$$\frac{|f(\omega) - f(z)|^2}{(1 - |f(\omega)|^2)(1 - |f(z)|^2)} \leq \frac{1}{|d|^2}.$$

Using the identity (1.3), we get:

$$\frac{\left|1-\overline{f(\omega)}f(z)\right|^2}{(1-|f(z)|^2)(1-|f(\omega)|^2)} \leq \frac{|1-\langle z,\omega\rangle|^2}{(1-\|z\|^2)(1-\|\omega\|^2)}.$$

□

6.2 The annulus

For $r \in [0,1[$, we denote by $\mathbb{A}_r := \{z \in \mathbb{C} : r < |z| < 1\}$ the annulus of radii r and 1. Moreover, if H is an Hilbert space, we denote by $\mathcal{F}_r := \left\{T \in \mathcal{B}(H) : r^2 T^{-1}\left(T^{-1}\right)^* + TT^* \leq (r^2+1)\cdot \mathrm{Id},\ \sigma(T) \subset \mathbb{A}_r\right\}$.

Recall that if S and T are two operators on a Hilbert space H, the notation $S \leq T$ means that $T - S$ is a positive operator. If $T = \lambda \cdot \mathrm{Id}$, for some $\lambda \in \mathbb{C}$, we write sometimes $S \leq \lambda$.

In [61], Tsikalas has shown that $\mathbb{A}_r$ is a $\sqrt{2}$-spectral set for every element of $\mathcal{F}_r$:

Theorem 6.2.1 (Tsikalas). *For every bounded holomorphic function $\phi \in H^\infty(\mathbb{A}_r)$, for every $T \in \mathcal{F}_r$,*

$$\|\phi(T)\| \leq \sqrt{2}\sup\{|\phi(z)| : z \in \mathbb{A}_r\} \tag{6.2}$$

where the constant $\sqrt{2}$ is the best possible.

Remark 6.2.2. In [61], it is proved that the constant $\sqrt{2}$ is the best possible if we require (6.2) to be true for every bounded linear operator $T \in \mathcal{F}_r$. However, if we restrict ourselves to 2×2 matrices, this constant may not be optimal.

We can characterize $\mathcal{F}_r$ in the case where $H = \mathbb{C}^2$. To do this, we first generalize Lemma 6.1.5 as follows:

Lemma 6.2.3. *Let $T = \begin{pmatrix} a & c \\ \overline{c} & b\end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$, with $a > 0$, $b \in \mathbb{R}$, and $\det(T) = ab - |c|^2 \geq 0$, and let $\rho > 0$.*

Then, $T \leq \rho \cdot \mathrm{Id}$ if and only if $|c|^2 \leq (\rho - a)(\rho - b)$.

Proof. As in Lemma 6.1.5, we consider the Cholesky factorization $T = S_T^* S_T$. If we denote $S_T = \begin{pmatrix} s_1 & s_2 \\ 0 & s_4\end{pmatrix}$, we have:

$$s_1 = \sqrt{a} \quad ; \quad s_2 = \frac{c}{s_1} = \frac{c}{\sqrt{a}} \quad ; \quad s_4 = \sqrt{b - |s_2|^2} = \sqrt{b - \frac{|c|^2}{a}}.$$

Now, we observe that $T \leq \rho$ if and only if, for all $x \in \mathbb{C}^2$, $\rho\|x\|^2 - \|S_T x\|^2 \geq 0$, which is equivalent to saying that $\|S_T\| \leq \sqrt{\rho}$ or, equivalently, that $\frac{1}{\sqrt{\rho}} S_T$ is a contraction.

By Lemma 1.2.14, this happens if and only if

$$\frac{|c|^2}{\rho a} \leq \left(1 - \frac{a}{\rho}\right)\left(1 - \frac{b}{\rho} + \frac{|c|^2}{\rho a}\right),$$

which is equivalent to

$$|c|^2 \le (\rho - a)(\rho - b).$$

□

Now, we are ready to prove the following lemma:

Proposition 6.2.4. *Let* $T = \begin{pmatrix} \omega_1 & \alpha \\ 0 & \omega_2 \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$*, with* $\omega_1, \omega_2 \in \mathbb{A}_r$*.*

$T \in \mathcal{F}_r$ *if and only if*

$$|\alpha|^2 \le \frac{(1-|\omega_1|^2)(1-|\omega_2|^2)(|\omega_1|^2 - r^2)(|\omega_2|^2 - r^2)}{|r^2 - \overline{\omega_2}\omega_1|^2 + r^2(1-|\omega_1|^2)(1-|\omega_2|^2)}.$$

Proof. Let $B_r := r^2 T^{-1}\left(T^{-1}\right)^* + TT^*$. First of all, observe that B_r is positive as it is the sum of two positive operators. Moreover, we have:

$$B_r = \frac{r^2}{|\omega_1\omega_2|^2}\begin{pmatrix} |\omega_2|^2 + |\alpha|^2 & -\alpha\overline{\omega_1} \\ -\overline{\alpha}\omega_1 & |\omega_1|^2 \end{pmatrix} + \begin{pmatrix} |\omega_1|^2 + |\alpha|^2 & \alpha\overline{\omega_2} \\ \overline{\alpha}\omega_2 & |\omega_2|^2 \end{pmatrix}.$$

By Lemma 6.2.3, $B_r \le (r^2+1)\cdot \mathrm{Id}$ if and only if

$$\left|\alpha\overline{\omega_2} - \frac{r^2}{|\omega_1\omega_2|^2}\alpha\overline{\omega_1}\right|^2 \le \left[r^2 + 1 - \frac{r^2}{|\omega_1|^2} - |\omega_1|^2 - |\alpha|^2\left(\frac{r^2}{|\omega_1\omega_2|^2} + 1\right)\right]\left[r^2 + 1 - \frac{r^2}{|\omega_2|^2} - |\omega_2|^2\right] \tag{6.3}$$

which is equivalent to

$$\begin{aligned} |\alpha|^2 \cdot \left[\left|\overline{\omega_2} - \frac{r^2}{|\omega_1\omega_2|^2}\overline{\omega_1}\right|^2 + \left(\frac{r^2}{|\omega_1\omega_2|^2} + 1\right)\left(r^2 + 1 - \frac{r^2}{|\omega_2|^2} - |\omega_2|^2\right)\right] \\ \le \left(r^2 + 1 - \frac{r^2}{|\omega_1|^2} - |\omega_1|^2\right)\left(r^2 + 1 - \frac{r^2}{|\omega_2|^2} - |\omega_2|^2\right) \end{aligned} \tag{6.4}$$

Let

$$C := \left|\overline{\omega_2} - \frac{r^2}{|\omega_1\omega_2|^2}\overline{\omega_1}\right|^2 + \left(\frac{r^2}{|\omega_1\omega_2|^2} + 1\right)\left(r^2 + 1 - \frac{r^2}{|\omega_2|^2} - |\omega_2|^2\right)$$

and

$$D := \left(r^2 + 1 - \frac{r^2}{|\omega_1|^2} - |\omega_1|^2\right)\left(r^2 + 1 - \frac{r^2}{|\omega_2|^2} - |\omega_2|^2\right).$$

On the one hand, we have:

$$\begin{aligned} C &= |\omega_2|^2 + \frac{r^4}{|\omega_1|^2\cdot|\omega_2|^4} - 2\mathrm{Re}\left(\frac{r^2}{\overline{\omega_1}\omega_2}\right) + \frac{r^4}{|\omega_1\omega_2|^2} + \frac{r^2}{|\omega_1\omega_2|^2} - \frac{r^4}{|\omega_1|^2\cdot|\omega_2|^2} - \frac{r^2}{|\omega_1|^2} + r^2 + 1 - \frac{r^2}{|\omega_2|^2} - |\omega_2|^2 \\ &= \left|1 - \frac{r^2}{\omega_1\overline{\omega_2}}\right|^2 + r^2\left(1 - \frac{1}{|\omega_1|^2} - \frac{1}{|\omega_2|^2} + \frac{1}{|\omega_1\omega_2|^2}\right) \\ &= \frac{1}{|\omega_1\omega_2|^2}\left[|r^2 - \overline{\omega_2}\omega_1|^2 + r^2(1-|\omega_1|^2)(1-|\omega_2|^2)\right]. \end{aligned}$$

On the other hand, we have:

$$D = \frac{1}{|\omega_1\omega_2|^2}(1-|\omega_1|^2)(|\omega_1|^2-r^2)(1-|\omega_2|^2)(|\omega_2|^2-r^2).$$

Thus, (6.4) is equivalent to:

$$|\alpha|^2 \leq \frac{(1-|\omega_1|^2)(1-|\omega_2|^2)(|\omega_1|^2-r^2)(|\omega_2|^2-r^2)}{|r^2-\overline{\omega_2}\omega_1|^2+r^2(1-|\omega_1|^2)(1-|\omega_2|^2)}.$$

□

Remark 6.2.5. In the case where $r=0$, $\mathcal{F}_r$ is the set of invertible contractions. If we take $r=0$ in Proposition 6.2.4, we recover the criteria of contractivity for 2×2 matrices (Lemma 1.2.14).

Now, let $K_2>0$ be the smallest constant such that $\mathbb{A}_r$ is a K_2–spectral set for every 2×2 matrix of $\mathcal{F}_r$ (by Tsikalas' inequality, we know that $K_2\leq\sqrt{2}$).

Applying the inequality

$$\|\phi(T)\| \leq K_2 \sup\{|\phi(z)| : z\in\mathbb{A}_r\}, \quad \forall\,\phi\in H^\infty(\mathbb{A}_r)$$

to the matrix $T=\begin{pmatrix}\omega_1 & \alpha\\ 0 & \omega_2\end{pmatrix}$, where $\omega_1,\omega_2\in\mathbb{A}_r$, and

$$\alpha=\sqrt{\frac{(1-|\omega_1|^2)(1-|\omega_2|^2)(|\omega_1|^2-r^2)(|\omega_2|^2-r^2)}{|r^2-\overline{\omega_2}\omega_1|^2+r^2(1-|\omega_1|^2)(1-|\omega_2|^2)}},$$

we obtain the following Schwarz-Pick type inequality:

Corollary 6.2.6. *Let $\omega_1,\omega_2\in\mathbb{A}_r$, and let $f:\mathbb{A}_r\to\mathbb{D}$ be an holomorphic map. Then, we have:*

$$\frac{|f(\omega_1)-f(\omega_2)|^2}{\left(K_2^2-|f(\omega_1)|^2\right)\left(K_2^2-|f(\omega_2)|^2\right)} \leq \frac{|\omega_1-\omega_2|^2}{K_2^2}\cdot\frac{|r^2-\overline{\omega_2}\omega_1|^2+r^2(1-|\omega_1|^2)(1-|\omega_2|^2)}{(1-|\omega_1|^2)(1-|\omega_2|^2)(|\omega_1|^2-r^2)(|\omega_2|^2-r^2)} \tag{6.5}$$

or, equivalently,

$$\begin{aligned}\frac{|f(\omega_1)-f(\omega_2)|^2}{\left(K_2^2-|f(\omega_1)|^2\right)\left(K_2^2-|f(\omega_2)|^2\right)} &\leq \frac{r^2\cdot|\omega_1-\omega_2|^4}{K_2^2(1-|\omega_1|^2)(1-|\omega_2|^2)(|\omega_1|^2-r^2)(|\omega_2|^2-r^2)}\\ &\quad+\frac{r^2\cdot|\omega_1-\omega_2|^2}{K_2^2(|\omega_1|^2-r^2)(|\omega_2|^2-r^2)}+\frac{|\omega_1-\omega_2|^2}{K_2^2(1-|\omega_1|^2)(1-|\omega_2|^2)}\end{aligned} \tag{6.6}$$

Proof. We proceed as in the proof of Schwarz-Pick's lemma introduced in Section 1.2.3, except that this time, Tsikalas' inequality gives that $\|f(T)\|\leq\sqrt{2}$ or, equivalently, that $\left\|\frac{1}{\sqrt{2}}f(T)\right\|\leq 1$.

For the second inequality, we use the identity

$$|r^2-\overline{\omega_2}\omega_1|^2 = r^2\cdot|\omega_1-\omega_2|^2+(|\omega_1|^2-r^2)(|\omega_2|^2-r^2) \tag{6.7}$$

□

Bibliography

[1] Jim Agler and John E. McCarthy. “Distinguished Varieties”. In: *Acta Mathematica* 194.2 (Jan. 2005), pp. 133–153. ISSN: 0001-5962, 1871-2509. DOI: 10.1007/BF02393219.

[2] Jim Agler and John E. McCarthy. *Pick Interpolation and Hilbert Function Spaces*. Vol. 44. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2002. xx+308. ISBN: 978-0-8218-2898-4. DOI: 10.1090/gsm/044.

[3] Jim Agler, John Edward McCarthy, and Nicholas Young. *Operator Analysis—Hilbert Space Methods in Complex Analysis*. Vol. 219. Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 2020. xv+375. ISBN: 978-1-108-48544-9. DOI: 10.1017/9781108751292.

[4] Tsuyoshi Ando and Ky Fan. “Pick-Julia Theorems for Operators”. In: *Mathematische Zeitschrift* 168.1 (1979), pp. 23–34. ISSN: 0025-5874,1432-1823. DOI: 10.1007/BF01214433.

[5] T. Andô. “On a Pair of Commutative Contractions”. In: *Acta Universitatis Szegediensis. Acta Scientiarum Mathematicarum* 24 (1963), pp. 88–90. ISSN: 0001-6969.

[6] Gr. Arsene and A. Gheondea. “Completing Matrix Contractions”. In: *Journal of Operator Theory* 7.1 (1982), pp. 179–189. ISSN: 0379-4024.

[7] Line Baribeau, Patrice Rivard, and Elias Wegert. “On Hyperbolic Divided Differences and the Nevanlinna-Pick Problem”. In: *Computational Methods and Function Theory* 9.2 (2009), pp. 391–405. ISSN: 1617-9447. DOI: 10.1007/BF03321735.

[8] A. F. Beardon and D. Minda. “A Multi-Point Schwarz-Pick Lemma”. In: *Journal d’Analyse Mathématique* 92 (2004), pp. 81–104. ISSN: 0021-7670,1565-8538. DOI: 10.1007/BF02787757.

[9] Rajendra Bhatia and Peter Rosenthal. “How and Why to Solve the Operator Equation AX-XB=Y”. In: *The Bulletin of the London Mathematical Society* 29.1 (1997), pp. 1–21. ISSN: 0024-6093,1469-2120. DOI: 10.1112/S0024609396001828.

[10] Kelly Bickel et al. “Crouzeix’s Conjecture and Related Problems”. In: *Computational Methods and Function Theory* 20.3-4 (2020), pp. 701–728. ISSN: 1617-9447,2195-3724. DOI: 10.1007/s40315-020-00350-9.

[11] David A. Brannan, Matthew F. Esplen, and Jeremy J. Gray. *Geometry*. Higher Education from Cambridge University Press. Dec. 22, 2011. DOI: 10.1017/CBO9781139003001.

[12] Kyung Hyun Cho, Seong-A Kim, and Toshiyuki Sugawa. “On a Multi-Point Schwarz-Pick Lemma”. In: *Computational Methods and Function Theory* 12.2 (2012), pp. 483–499. ISSN: 1617-9447. DOI: 10.1007/BF03321839.

[13] David Cohen. *Dilations of matrices*. MA thesis. 2015. arXiv: 1503.07334 [math.FA]. URL: https://arxiv.org/abs/1503.07334.

[14] John B. Conway. *A Course in Functional Analysis*. 2nd ed. Vol. 96. Graduate Texts in Mathematics. Springer-Verlag, New York, 1990. xvi+399. ISBN: 978-0-387-97245-9.

[15] M. J. Crabb and A. M. Davie. “Von Neumann’s Inequality for Hilbert Space Operators”. In: *The Bulletin of the London Mathematical Society* 7 (1975), pp. 49–50. ISSN: 0024-6093,1469-2120. DOI: 10.1112/blms/7.1.49.

[16] Chandler Davis. “Explicit Functional Calculus”. In: *Linear Algebra and its Applications* 6 (1973), pp. 193–199. ISSN: 0024-3795,1873-1856. DOI: 10.1016/0024-3795(73)90019-0.

[17] Chandler Davis, W. M. Kahan, and H. F. Weinberger. “Norm-Preserving Dilations and Their Applications to Optimal Error Bounds”. In: *SIAM Journal on Numerical Analysis* 19.3 (1982), pp. 445–469. ISSN: 0036-1429. DOI: 10.1137/0719029.

[18] Carl de Boor. “Divided Differences”. In: *Surveys in Approximation Theory* 1 (2005), pp. 46–69. ISSN: 1555-578X.

[19] R. G. Douglas. “On Majorization, Factorization, and Range Inclusion of Operators on Hilbert Space”. In: *Proceedings of the American Mathematical Society* 17 (1966), pp. 413–415. ISSN: 0002-9939,1088-6826. DOI: 10.2307/2035178.

[20] S. W. Drury. “Remarks on von Neumann’s Inequality”. In: *Banach Spaces, Harmonic Analysis, and Probability Theory (Storrs, Conn., 1980/1981)*. Vol. 995. Lecture Notes in Math. Springer, Berlin, 1983, pp. 14–32. ISBN: 978-3-540-12314-9. DOI: 10.1007/BFb0061886.

[21] Nelson Dunford and Jacob T. Schwartz. *Linear Operators. Part I*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1988. xiv+858. ISBN: 978-0-471-60848-6.

[22] Ky Fan. “Analytic Functions of a Proper Contraction”. In: *Mathematische Zeitschrift* 160.3 (Oct. 1, 1978), pp. 275–290. ISSN: 1432-1823. DOI: 10.1007/BF01237041.

[23] Ciprian Foias and Arthur E. Frazho. *The Commutant Lifting Approach to Interpolation Problems*. Vol. 44. Operator Theory: Advances and Applications. Birkhäuser Verlag, Basel, 1990. xxiv+632. ISBN: 978-3-7643-2461-2. DOI: 10.1007/978-3-0348-7712-1.

[24] Emmanuel Fricain and Javad Mashreghi. *The Theory of $\mathcal{H}(b)$ Spaces. Vol. 1*. Vol. 20. New Mathematical Monographs. Cambridge University Press, Cambridge, 2016. xix+681. ISBN: 978-1-107-02777-0. DOI: 10.1017/CBO9781139226752.

[25] Siqi Fu and Bernard Russo. “Spectral Domains in Several Complex Variables”. In: *The Rocky Mountain Journal of Mathematics* 27.4 (1997), pp. 1095–1116. ISSN: 0035-7596,1945-3795. DOI: 10.1216/rmjm/1181071863.

[26] Stephan Ramon Garcia, Javad Mashreghi, and William T. Ross. *Finite Blaschke Products and Their Connections*. Cham: Springer International Publishing, 2018. ISBN: 978-3-319-78246-1 978-3-319-78247-8. DOI: 10.1007/978-3-319-78247-8.

[27] Stephan Ramon Garcia, Javad Mashreghi, and William T. Ross. *Introduction to Model Spaces and Their Operators*. Vol. 148. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2016. xv+322. ISBN: 978-1-107-10874-5. DOI: 10.1017/CBO9781316258231.

[28] Stephan Ramon Garcia and William T. Ross. “Model Spaces: A Survey”. In: *Invariant Subspaces of the Shift Operator*. Vol. 638. Contemp. Math. Amer. Math. Soc., Providence, RI, 2015, pp. 197–245. ISBN: 978-1-4704-1045-2. DOI: 10.1090/conm/638/12811.

[29] Rajeev Gupta. “The Carathéodory-Féjer Interpolation Problems and the von-Neumann Inequality”. PhD thesis. Department of Mathematics, Indian Institute of Science, Bangalore, 2015.

[30] Michael Hartz, Stefan Richter, and Orr Moshe Shalit. “Von Neumann’s Inequality for Row Contractive Matrix Tuples”. In: *Mathematische Zeitschrift* 301.4 (Aug. 1, 2022), pp. 3877–3894. ISSN: 1432-1823. DOI: 10.1007/s00209-022-03044-1.

[31] Nicholas J. Higham. *Functions of Matrices*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008. xx+425. ISBN: 978-0-89871-646-7. DOI: 10.1137/1.9780898717778.

[32] John A. Holbrook. “Schur Norms and the Multivariate von Neumann Inequality”. In: *Recent Advances in Operator Theory and Related Topics (Szeged, 1999)*. Vol. 127. Oper. Theory Adv. Appl. Birkhäuser, Basel, 2001, pp. 375–386. ISBN: 978-3-7643-6607-0.

[33] Danko R. Jocić. “Noncommutative Schwarz Lemma and Pick-Julia Theorems for Generalized Derivations in Norm Ideals of Compact Operators”. In: *Complex Analysis and Operator Theory* 16.8 (2022), Paper No. 111, 23. ISSN: 1661-8254,1661-8262. DOI: 10.1007/s11785-022-01287-8.

[34] Michael T. Jury, Greg Knese, and Scott McCullough. “Nevanlinna–Pick Interpolation on Distinguished Varieties in the Bidisk”. In: *Journal of Functional Analysis* 262.9 (May 1, 2012), pp. 3812–3838. ISSN: 0022-1236. DOI: 10.1016/j.jfa.2012.01.028.

[35] Richard V. Kadison and John R. Ringrose. *Fundamentals of the Theory of Operator Algebras. Vol. I*. Vol. 100. Pure and Applied Mathematics. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, 1983. xv+398. ISBN: 978-0-12-393301-0.

[36] Dmitry S. Kaliuzhnyi-Verbovetskyi and Victor Vinnikov. *Foundations of Free Noncommutative Function Theory*. Vol. 199. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2014. vi+183. DOI: 10.1090/surv/199.

[37] Seong-A Kim and Toshiyuki Sugawa. “Invariant Differential Operators Associated with a Conformal Metric”. In: *Michigan Mathematical Journal* 55.2 (2007), pp. 459–479. ISSN: 0026-2285,1945-2365. DOI: 10.1307/mmj/1187647003.

[38] Greg Knese. “A Schwarz Lemma on the Polydisk”. In: *Proceedings of the American Mathematical Society* (2007).

[39] Greg Knese. “The von Neumann Inequality for 3×3 Matrices”. In: *Bulletin of the London Mathematical Society* 48.1 (2016), pp. 53–57. ISSN: 0024-6093,1469-2120. DOI: 10.1112/blms/bdv087.

[40] Vitalii G. Kurbatov and Irina V. Kurbatova. “Green’s Function of the Problem of Bounded Solutions in the Case of a Block Triangular Coefficient”. In: *Operators and Matrices* 13.4 (2019), pp. 981–1001. ISSN: 1846-3886,1848-9974. DOI: 10.7153/oam-2019-13-69.

[41] Vladimir Müller. *Spectral Theory of Linear Operators and Spectral Systems in Banach Algebras*. Vol. 139. Operator Theory: Advances and Applications. Birkhäuser Verlag, Basel, 2003. x+381. ISBN: 978-3-7643-6912-5. DOI: 10.1007/978-3-0348-7788-6.

[42] Rolf Nevanlinna. *Über beschränkte Funktionen die in gegebenen Punkten vorgeschriebene Werte annehmen*. Buchdruckerei-a.-g. Sana, 1919. 71 pp. Google Books: VDtXmgEACAAJ.

[43] Bruce P. Palka. *An Introduction to Complex Function Theory*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1991. xviii+559. ISBN: 978-0-387-97427-9. DOI: 10.1007/978-1-4612-0975-1.

[44] Stephen Parrott. “On a Quotient Norm and the Sz.-Nagyþinspace -Þinspace Foiaş Lifting Theorem”. In: *Journal of Functional Analysis* 30.3 (1978), pp. 311–328. ISSN: 0022-1236. DOI: 10.1016/0022-1236(78)90060-5.

[45] Stephen Parrott. “Unitary Dilations for Commuting Contractions”. In: *Pacific Journal of Mathematics* 34 (1970), pp. 481–490. ISSN: 0030-8730,1945-5844.

[46] Vern Paulsen. *Completely Bounded Maps and Operator Algebras*. Vol. 78. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2002. xii+300. ISBN: 978-0-521-81669-4.

[47] Vern I. Paulsen, Gelu Popescu, and Dinesh Singh. “On Bohr’s Inequality”. In: *Proceedings of the London Mathematical Society. Third Series* 85.2 (2002), pp. 493–512. ISSN: 0024-6115,1460-244X. DOI: 10.1112/S0024611502013692.

[48] Vern I. Paulsen and Mrinal Raghupathi. *An Introduction to the Theory of Reproducing Kernel Hilbert Spaces*. Vol. 152. Cambridge Studies in Advanced Mathematics. Cambridge

University Press, Cambridge, 2016. x+182. ISBN: 978-1-107-10409-9. DOI: 10.1017/CBO9781316219232.

[49] Georg Pick. "Über Die Beschränkungen Analytischer Funktionen, Welche Durch Vorgegebene Funktionswerte Bewirkt Werden". In: *Mathematische Annalen* 77.1 (1915), pp. 7–23. ISSN: 0025-5831,1432-1807. DOI: 10.1007/BF01456817.

[50] Gilles Pisier. *Similarity Problems and Completely Bounded Maps*. expanded. Vol. 1618. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2001. viii+198. ISBN: 978-3-540-41524-4. DOI: 10.1007/b55674.

[51] Marvin Rosenblum. "On the Operator Equation BX-XA=Q". In: *Duke Mathematical Journal* 23 (1956), pp. 263–269. ISSN: 0012-7094,1547-7398.

[52] Marvin Rosenblum and James Rovnyak. *Hardy Classes and Operator Theory*. Dover Publications, Inc., Mineola, NY, 1997. xiv+161. ISBN: 978-0-486-69536-5.

[53] James Rovnyak. "A Converse to von Neumann's Inequality". In: *Proceedings of the American Mathematical Society* 84.3 (1982), pp. 370–372. ISSN: 0002-9939,1088-6826. DOI: 10.2307/2043562.

[54] Walter Rudin. *Function Theory in Polydiscs*. W. A. Benjamin, Inc., New York-Amsterdam, 1969. vii+188.

[55] Walter Rudin. *Function Theory in the Unit Ball of* $\mathbb{C}^n$. New York, NY: Springer, 1980. ISBN: 978-3-540-68272-1 978-3-540-68276-9. DOI: 10.1007/978-3-540-68276-9.

[56] St Ruscheweyh. "Two Remarks on Bounded Analytic Functions". In: *Serdica* 11.2 (1985), pp. 200–202.

[57] J. Schur. "Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind." In: *Journal für die reine und angewandte Mathematik* 148 (1918), pp. 122–145. ISSN: 0075-4102; 1435-5345/e. URL: https://eudml.org/doc/149476.

[58] Barry Simon. *Loewner's Theorem on Monotone Matrix Functions*. Vol. 354. Grundlehren Der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Cham, 2019. xi+459. ISBN: 978-3-030-22421-9 978-3-030-22422-6. DOI: 10.1007/978-3-030-22422-6.

[59] Béla Sz.-Nagy et al. *Harmonic Analysis of Operators on Hilbert Space*. enlarged. Universitext. Springer, New York, 2010. xiv+474. ISBN: 978-1-4419-6093-1. DOI: 10.1007/978-1-4419-6094-8.

[60] Joseph L. Taylor. "A joint spectrum for several commuting operators". In: *J. Functional Analysis* 6 (1970), pp. 172–191. ISSN: 0022-1236. DOI: 10.1016/0022-1236(70)90055-8. URL: https://doi.org/10.1016/0022-1236(70)90055-8.

[61] Georgios Tsikalas. "A von Neumann Type Inequality for an Annulus". In: *Journal of Mathematical Analysis and Applications* 506.2 (2022), Paper No. 125714, 12. ISSN: 0022-247X,1096-0813. DOI: 10.1016/j.jmaa.2021.125714.

[62] N. Th. Varopoulos. "On an Inequality of von Neumann and an Application of the Metric Theory of Tensor Products to Operators Theory". In: *Journal of Functional Analysis* 16 (1974), pp. 83–100. ISSN: 0022-1236. DOI: 10.1016/0022-1236(74)90071-8.

[63] Johann von Neumann. "Eine Spektraltheorie Für Allgemeine Operatoren Eines Unitären Raumes". In: *Mathematische Nachrichten* 4 (1951), pp. 258–281. ISSN: 0025-584X,1522-2616. DOI: 10.1002/mana.3210040124.

[64] James S. Walker. *Fourier Analysis*. Oxford University Press, New York, 1988. xx+440. ISBN: 978-0-19-504300-6.

[65] Shinji Yamashita. "The Pick Version of the Schwarz Lemma and Comparison of the Poincaré Densities". In: *Annales Academiae Scientiarum Fennicae. Series A I. Mathematica* 19.2 (1994), pp. 291–322. ISSN: 0066-1953.

[66] Nicholas Young. *An Introduction to Hilbert Space*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1988. x+239. ISBN: 978-0-521-33071-8 978-0-521-33717-5. DOI: 10.1017/CBO9781139172011.

Table des matières

Model spaces, spectral sets and Schwarz-Pick type inequalities

Abstract

The Schwarz-Pick inequality for holomorphic functions of one variable is a classical topic in complex analysis and, more specifically, in hyperbolic geometry. Schwarz-Pick type inequalities for three points and, subsequently, for n points in the open unit disc have been established by Beardon and Minda, and by Baribeau, Rivard, and Wegert, respectively. The aim of this PhD thesis is to study Schwarz-Pick type inequalities from an operator-theoretic perspective. By employing the von Neumann inequality for operators on Hilbert spaces and its generalizations, the Schwarz-Pick inequality and the Beardon-Minda inequality are derived in a unified manner. The connection with model operators is explored, and new Schwarz-Pick type inequalities for functions in one or several complex variables are established. The thesis also focuses on operator versions of Schwarz-Pick type inequalities, building on the work of Ky Fan and D. Joćić. To this end, we discuss related topics such as explicit holomorphic functional calculus, non-commutative divided differences, and contractivity criteria for matrices. The manuscript concludes with a brief exploration of some other spectral sets.

Keywords: functional analysis, operator theory, mathematical analysis, complex analysis, holomorphic functions, spectral theory, Hilbert spaces, hyperbolic geometry, Schwarz-Pick lemma

---

Espaces modèles, ensembles spectraux et inégalités de type Schwarz-Pick

Résumé

L'inégalité de Schwarz-Pick pour une fonction holomorphe d'une variable est un sujet classique en analyse complexe et, plus spécifiquement, en géométrie hyperbolique. Des inégalités de type Schwarz-Pick pour trois points et, ensuite, pour n points du disque unité ouvert ont été établies respectivement par Beardon et Minda, et par Baribeau, Rivard, et Wegert. L'objectif de cette thèse de doctorat est d'étudier les inégalités de type Schwarz-Pick dU point de vue de la théorie des opérateurs. En utilisant l'inégalité de von Neumann pour les opérateurs sur les espaces de Hilbert et ses généralisations, les inégalités de Schwarz-Pick et de Beardon-Minda sont obtenues de manière unifiée. Le lien avec les opérateurs modèles est mis en évidence, et de nouvelles inégalités de type Schwarz-Pick en une ou plusieurs variables complexes sont établies. Nous nous concentrons également sur les versions opérateurs des inégalités de type Schwarz-Pick, en suivant les travaux de Ky Fan et D. Joćić. Pour ce faire, nous discutons d'autres sujets d'intérêt, tels que le calcul fonctionnel holomorphe explicite, les différences divisées non commutatives et les critères de contractivité pour les matrices. Le manuscrit se termine par un aperçu rapide d'autres ensembles spectraux.

Mots clés : analyse fonctionnelle, théorie des opérateurs, analyse mathématique, analyse complexe, fonctions holomorphes, théorie spectrale, espaces de Hilbert, géométrie hyperbolique, lemme de Schwarz-Pick

---

Laboratoire Paul Painlevé

CNRS UMR 8524 – Cité Scientifique, Bâtiment M2 – 59655 Villeneuve d'Ascq Cedex