Departamento de Matemáticas
Universidad Autónoma de Madrid

Curso 2012–2013

Trabajo Fin de Máster

Complex analysis techniques in the spectral theory of linear operators

Daniel Estévez Sánchez

September 13, 2013

Advisor: Dmitry Yakubovich

Contents

Acknowledgements ii

Preface iii

Notation and preliminaries ix

1. Dilations of a single operator: Sz.-Nagy–Foiaș theory 1
- 1.1. Extensions, liftings and dilations . . . 1
- 1.2. Vector valued Hardy spaces . . . 3
- 1.3. Unilateral and bilateral shifts . . . 6
- 1.4. Isometric and unitary dilations . . . 8
- 1.5. The Sz.-Nagy–Foiaş functional model of a $C_{\cdot 0}$ contraction . . . 11
- 1.6. Invariant subspaces . . . 15

2. Dilations of several operators 18
- 2.1. The Commutant Lifting Theorem . . . 18
- 2.2. An application: Nevanlinna-Pick interpolation . . . 21
- 2.3. Extensions, liftings and dilations . . . 22
- 2.4. Dilations of a pair of operators: Andô's theorem . . . 23
- 2.5. Dilations of a system of isometries . . . 26
- 2.6. Von Neumann's inequality . . . 27
- 2.7. Regular dilations . . . 30
- 2.8. Non-existence of dilations . . . 33
- 2.9. A dilation existence result for diagonalizable contractions . . . 35

3. Livšic-Vinnikov theory 38
- 3.1. Colligations of a single operator . . . 38
- 3.2. Colligations of several commuting operators . . . 41
- 3.3. Vessels of several commuting operators . . . 44
- 3.4. Vessels of two commuting operators . . . 48

4. Separating structures 52
- 4.1. Operator pools . . . 52
- 4.2. The discriminant curve . . . 54
- 4.3. Affine separating structures . . . 56
- 4.4. Orthogonal separating structures . . . 62
- 4.5. Generalized compression . . . 72

A. Functional calculus 83
- A.1. Riesz-Dunford functional calculus . . . 83
- A.2. Spectral theorem and functional calculus for normal operators . . . 85

Acknowledgements

I would like to thank the Mathematics Department of the Universidad Autónoma de Madrid, which has supported me during this academical year with a grant for the study of the Master's Degree. I am also very grateful to my advisor, Dmitry Yakubovich, for all his valuable insight during the completion of this work and for his thorough reading of the manuscript.

Finally, I would also like to thank my teachers and colleagues this year, who have made the Master's Degree courses a very enjoyable experience.

Preface

It is known that the theory of linear operators on a Hilbert space is intimately related with the complex analysis. In this work, we will expose the classical Sz.-Nagy–Foiaş theory and its relations with the theory of Hardy H^p spaces of the disk. We will also treat commutative tuples of operators (selfadjoint or non-selfadjoint) and how they can be related to the function theory on complex algebraic curves.

This work is divided in four chapters, which are devoted to the following four topics respectively: the Sz.-Nagy–Foiaş theory of contractions, the theory of dilations of several operators, the Livŝic-Vinnikov theory of commuting non-selfadjoint operators, and the theory of separating structures.

The first chapter is an introduction to the Sz.-Nagy–Foiaş theory of contractions on a Hilbert space. From a certain point of view, it is based on the spectral theory for isometries. Given an isometry, the Kolmogorov-von Neumann-Wold decomposition theorem states that it is a direct sum of a unitary operator and a unilateral shift. The unitary operator is well understood, via the spectral theory for normal operators, and the unilateral shift can be realized as the operator M_z of multiplication by the independent variable z in some Hardy space $H^2(\mathscr{U})$ of $\mathscr{U}$-valued analytic functions on the disk, where $\mathscr{U}$ is a suitable Hilbert space.

Since isometries can be well understood using analytic function theory, the main idea of the Sz.-Nagy–Foiaş theory is to represent an arbitrary contraction T (i.e., an operator with $\|T\| \leq 1$) on a Hilbert space H as a *compression* of an isometry V on a larger Hilbert space $K \supset H$. This means that $T = P_H V|H$, where P_H is the orthogonal projection of onto H. Hence, we can think of T as being a "piece" of V, and we can study the operator T by studying the larger operator V. Note that T must be a contraction if it is going to be a compression of some isometry. However, this is no restriction, because any operator can be assumed to be a contraction after multiplication by a suitable scalar. The dual concept to the compression is that of the *dilation*. Given an operator T on a Hilbert space H, we say that an operator V acting on a Hilbert space $K \supset H$ is its dilation if $T^n = P_H V^n|H$ for all $n \geq 0$. An equivalent definition of the dilation is to require that V has the structure

$$V = \begin{bmatrix} * & * & * \\ 0 & T & * \\ 0 & 0 & * \end{bmatrix}$$

with respect to some decomposition $K = H_1 \oplus H \oplus H_2$. This condition is stronger than that of T being a compression of V, because we have included all the powers of T and V. However, this is convenient to develop the theory.

The Sz.-Nagy–Foiaş theory shows that every contraction T can be dilated to a suitable isometry V. Then, it uses the spectral theory of isometries to produce a model for T using analytic functions. The model has a simpler form when $T^{*n} \to 0$ strongly as $n \to \infty$. It turns out that in this case T can be realized as a compression of the operator M_z of multiplication by z in $H^2(\mathscr{U})$ to a coinvariant subspace $H^2(\mathscr{U}) \ominus \Theta_T H^2(\mathscr{Y})$. Here, Θ_T is an inner function, which is a bounded analytic function on the disk whose values are linear operators taking $\mathscr{Y}$ into $\mathscr{U}$, and whose boundary values are isometric a.e. (this generalizes the scalar-valued inner functions in $H^\infty(\mathbb{D})$). The function Θ_T is called the *characteristic function* of T and contains all the information about the operator T. In the general case, the model space is more complicated, but it also involves the characteristic function Θ_T, which in this case is no longer inner.

Although it is customary to formulate this theory for contractions T, a parallel theory exists for maximal dissipative operators A. These are (possibly unbounded) operators such that $\operatorname{Im} A = (A - A^*)/2i \geq 0$. One can pass from dissipative operators to contractions by means of the Cayley transform $T = (A - iI)(A + iI)^{-1}$, which is constructed from the fractional transformation taking the upper half-plane onto the disk. The space $\mathscr{U}$ in the analytic model of the operator A will be finite-dimensional if and only if the condition $\operatorname{rank} \operatorname{Im} A < \infty$ holds.

Some of the main achievements of the Sz.-Nagy–Foiaş theory for contractions are the following:

- A functional calculus for a contraction T can be defined for a wide class H^∞_T of scalar-valued bounded analytic functions on the disk. For instance, in the case when T has no unitary part, H^∞_T is the whole $H^\infty(\mathbb{D})$ space.

- For contractions T such that there exists a non-zero $\varphi \in H^\infty_T$ such that $\varphi(T) = 0$, a minimal function $m_T \in H^\infty_T$ is defined. This minimal function plays a similar role to the minimal polynomial in linear algebra.

- There is a relation between the invariant subspaces of the contraction T and a special kind of factorizations of its characteristic function Θ_T, the so called regular factorizations.

Indeed, the Sz.-Nagy–Foiaş theory is one of the most informative spectral theories, apart from the spectral theory for normal operators.

Another important application of the Sz.-Nagy–Foiaş theory is the Commutant Lifting Theorem. Given a contraction T, an operator A commuting with T, and V an isometric dilation of T, this theorem constructs an operator B which is a lifting (a particular kind of dilation) of A and commutes with V. This allows us to represent A in the functional model of T and to completely describe the commutant of T.

The Commutant Lifting Theorem has also many applications to interpolation problems. For instance, it can be used to solve the Nevanlinna-Pick interpolation problem and the Caratheodory interpolation problem, among many others. The Caratheodory problem is related to geophysics, because it can be understood in terms of a seismic wave travelling through a layered medium. For a treatment of these topics, we refer to the monograph [FF90].

A natural generalization of this theory is its generalization to commutative tuples of contractions $(T_1, \dots, T_n)$. The second chapter of this work is devoted to this topic. Here, we try to find a dilation of the tuple to a commutative tuple of isometries $(V_1, \dots, V_n)$. For a pair of contractions, Andô's theorem shows that the dilation always exists. However, when $n \geq 3$, the dilation may or may not exist. There are several counterexamples in which the dilation does not exist, but the reasons behind the existence or non-existence of the dilation are not very well understood.

There is a relation between the existence of the dilation and the so called von Neumann's inequality. There are two versions of the inequality: the scalar-valued and the matrix-valued. The inequality for a single contraction T is

$$\|p(T)\| \leq \sup_{z \in \mathbb{D}} \|p(z)\|, \qquad \forall p.$$

Here, p is either a scalar-valued polynomial or a matrix-valued polynomial, depending on which version of the inequality we are considering. By a matrix-valued polynomial $p(z) = [p_{jk}(z)]_{jk}$ we mean a square matrix whose entries are scalar-valued polynomials p_{jk}. Then, $p(T)$ is the operator defined by the block operator matrix $[p_{jk}(T)]_{jk}$. Both versions of the von Neumann inequality hold for every contraction T, and this can be easily proved by using the Sz.-Nagy–Foiaş theory.

One can also consider the following generalization of this inequality for a tuple of commuting contractions $(T_1,\dots,T_n)$:

$$\|p(T_1,\dots,T_n)\| \leq \sup_{z_1,\dots,z_n\in\mathbb{D}} \|p(z_1,\dots,z_n)\|, \qquad \forall p,$$

where p is either a polynomial in n variables, or a matrix-valued polynomial in n-variables, depending on whether we want to consider the scalar-valued inequality or the matrix-valued inequality. However, this equality does not hold for every tuple of commuting contractions. It turns out that the existence of the dilation implies the scalar-valued version and that the matrix-valued version is equivalent to the existence of the dilation.

An important result of Agler [Agl90] gives a characterization of those polynomials p in n variables (either scalar valued or matrix valued) for which von Neumann's inequality holds for every tuple of commuting contractions $(T_1,\dots,T_d)$. We denote by $H^\infty(\mathbb{D}^n,G_1,G_2)$ the class of functions analytic and bounded on the polydisk $\mathbb{D}^n$, and which take values in $\mathcal{B}(G_1,G_2)$, where G_1, G_2 are Hilbert spaces (here, $\mathcal{B}(G_1,G_2)$ denotes the space of bounded linear operators mapping G_1 into G_2). By definition, a function $f\in H^\infty(\mathbb{D}^n,G_1,G_2)$ belongs to the Schur-Agler class $\mathcal{SA}_n(G_1,G_2)$ if there exist positive sesquianalytic $\mathcal{B}(G_2)$-valued kernels $K_j(z,w)$ on $\mathbb{D}^n$, for $j=1,\dots,n$, such that

$$I_{G_2} - f(z)f(w)^* = \sum_{j=1}^{n}(1-z_j\overline{w_j})K_j(z,w). \tag{$*$}$$

Here, a positive sesquianalytic $\mathcal{B}(G_2)$-valued kernel $K(z,w)$ is, by definition, a function which admits a factorization $K(z,w)=H(z)H(w)^*$, with H an analytic function on $\mathbb{D}^n$ which takes values on $\mathcal{B}(L,G_2)$, where L is an auxiliary Hilbert space.

A polynomial p in n variables whose values are $s\times s$ matrices belongs to the class $H^\infty(\mathbb{D}^n,\mathbb{C}^s,\mathbb{C}^s)$. The result of Agler is that the von Neumann inequality for a fixed p holds for every tuple of commuting contractions $(T_1,\dots,T_n)$ if and only if $p/\|p\|_\infty$ belongs to the Schur-Agler class $\mathcal{SA}_n(\mathbb{C}^s,\mathbb{C}^s)$.

Moreover, if $f\in H^\infty(\mathbb{D}^n,G_1,G_2)$, the operator $f(T_1,\dots,T_n)$ can be defined for a tuple of commuting strict contractions $(T_1,\dots,T_n)$ by using the power series of f. Therefore, it makes sense to ask when the von Neumann inequality for a fixed f holds for an arbitrary tuple of commuting strict contractions $(T_1,\dots,T_n)$. Once again, it holds if and only if $f/\|f\|_\infty$ belongs to the corresponding Schur-Agler class.

In the recent paper [GKVVW09], Grinshpan, Kaliuzhnyi-Verbovetskyi, Vinnikov and Woerdeman construct, for an arbitrary $f\in H^\infty(\mathbb{D}^n,G_1,G_2)$, a decomposition of $I_{G_2}-f(z)f(w)^*$ which is somewhat related to ($*$). This allows them to prove that if a tuple $(T_1,\dots,T_n)$ satisfies certain inequalities, then it satisfies the matrix-valued von Neumann's inequality. The inequalities required are similar to those involved in the theory of regular dilations (see (2.12) in Chapter 2).

A different approach to the theory of tuples of commuting operators is the Livšic-Vinnikov theory, which we briefly expose in the third chapter. The general setting of this theory is that of commuting non-selfadjoint operators. This is a tuple of commuting operators $(A_1,\dots,A_n)$ such that $\operatorname{rank}\operatorname{Im}A_j<\infty$, for $j=1,\dots,n$. The theory embeds this tuples into some structures, called colligations and vessels, which allow us to study the operators in terms of auxiliary matrices acting on a finite-dimensional space.

An algebraic variety in $\mathbb{C}^n$ called the *discriminant variety* is associated with each vessel. This gives a connection between operator theory and algebraic geometry. When the vessel consists of only two operators A_1, A_2, the discriminant variety is an algebraic curve in $\mathbb{C}^2$, and is called the *discriminant curve*. It is a real algebraic curve having a determinantal representation.

By means of the procedure of desingularization or blow-up, we can consider the discriminant curve as a compact Riemann surface. We say that a component of this Riemann surface is separated if when we remove from it the real points of the curve, we obtain a disconnected surface. It turns out that this disconnected surface has precisely two connected components, which we call its "halves". We say that the whole curve is separated if all of its components are separated. In this case, we can define one of its halves by picking out one of the two halves from each component (and the other half is the union of all the halves of the components that we did not pick).

In some cases, the discriminant curve of a vessel is separated. For instance, it follows from the work in [SV05] that this happens if a vessel is strict and $r(A_1, A_2)$ is defined and dissipative for some rational function r. When the discriminant curve is separated, its halves play an analogous role to the disk in the case of the theory of a single operator. Another important property of the discriminant curve is the generalized Cayley-Hamilton theorem. If Δ is the polynomial which defines the discriminant curve, then $\Delta(A_1, A_2) = 0$. The classical Cayley-Hamilton theorem can be obtained as a direct consequence of this.

The Livšic-Vinnikov theory is also related to the systems theory and control theory. Indeed, vessels have an interpretation as a control system, and many of the concepts of the theory have a more natural explanation in terms of this system. There is also some relation between vessels and quantum physics, because there seems to be an interesting interpretation of a quantum particle in terms of vessels. A treatment of this can be found in [LKMV95, Section 4.6] or in the paper [LA89].

One can think of the Sz.-Nagy–Foiaş theory as having two parts. The first one is the spectral theory of isometries, which have an analytic model consisting of a Hardy space on the disk. The second part is the relation between contractions and isometries: one can pass back and forth between them using compressions and dilations. This allows one to construct an analytic model for contractions using the Hardy space of the disk.

When passing from a single operator to a tuple of commuting operators, the key idea of the Livšic-Vinnikov theory is to use a complex algebraic curve (equivalently, a Riemann surface) instead of the complex plane. Vessels can be seen as the analogue of contractions, but there is no analogue for the spectral theory of isometries.

The fourth chapter of this work is an attempt to build this analogue. It is the original part of this work and it will be the starting point for the PhD thesis of the author. This part is a development of some previously unpublished ideas of Vinnikov and Yakubovich. We consider a construction called operator *pool*, which allows one to assign an algebraic curve in $\mathbb{C}^2$ to a pair of selfadjoint operators A_1, A_2 in a similar way to the Livšic-Vinnikov theory. The main construction of this chapter will be that of a *separating structure*. This is a pair of selfadjoint operators A_1, A_2 on some Hilbert space K, together with a decomposition $K = H_- \oplus H_+$ with the additional property that this decomposition "almost reduces" the operators A_1, A_2. This means that the operators $P_{H_-} A_j P_{H_+}$, $j = 1, 2$ have finite rank. There is a canonical way to produce a pool from a separating structure, so that we can assign an algebraic curve to it. In many case, this algebraic curve is separated. Its two halves can be used to model, in some sense, the two spaces H_- and H_+. Hence, we should be able to construct an analytic model of the separating structure by using Hardy spaces on its halves.

We also give the definition and first facts about a generalized notion of compression, which allows us to obtain a vessel by compressing a separating structure. We also hope to be able to define a dual notion of dilation of a vessel to a separating structure, so that we can pass back and forth between vessels and separating structures.

The theory of separating structures can be seen as a generalization of the theory of subnormal operators of finite type developed by Yakubovich in [Yak98a, Yak98b]. There, an analytic model for a subnormal operator is constructed using Hardy spaces on the halves of a separated algebraic curve. A subnormal operator S of finite type generates a separating structure, and the construction of the algebraic curve of this separating structure is equivalent to that done in [Yak98a, Yak98b] for the subnormal operator. However, separating structures are much more general. For instance, any linear combination $\alpha S + \beta S^*$ also generates a separating structure. Therefore, many new phenomena appear when considering separting structures.

One of the motivations behind the theory of separating structures is to try to shed some light on the existence and non-existence of the dilation of a tuple of operators. A future goal would be to obtain an analogue of Andô's theorem in the context of this theory and to use the theory to obtain more information about the dilation.

Let us mention some recent papers related to this work. Agler, Knese and McCarthy consider in [AKM12] algebraic pairs of isometries (V_1, V_2), which are those satisfying $q(V_1, V_2) = 0$ for some polynomial q. They show that there is a minimal polynomial q satifying this relation and that this minimal q is inner-toral. This means that its zero set Z_q, which is an algebraic curve, is contained in $\mathbb{D}^2 \cup \mathbb{T}^2 \cup (\mathbb{C} \setminus \mathbb{D})^2$ (such an algebraic curve is called a distinguished variety, because it exits the bidisk $\mathbb{D}^2$ only at the distinguished boundary $\mathbb{T}^2 \subset \partial\mathbb{D}^2$). They prove that under certain conditions, the pair (V_1, V_2) can be realized as multiplication by the coordinates functions on some Hardy space on the distinguished variety Z_q. Jury, Knese and McCullough give in [JKM12] an analogue of the Nevanlinna-Pick interpolation theorem in distinguished varieties. Then, they use this theorem to prove a result about dilation of algebraic pairs of isometries modelled on the distinguished variety.

Now we will give a more detailed summary of the contents of this work by chapters. In Chapter 1, we make an introductory exposition of the Sz.-Nagy–Foiaş theory. We define the vector valued Hardy spaces and give their main properties. We prove the results about isometric and unitary dilations of a contraction, giving several different but equally interesting proofs. Then, we pass to the construction of the Sz.-Nagy–Foiaş model for the case of $C_{\cdot 0}$ contractions, which has a simpler form. The model for the case of a general contraction is also given without proof. Finally, we show an application of the model to the study of invariant subspaces, and give a relation between factorizations of the characteristic function and invariant subspaces.

Chapter 2 is devoted to results concerning the simultaneous dilation of several operators. We start this chapter with a proof of the Commutant Lifting Theorem. This theorem is usually considered to be part of the Sz.-Nagy–Foias theory, but since it concerns the dilation of two operators and it can be used to give a short proof of Andô's Theorem, we have decided to give it in this chapter. We also show an application of the Commutant Lifting Theorem to the Nevanlinna-Pick interpolation problem. Then, we turn to Andô's Theorem, which can be considered to be the main result in the theory of dilations of several operators. We show how it can be obtained from the Commutant Lifting Theorem easily and, conversely, how the Commutant Lifting Theorem also follows from Andô's Theorem. Then we introduce von Neumann's inequality, and its relation with the existence of the dilation. We also make a brief exposition of the theory of regular dilations. Next, we give the main examples of non existence of the dilation from the literature. Finally, we give a positive existence result by Lotto.

In Chapter 3, we give a brief exposition of the Livšic-Vinnikov theory of commuting nonselfadjoint operators. We start by introducing the concept of a colligation of a single operator, showing some of the main tools of the colligation: the projection and coupling, the system theoretical interpretation and the characteristic function. Then, we give the notion of a colligation of several operators, showing that many of the tools which were available for a single operator do not have a good generalization to colligations of several operators. This motivates the definition of a vessel, which is a colligation of several operators with some extra extructure that allows one to generalize these tools. We also introduce the discriminant curve of a vessel and the restoration formula.

Chapter 4 is the original part of this work and concerns the theory of separating structures. We start with the theory of pools, showing how their discriminat curve can be defined. Then we pass to the definition and main results of separating structures. We first treat the affine case separately to show which properties depend just on the linear algebra. We define the mosaic function, which is an analytic projection-valued function that, in a certain sense, models the projection onto the factor H_+ of the decomposition $K = H_- \dotplus H_+$. Then we pass to the orthogonal case, proving that every separating structure defines a pool. Then we show that, under some mild conditions on the structure, the associated algebraic curve is separated, and we prove a restoration formula for the mosaic function of the structure in terms of its algebraic curve. Many of these constructions generalize those done in [Yak98a, Yak98b] for subnormal operators. Therefore, we will also give examples of our constructions using subnormal operators and show how our results relate to those in [Yak98a, Yak98b]. Finally, we give the definition of the compression of a separating structure and show some of its properties. In particular, we prove that the compression of a separating structure is a vessel.

Notation and preliminaries

Complex analysis

The open unit disk in the complex plane $\mathbb{C}$ will be denoted by $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$. The one dimensional torus will be denoted by $\mathbb{T} = \partial\mathbb{D} = \{z \in \mathbb{C} : |z| = 1\}$.

Hilbert spaces

The scalar product in a Hilbert space will be denoted by $\langle \cdot, \cdot \rangle$. By a subspace of a Hilbert space, we will always mean a linear submanifold which is closed (so that the subspace is also a Hilbert space). If $H_1 \subset H_2$ are subspaces of some Hilbert space H, we define $H_2 \ominus H_1 = H_2 \cap H_1^{\perp}$, where $H_1^{\perp}$ is the orthogonal complement of H_1 in H.

If A_j are subsets of a Hilbert space H, we will denote by $\bigvee_j A_j$ the smallest subspace of H which contains every set A_j.

The orthogonal direct sum of Hilbert spaces H_n, $n \geq 0$ will be denoted by $\bigoplus_{n\geq 0} H_n$. An element in the direct sum will be written as $(h_0, h_1, \ldots)$.

If M_1, M_2 are subspaces of a Hilbert space H, we will usually write $M_1 + M_2$ for the smallest (not necessarily closed) linear submanifold of H which contains both. If $M_1 \cap M_2 = 0$, we can consider the parallel projections $P_j : M_1 + M_2 \to M_j$. For instance, by definition, P_1 has range M_1 and kernel M_2. If these parallel projections are bounded (equivalently, if $M_1 + M_2$ is closed), we say that M_1 and M_2 are in direct sum and write $M_1 \dotplus M_2$. If at least one of the two subspaces M_j is finite-dimensional, then $M_1 + M_2$ is always closed. This notation applies also to sums of more than two summands.

Operators

By an operator between Hilbert spaces, we will always mean a linear and bounded operator. The set of operators between Hilbert spaces H and K will be denoted by $\mathcal{B}(H, K)$. We will write $\mathcal{B}(H)$ instead of $\mathcal{B}(H, H)$.

For an operator $T \in \mathcal{B}(H, K)$ we will define its adjoint operator $T^* \in \mathcal{B}(K, H)$ by the condition $\langle Th, k \rangle = \langle h, T^*k \rangle$ for all $h \in H$ and $k \in K$.

If $T \in \mathcal{B}(H)$ and L is a subspace of H, we will say that L is invariant for T if $TL \subset L$. This is equivalent to $P_L T P_L = T P_L$, where P_L is the orthogonal projection onto L. It is easy to see that L is invariant for T if and only if $H \ominus L$ is invariant for the adjoint operator T^*. We will say that the subspace L reduces T if and only if L is invariant for both T and T^*. This is equivalent to $P_L T = T P_L$.

An operator $A \in \mathcal{B}(H)$ will be called selfadjoint if $A = A^*$. If $A, B \in \mathcal{B}(H)$ are selfadjoint operators, we will write $A \leq B$ if $\langle Ah, h \rangle \leq \langle Bh, h \rangle$ for all $h \in H$. If $A \geq 0$, there is a unique operator $B \geq 0$ such that $A = B^2$. This operator B will be denoted by $A^{\frac{1}{2}}$.

An operator $V \in \mathcal{B}(H, K)$ will be called an isometry if $\|Vh\| = \|h\|$ for all $h \in H$. This is equivalent to the condition $V^*V = I$. An operator $U \in \mathcal{B}(H, K)$ will be called a unitary if $U^* = U^{-1}$. It is easy to see that an isometry is a unitary if and only if it is onto.

Two operators $T \in \mathcal{B}(H, K)$ and $T' \in \mathcal{B}(H', K')$ are called unitarily equivalent if there exist unitaries $U_1 \in \mathcal{B}(H, H')$ and $U_2 \in \mathcal{B}(K, K')$ such that $T'U_1 = U_2 T$. If $K = H$ and $K' = H'$, it is usual to require that $U_2 = U_1$.

An operator $T \in \mathcal{B}(H, K)$ will be called a contraction if $\|T\| \leq 1$. This is equivalent to the condition $T^*T \leq I$. If T is a contraction, one can define its defect operator $D_T = (I - T^*T)^{\frac{1}{2}}$ and its defect subspace $\mathfrak{D}_T = \overline{D_T H}$. Note that if T is a contraction, then T^* is also a contraction.

A sequence of operators $\{T_n\} \subset \mathcal{B}(H, K)$ will be said to converge uniformly (or in the operator norm) to $T \in \mathcal{B}(H, K)$ if $\|T_n - T\| \to 0$. It will be said to converge strongly to T if $\|T_n h - Th\| \to 0$ for all $h \in H$, and it will be said to converge weakly to T if $\langle T_n h, k\rangle \to \langle Th, k\rangle$ for all $h \in H$ and $k \in K$.

Miscellanea

The end of a proof will be marked with the usual symbol $\square$. Some examples will be given in Chapter 4. Then end of those examples will be marked with the symbol $\spadesuit$.

1. Dilations of a single operator: Sz.-Nagy–Foiaş theory

This chapter is devoted to a brief exposition of the Sz.-Nagy–Foiaş theory of contractions. First, we will define the concepts of extension, lifting and dilation, which will be important throughout all this work, especially the first two chapters. Then, we give a very brief introduction to vector valued Hardy spaces, which will be needed later. We give some general facts about unilateral and bilateral shifts, and their relation with Hardy spaces. Then we show how every contraction can be dilated to an isometry or a unitary. This allows us to construct the Sz.-Nagy–Foiaş functional model of a contraction. We will do this only for $C_{\cdot 0}$ contractions, which is a somewhat simpler case. However, the general statement of the model is also given at the end of Section 1.5. Finally, we show an application of the model to the study of invariant subspaces.

The principal exposition of this theory can be found in the book by Sz.-Nagy and Foiaş [SNFBK10]. Alternative expositions from slightly different points of view appear in the book by Foiaş and Frazho [FF90] and in the books by Nikolski [Nik86, Nik02b].

1.1. Extensions, liftings and dilations

Let H, H' be Hilbert spaces, $A \in \mathcal{B}(H, H')$ and K, K', larger Hilbert spaces $K \supset H$, $K' \supset H'$. The operator $B \in \mathcal{B}(K, K')$ is said to be an *extension* of A if $A = B|H$. In this case, A is called a *restriction* of B. It is easy to see that this is equivalent to the condition that B has the matrix representation

$$B = \begin{array}{c c} & \begin{array}{cc} H & K\ominus H \end{array} \\ \begin{array}{c} H' \\ K'\ominus H' \end{array} & \begin{bmatrix} A & * \\ 0 & * \end{bmatrix} \end{array},$$

with respect to the decompositions $K = H \oplus (K \ominus H)$ and $K' = H' \oplus (K' \ominus H')$.

We say that B is a *lifting* of A if $P_{H'}B = AP_H$, where P_H and P'_H denote the orthogonal projections onto H and H' respectively. It is easy to check that B is a lifting of A if and only if the adjoint operator B^* is an extension of A^*, and also if and only if

$$B = \begin{array}{c c} & \begin{array}{cc} H & K\ominus H \end{array} \\ \begin{array}{c} H' \\ K'\ominus H' \end{array} & \begin{bmatrix} A & 0 \\ * & * \end{bmatrix} \end{array}.$$

The operator B is called a *weak dilation* of A if $A = P_{H'}B|H$. In this case, A is called a *compression* of B. This notion corresponds to the matrix representation

$$B = \begin{array}{c c} & \begin{array}{cc} H & K\ominus H \end{array} \\ \begin{array}{c} H' \\ K'\ominus H' \end{array} & \begin{bmatrix} A & * \\ * & * \end{bmatrix} \end{array}.$$

Assume now that $H = H'$ and $K = K'$. We say that B is a *dilation* of A if B^n is a weak dilation of A^n for each $n \geq 0$. It is clear that if B is an extension of A, then B is also a dilation of A.

The following Lemma, due to Sarason, gives the matrix representation of a dilation. We will formulate it in a more general context, because it will be useful later in this form. Recall that an algebra $\mathcal{A}$ is called unital if it has a unit and that a homomorphism between two unital algebras is called unital if it takes the unit of the first algebra to the unit of the second one.

Lemma 1.1 (Sarason)**.** *Let $\mathcal{A}$ be a complex unital algebra, $\varphi : \mathcal{A} \to \mathcal{B}(K)$ a unital homomorphism, and H a subspace of K. The map $\widetilde{\varphi} : \mathcal{A} \to \mathcal{B}(H)$ defined by $\widetilde{\varphi}(a) = P_H\varphi(a)|H$ is a homomorphism if and only if $H = H_1 \ominus H_2$, where H_1 and H_2 are subspaces of K which are $\varphi(a)$-invariant for every $a \in \mathcal{A}$.*

Proof. First assume that $\widetilde{\varphi}$ is a homomorphism. We define

$$H_1 = \bigvee_{a\in\mathcal{A}} \varphi(a)H.$$

Note that $H \subset H_1$ because $\mathcal{A}$ is unital. Then we define $H_2 = H_1 \ominus H$, so that $H = H_1 \ominus H_2$. It is clear that H_1 is invariant for every $\varphi(a)$, because φ is a homomorphism.

To show that H_2 is invariant for $\varphi(a)$, we must check that

$$(P_{H_1} - P_H)\varphi(a)(P_{H_1} - P_H) = \varphi(a)(P_{H_1} - P_H). \tag{1.1}$$

Since H_1 is invariant for $\varphi(a)$, we have $P_{H_1}\varphi(a)P_{H_1} = \varphi(a)P_{H_1}$. Also, $P_{H_1}\varphi(a)P_H = \varphi(a)P_H$, by definition of H_1. Hence, (1.1) is equivalent to

$$P_H\varphi(a)P_{H_1} = P_H\varphi(a)P_H. \tag{1.2}$$

It suffices to check equation (1.2) applied to vectors of the form $\varphi(b)h$, for $b \in \mathcal{A}$ and $h \in H$, because the family of those vectors spans H_1. Therefore, we need to check that

$$P_H\varphi(a)\varphi(b)|H = P_H\varphi(a)P_H\varphi(b)|H,$$

for every $b \in \mathcal{A}$. Using the fact that φ is a homomorphism, this equation rewrites as

$$\widetilde{\varphi}(ab) = \widetilde{\varphi}(a)\widetilde{\varphi}(b), \tag{1.3}$$

which is true precisely because $\widetilde{\varphi}$ is a homomorphism.

To prove the converse, note that the hypothesis imply (1.1), and that from this we can obtain (1.3) in roughly the same way as before (note that $\varphi(b)H \subset H_1$ for every $b \in \mathcal{A}$). This implies that $\widetilde{\varphi}$ is multiplicative. Since $\widetilde{\varphi}$ is also linear, it is a homomorphism. □

Applying the Lemma to the case when $\mathcal{A} = \mathbb{C}[z]$ is the algebra of complex polynomials and $\varphi(p) = p(B)$ (see Appendix A for the definition of the polynomial functional calculus), we see that B is a dilation of A if and only if

$$B = \begin{bmatrix} * & * & * \\ 0 & A & * \\ 0 & 0 & * \end{bmatrix},$$

with respect to the decomposition $K = H_2 \oplus H \oplus (K \ominus H_1)$. This shows that if B is a lifting of A, then B is also a dilation of A (just put $H_1 = K$).

1.2. Vector valued Hardy spaces

The theory of Hardy spaces of functions with values in a Hilbert space will be needed later. In this section, which has an auxiliary character, we will give the main definitions and results of this theory. A more detailed exposition can be found in [SNFBK10, Chapter V].

Let $\mathscr{U}$ be a separable Hilbert space. If X is a measure space, a function $f : X \to \mathscr{U}$ will be called measurable if each of the scalar-valued functions $\langle f(\cdot), u\rangle$, $u \in \mathscr{U}$, is measurable. This notion is usually called weak measurability. There exists another notion of measurability called Bochner measurability or strong measurability. In the case when $\mathscr{U}$ is separable, this two notions coincide, by the Pettis Theorem. A good treatment of this topic can be found in [ABHN11, Section 1.1].

The space $L^2(\mathscr{U})$ is defined as the space of $\mathscr{U}$-valued measurable functions f on $\mathbb{T}$ such that

$$\|f\|^2_{L^2(\mathscr{U})} = \frac{1}{2\pi}\int_0^{2\pi} \|f(e^{it})\|^2_{\mathscr{U}}\, dt < \infty.$$

With this norm, $L^2(\mathscr{U})$ becomes a Hilbert space.

Any function $f \in L^2(\mathscr{U})$ can be written in Fourier series as

$$f(e^{it}) = \sum_{n\in\mathbb{Z}} e^{int} c_n, \qquad c_n \in \mathscr{U},$$

where this series is convergent in the sense of the $L^2(\mathscr{U})$-norm. The Fourier coefficients are determined by

$$\langle c_n, u\rangle_{\mathscr{U}} = \frac{1}{2\pi}\int_0^{2\pi} e^{-int}\langle f(e^{it}), u\rangle_{\mathscr{U}}\, dt = \langle f, e^{int}u\rangle_{L^2(\mathscr{U})}, \qquad u \in \mathscr{U},$$

and the Parseval identity holds:

$$\|f\|^2_{L^2(\mathscr{U})} = \sum_{n\in\mathbb{Z}} \|c_n\|^2_{\mathscr{U}}.$$

The operator $M_{e^{it}}$ of multiplication by e^{it} on $L^2(\mathscr{U})$ will be important in the sequel. It is defined by $(M_{e^{it}}f)(e^{it}) = e^{it}f(e^{it})$.

If $\Omega \subset \mathbb{C}$ is an open set and X is a Banach space, a function $g : \Omega \to X$ will be called analytic if each of the scalar-valued functions $\varphi \circ g$, $\varphi \in X^*$, is analytic (i.e., holomorphic). This notion is usually called weak analyticity, and there exists also a notion of strong analyticity. The two notions are equivalent when X is a Banach space. See, for instance, [Rud91, Chapter 3] for a discussion of these notions in the context of topological vector spaces.

The space $H^2(\mathscr{U})$ is defined as the space of analytic $\mathscr{U}$-valued functions g on $\mathbb{D}$ such that

$$\|g\|^2_{H^2(\mathscr{U})} = \sup_{0<r<1} \frac{1}{2\pi}\int_0^{2\pi} \|g(re^{it})\|^2_{\mathscr{U}}\, dt < \infty. \tag{1.4}$$

With this norm, $H^2(\mathscr{U})$ is a Hilbert space.

Any function $g \in H^2(\mathscr{U})$ can be written in power series as

$$g(z) = \sum_{n\geq 0} z^n a_n, \qquad a_n \in \mathscr{U}. \tag{1.5}$$

This power series converges in the $\mathscr{U}$-norm uniformly on compact subsets of $\mathbb{D}$. Direct computation using (1.4) shows that

$$\|g\|^2_{H^2(\mathscr{U})} = \sum_{n\geq 0} \|a_n\|^2_{\mathscr{U}}.$$

If $g \in H^2(\mathscr{U})$ has the power series expansion (1.5), we can also consider the function $f \in L^2(\mathscr{U})$ given by

$$f(e^{it}) = \sum_{n\geq 0} e^{int} a_n.$$

Then $\|f\|_{L^2(\mathscr{U})} = \|g\|_{H^2(\mathscr{U})}$. Defining $g_r(e^{it}) = g(re^{it})$, for $0 < r < 1$, one can easily check that $g_r \in L^2(\mathscr{U})$ and that $g_r \to f$ in the $L^2(\mathscr{U})$-norm as $r \to 1^-$. This function f is called the *boundary value function* of g, and g can be recovered from f by the Poisson formula

$$g(re^{it}) = \frac{1}{2\pi} \int_0^{2\pi} P_r(t-s) f(e^{is})\, ds, \tag{1.6}$$

where $P_r(t)$ is Poisson's kernel

$$P_r(t) = \frac{1-r^2}{1-2r\cos t + r^2}.$$

Equation (1.6) can be checked directly for $g(z) = z^n$, $n \geq 0$, and then for an arbitrary $g \in H^2(\mathscr{U})$ by using linearity and taking limits. Moreover, by using the properties of Poisson's kernel, one can prove that f is also a radial limit of g almost everywhere:

$$f(e^{it}) = \lim_{r\to 1^-} g(re^{it}), \qquad \text{a.e. } e^{it} \in \mathbb{T},$$

where the limit is taken in the $\mathscr{U}$-norm. The proof of this fact is the same as for the scalar-valued $H^2(\mathbb{D})$. This is the so called Fatou's Theorem, and can be found, for instance, in [MAR07, Theorem 1.1.26] and [Dur70, Theorem 1.2].

For each function $f \in L^2(\mathscr{U})$ whose negative Fourier coefficients vanish, it is easy to see that (1.6) defines a function $g \in H^2(\mathscr{U})$. Hence, such a function arises as the boundary value function of some $g \in H^2(\mathscr{U})$. It is customary to identify a function in $H^2(\mathscr{U})$ with its boundary value function in $L^2(\mathscr{U})$ and to identify $H^2(\mathscr{U})$ with the subspace of $L^2(\mathscr{U})$ of functions with vanishing negative Fourier coefficients. We will do so in the sequel without explicit mention.

We define the operator M_z of multiplication by z in $H^2(\mathscr{U})$ by $(M_z g)(z) = zg(z)$. Note that the operator $M_{e^{it}}$ is an extension of M_z.

If $\mathscr{U}$ and $\mathscr{Y}$ are separable Hilbert spaces, the space $H^\infty(\mathscr{U},\mathscr{Y})$ is defined as the space of analytic functions Θ in $\mathbb{D}$ which take values in $\mathcal{B}(\mathscr{U},\mathscr{Y})$, and such that

$$\|\Theta\|_{H^\infty(\mathscr{U},\mathscr{Y})} = \sup_{z\in\mathbb{D}} \|\Theta(z)\| < \infty.$$

Such a function $\Theta \in H^\infty(\mathscr{U},\mathscr{Y})$ is called a *bounded analytic function*.

If $\Theta \in H^\infty(\mathscr{U},\mathscr{Y})$, it has a power series representation

$$\Theta(z) = \sum_{n\geq 0} z^n \Theta_n, \qquad \Theta_n \in \mathcal{B}(\mathscr{U},\mathscr{Y}).$$

This series converges in the $\mathcal{B}(\mathscr{U},\mathscr{Y})$-norm uniformly on compact subsets of $\mathbb{D}$.

A bounded analytic function Θ defines an operator of multiplication by Θ, which we will also denote by Θ. The operator $\Theta : H^2(\mathscr{U}) \to H^2(\mathscr{Y})$ is defined $(\Theta g)(z) = \Theta(z)g(z)$, for $g \in H^2(\mathscr{U})$. The power series of Θg can be computed by

$$(\Theta g)(z) = \sum_{\substack{0\le n\\ 0\le m\le n}} z^n \Theta_m a_{n-m},$$

where (1.5) is the power series of g.

Moreover, one can also prove that the strong limit of $\Theta(re^{it})$ as $t \to 1^-$ exists a.e. $e^{it} \in \mathbb{T}$. Hence, we can define a.e. a boundary value function, which we will also denote by Θ, by

$$\Theta(e^{it}) = \lim_{r\to 1^-} \Theta(re^{it}).$$

This allows us to define also the operator of multiplication by Θ in $L^2(\mathscr{U})$, $\Theta : L^2(\mathscr{U}) \to L^2(\mathscr{Y})$, by $(\Theta f)(e^{it}) = \Theta(e^{it})f(e^{it})$. Note that when $H^2(\mathscr{U})$ is identified with a subspace of $L^2(\mathscr{U})$, this operator is an extension of the operator Θ in $H^2(\mathscr{U})$ previously defined.

A function $\Theta \in H^\infty(\mathscr{U},\mathscr{Y})$ will be called a *contractive analytic function* if $\Theta(z)$ is a contraction for all $z \in \mathbb{D}$. A contractive analytic function $\Theta(z)$ is called *inner* if $\Theta(e^{it})$ is an isometry a.e. on $\mathbb{T}$.

The following Lemma characterizes which operators can arise as multiplication by a contractive analytic function. It will be useful later. Note that the converse of this Lemma is obvious.

Lemma 1.2. *Let $Q : H^2(\mathscr{U}) \to H^2(\mathscr{Y})$ be a contraction such that $QM_z = M_zQ$. Then there exists a contractive analytic function $\Theta \in H^\infty(\mathscr{U},\mathscr{Y})$ such that $Q = \Theta$ (in the sense that Q is the operator of multiplication by Θ). Moreover, if Q is an isometry, then Θ is inner.*

Proof. Fix an arbitary $u \in \mathscr{U}$. We will also denote by u the function in $H^2(\mathscr{U})$ which is constantly u. If $p(z)$ is a polynomial we have

$$\|p(z)Qu\|^2_{H^2(\mathscr{Y})} = \|Qp(z)u\|^2_{H^2(\mathscr{Y})} \le \|p(z)u\|^2_{H^2(\mathscr{U})},$$

which implies

$$\frac{1}{2\pi}\int_0^{2\pi} |p(e^{it})|^2 \|(Qu)(e^{it})\|^2_{\mathscr{Y}}\, dt \le \frac{1}{2\pi}\int_0^{2\pi} |p(e^{it})|^2 \|u\|^2_{\mathscr{U}}\, dt.$$

It is clear that this inequality must be also true when p is a trigonometric polynomial. Hence, by the Weierstrass approximation theorem,

$$\frac{1}{2\pi}\int_0^{2\pi} \varphi(e^{it}) \|(Qu)(e^{it})\|^2_{\mathscr{Y}}\, dt \le \frac{1}{2\pi}\int_0^{2\pi} \varphi(e^{it}) \|u\|^2_{\mathscr{U}}\, dt,$$

for every non-negative continuous function φ in $\mathbb{T}$. This implies that

$$\|(Qu)(e^{it})\|^2_{H^2(\mathscr{Y})} \le \|u\|^2_{\mathscr{U}}, \qquad \text{a.e. } t \in [0, 2\pi].$$

By the Poisson formula (1.6), since $\int_0^{2\pi} P_r(t)\, dt = 2\pi$, we get

$$\|(Qu)(z)\|^2_{\mathscr{Y}} \le \|u\|^2_{\mathscr{U}}, \qquad z \in \mathbb{D}. \tag{1.7}$$

Now it is clear that

$$\Theta(z)u = (Qu)(z), \qquad u \in \mathscr{U}, \tag{1.8}$$

defines a contractive analytic function.

Finally, assume that Q is an isometry. Then all the inequalities in the reasoning above up to (1.7) are indeed equalities. Letting $z \to e^{it} \in \mathbb{T}$ in (1.8), we get $\Theta(e^{it})u = (Qu)(e^{it})$ a.e. on $\mathbb{T}$. Since equality holds in (1.7), we see that $\Theta(e^{it})$ is an isometry a.e. on $\mathbb{T}$. Hence, Θ is inner. □

The next Lemma gives alternative characterizations of contractive analytic functions and inner functions which are quite useful.

Lemma 1.3. *The following conditions are equivalent for a bounded analytic function $\Theta \in H^\infty(\mathscr{U}, \mathscr{Y})$:*

(i) *Θ is a contractive analytic function.*

(ii) *The operator $\Theta : H^2(\mathscr{U}) \to H^2(\mathscr{Y})$ is a contraction.*

(iii) *The operator $\Theta : L^2(\mathscr{U}) \to L^2(\mathscr{Y})$ is a contraction.*

The following conditions are also equivalent:

(a) *Θ is an inner function.*

(b) *The operator $\Theta : H^2(\mathscr{U}) \to H^2(\mathscr{Y})$ is an isometry.*

(c) *The operator $\Theta : L^2(\mathscr{U}) \to L^2(\mathscr{Y})$ is an isometry.*

Proof. The implication (iii) $\Rightarrow$ (ii) is obvious. To prove (ii) $\Rightarrow$ (i), note that if we let Q be the operator of multiplication by Θ in Lemma 1.2, then we get a contractive analytic function Θ' such that the operators of multiplication by Θ and Θ' are equal. Then, Θ and Θ' must be the same contractive analytic function. Finally, (i) $\Rightarrow$ (iii) is trivial, because $\Theta(e^{it})$, being a strong limit of contractions a.e. on $\mathbb{T}$, must be a contraction a.e. on $\mathbb{T}$.

To prove the second part, note that (a) $\Rightarrow$ (c) and (c) $\Rightarrow$ (b) are trivial and (b) $\Rightarrow$ (a) follows from Lemma 1.2 as in the proof of (ii) $\Rightarrow$ (i). $\square$

A particular kind of contractive analytic functions are those which are constant, i.e., those contractive analytic functions $\Theta \in H^\infty(\mathscr{U}, \mathscr{Y})$ such that $\Theta(z) = \Theta_0 \in \mathcal{B}(\mathscr{U}, \mathscr{Y})$ for all $z \in \mathbb{D}$. If the constant value of such a Θ is an isometry (a unitary), then Θ is called a constant isometry (a constant unitary). The following Lemma gives a characterization of constant isometries using the maximum modulus principle.

Lemma 1.4. *If $\Theta \in H^\infty(\mathscr{U}, \mathscr{Y})$ is a contractive analytic function such that $\Theta(0)$ is an isometry, then Θ is a constant isometry.*

Proof. Fix $u \in \mathscr{U}$ with $\|u\|_{\mathscr{U}} = 1$ and put $\varphi(z) = \langle \Theta(z)u, \Theta(0)u \rangle_{\mathscr{Y}}$. Then $\varphi(z)$ is scalar-valued and holomorphic in $\mathbb{D}$, and $|\varphi(z)| \leq 1$ for all $z \in \mathbb{D}$, because $\Theta(z)$ is a contraction. Moreover, $\varphi(0) = 1$, because $\Theta(0)$ is an isometry. By the maximum modulus principle, it follows that $\varphi(z) = 1$ for all $z \in \mathbb{D}$. Since $\Theta(z)$ is a contraction, this implies $\Theta(z)u = \Theta(0)u$. Therefore, Θ is constant, because u was arbitrary. $\square$

Note that in the hypothesis of this Lemma, if $\Theta(0)$ is in fact unitary, then it follows that Θ is a constant unitary.

1.3. Unilateral and bilateral shifts

In this section we will give the definition and basic facts of unilateral and bilateral shifts. These operators will be important in the sequel and are also good examples of isometries and unitaries, respectively.

If V is an isometry on H, a subspace $L \subset H$ is called a *wandering space* for V if $V^n L \perp V^m L$ for $n \neq m$, $n, m \geq 0$. Since V is an isometry, it is easy to see that this is equivalent to the condition $V^n L \perp L$ for $n \geq 1$. Indeed, if $n \geq m \geq 0$, since we have $V^{n-m} L \perp L$, and application of V preserves orthogonality, applying V^m to both subspaces we get $V^n L \perp V^m L$. Given a wandering space L, one can form the subspace

$$M_+(L) = \bigoplus_{n \geq 0} V^n L.$$

An isometry $V \in \mathcal{B}(H)$ is called a unilateral shift if there exists a wandering subspace L such that $H = M_+(L)$. In this case, L is uniquely determined by V. In fact, $L = H \ominus VH$. To see this, just note that if $H = M_+(V)$, then

$$H \ominus VH = \Big(\bigoplus_{n \geq 0} V^n L \Big) \ominus V \Big(\bigoplus_{n \geq 0} V^n L \Big) = \Big(\bigoplus_{n \geq 0} V^n L \Big) \ominus \Big(\bigoplus_{n \geq 1} V^n L \Big) = L.$$

The dimension of L is called the multiplicity of the shift, and a unilateral shift is uniquely determined up to unitary equivalence by its multiplicity (it is very easy to construct a unitary between two shifts of the same multiplicity).

An example of a unilateral shift is the operator M_z in $H^2(\mathscr{U})$, where $\mathscr{U}$ is a Hilbert space. Indeed, this operator is the canonical example of a unilateral shift. If $V \in \mathcal{B}(H)$ is a unilateral shift with wandering space L and $W : \mathscr{U} \to L$ is a unitary, then we can define a unitary operator $Z_+ : H \to H^2(\mathscr{U})$ by

$$Z_+ \sum_{n \geq 0} V^n l_n = \sum_{n \geq 0} z^n W^* l_n, \qquad l_n \in L. \tag{1.9}$$

Then Z_+ transforms the unilateral shift V into the operator M_z on $H^2(\mathscr{U})$, in the sense that $Z_+ V = M_z Z_+$.

If U is a unitary and L is wandering for U, then note that it follows that $U^n L \perp U^m L$ for $n \neq m$, $n, m \in \mathbb{Z}$. Hence, one can define

$$M(L) = \bigoplus_{n \in \mathbb{Z}} U^n L.$$

The operator $U \in \mathcal{B}(K)$ is called a bilateral shift if $M(L) = K$ for some wandering space L. In this case L is not uniquely determined by U. Note that if L is wandering, then $U^n L$ is also wandering for every $n \in \mathbb{Z}$ and $M(L) = M(U^n L)$. However, one can check that the dimension of any two possible such subspaces L is the same. This dimension is called the multiplicity of the shift, and two bilateral shifts of the same multiplicity are unitarily equivalent.

From this, it is clear how a unilateral shift $V \in \mathcal{B}(H)$ can be extended to a bilateral shift of the same multiplicity $U \in \mathcal{B}(K)$. We take L the wandering subspace of V, put $K = \bigoplus_{n \in \mathbb{Z}} L$ and define U by $U(\dots, l_{-1}, \check{l_0}, l_1, \dots) = (\dots, l_{-2}, \widecheck{l_{-1}}, l_0, \dots)$ (here, the symbol $\check{}$ marks the vector in the 0-th position). Then we embed H into K by $\sum_{n=0}^{\infty} V^n l_n \mapsto (\dots, 0, 0, \check{l_0}, l_1, \dots)$. It is easy to see that the embedding is well defined and U is a bilateral shift and a extension of V.

An example of a bilateral shift is the operator $M_{e^{it}}$ in $L^2(\mathscr{U})$, where $\mathscr{U}$ is a Hilbert space. As for the case of unilateral shifts, this operator is the canonical example of a bilateral shift. If $U \in \mathcal{B}(K)$ is a bilateral shift, L is a wandering subspace of U, and $W : L \to \mathscr{U}$ is a unitary, then the unitary $Z : K \to L^2(\mathscr{U})$ defined by

$$Z \sum_{n \in \mathbb{Z}} U^n l_n = \sum_{n \in \mathbb{Z}} e^{int} W^* l_n, \qquad l_n \in L \tag{1.10}$$

transforms the bilateral shift U into the operator $M_{e^{it}}$ on $L^2(\mathscr{U})$, in the sense that $ZV = M_{e^{it}}Z$. Note that if we put $V = U|M_+(L)$, then V is a unilateral shift and the restriction $Z|M_+(L) : M_+(L) \to H^2(\mathscr{U})$ is a unitary which takes V into M_z. This unitary is precisely the operator Z_+ defined in (1.9)

If we view a unilateral shift V as the operator M_z in some $H^2(\mathscr{U})$, then the procedure of extension of V to a bilateral shift U describe above corresponds to extending M_z by the operator $M_{e^{it}}$ in $L^2(\mathscr{U})$.

1.4. Isometric and unitary dilations

This section deals with the existence and the construction of isometric and unitary dilations of a contraction. Recall that if $T \in \mathcal{B}(H)$ is a contraction, its defect operator is $D_T = (I - T^*T)^{\frac{1}{2}}$, and its defect subspace is $\mathfrak{D}_T = \overline{D_T H}$. This will play an important role in the sequel.

We will start by proving a theorem of Sz.-Nagy which shows that every contraction can be lifted to an isometry.

Theorem 1.5 (Sz.-Nagy)**.** *Any contraction $T \in \mathcal{B}(H)$ can be lifted to an isometry $V \in \mathcal{B}(K_+)$. This lifting can be chosen to be minimal, in the sense that*

$$K_+ = \bigvee_{n\geq 0} V^n H. \tag{1.11}$$

Moreover, any two minimal isometric dilations of T are isomorphic. In particular, every minimal isometric dilation of T is indeed a lifting.

Proof. We put $K_+ = \bigoplus_{n\geq 0} H$, and embed H into K_+ by $h \mapsto (h, 0, \ldots)$. We define V by

$$V = \begin{bmatrix} T & 0 & 0 & \cdots \\ D_T & 0 & 0 & \cdots \\ 0 & I & 0 & \cdots \\ 0 & 0 & I & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Direct calculation shows that $V^*V = I$. Hence, V is isometric, and it is apparent from its definition that V is a lifting of T, and therefore, a dilation of T.

This lifting is not minimal in general. It is possible to check that if one repeats the preceding construction with $K_+ = H \oplus (\bigoplus_{n\geq 1} \mathfrak{D}_T)$, then the lifting obtained is minimal. However, it is easier to note that for every isometric dilation (lifting) V of T, the subspace $\widehat{K}_+ = \bigvee_{n\geq 0} V^n H$ is invariant for V and the restriction $V|\widehat{K}_+$ continues to be an isometric dilation (lifting) of T.

Observe that if $V \in \mathcal{B}(K_+)$ is some isometric dilation of T and $h, h' \in H$, then

$$\langle V^n h, V^m h' \rangle = \begin{cases} \langle V^{n-m} h, h' \rangle = \langle T^{n-m} h, h' \rangle, & \text{if } n \geq m \geq 0, \\ \langle h, V^{m-n} h' \rangle = \langle h, T^{m-n} h' \rangle, & \text{if } m \geq n \geq 0, \end{cases}$$

so $\langle V^n h, V^m h' \rangle$ does not depend on the particular choice of V.

Now assume that $V_1 \in \mathcal{B}(K_1)$ and $V_2 \in B(K_2)$ are minimal isometric dilations of T. Then the linear map U defined by

$$UV_1^n h = V_2^n h, \qquad h \in H,\ n \geq 0,$$

is well defined and continues to a unitary $U : K_1 \to K_2$ such that $Uh = h$ for every $h \in H$. Hence, the isometric dilations V_1 and V_2 are isomorphic. $\square$

The following lemma regarding isometric liftings will be useful in the next chapter.

Lemma 1.6. *Let $T \in \mathcal{B}(H)$ be a contraction, $V \in \mathcal{B}(K)$ an isometric lifting of T. Put*

$$K_+ = \bigvee_{n \geq 0} V^n H.$$

Then K_+ reduces V and $V|K_+$ is the minimal isometric dilation of T.

Proof. Clearly, K_+ is invariant for V. We show that it is also invariant for V^*. Since V is a lifting of T, we have $V^*|H = T^*$. This implies that $V^*H = T^*H \subset H$. Also, since V is an isometry, for $n \geq 1$, we get $V^*V^nH = V^{n-1}H \subset K_+$. This implies that K_+ is invariant for V^*.

The fact that $V|K_+$ is the minimal isometric dilation of T was already observed in the proof of Theorem 1.5. □

Now we would like to show that an isometry V can be dilated to a unitary. Applying this to the isometry V obtained in the preceding Theorem 1.5, we will obtain a unitary dilation of a contraction T. The unitary dilation of an isometry V can be constructed in three different but related ways. The first one is by applying Theorem 1.5 to V^*, the second one uses the so called Kolmogorov-von Neumann-Wold decomposition of an isometry, and the third one constructs the dilation using a matrix representation, as in the proof of the Theorem 1.5 above.

Lemma 1.7. *If $V \in \mathcal{B}(K_+)$ is an isometry, and $U \in \mathcal{B}(K)$ is the minimal isometric dilation of V^*, then U is unitary.*

Proof. We must show that U^* is an isometry. If $n \geq 1$ and $k' \in K_+$, we have

$$\|U^*U^nk\| = \|U^{n-1}k\| = \|U^nk\|,$$

because U is an isometry. Also, since U is in fact a lifting of V^*, we have $U^*|K_+ = V$, so

$$\|U^*k\| = \|Vk\| = \|k\|,$$

because V is an isometry. It follows that $D_{U^*}U^n|K_+ = 0$ for $n \geq 0$. Since $K = \bigvee_{n\geq 0} U^nK_+$, this implies that U^* is an isometry. □

To obtain a unitary dilation of a contraction $T \in \mathcal{B}(H)$, we first use Theorem 1.5 to obtain $V \in \mathcal{B}(K_+)$ an isometric dilation of T. Then we construct U the minimal isometric dilation of V^*, using again Theorem 1.5. By Lemma 1.7, U^* is a unitary extension of V. It remains to see that U^* is a unitary dilation of V, but this is true because U^* is a dilation of V and V is a dilation of T.

We can also construct a unitary dilation $W \in \mathcal{B}(\widehat{K})$ of $T \in \mathcal{B}(H)$ which is minimal, in the sense that $\widehat{K} = \bigvee_{n\in\mathbb{Z}} W^nH$. To do this, we put $\widehat{K} = \bigvee_{n\in\mathbb{Z}} U^{*n}H$ and observe that $\widehat{K}$ reduces U^*, so that $W = U^*|\widehat{K}$ is a minimal unitary dilation of T.

If in the previous construction one chooses V to be the minimal isometric dilation of T, it is easy to see that the obtained U^* is minimal. In fact, note that $U^{*n}H = V^nH$ for $n \geq 0$, so that

$$K = \bigvee_{n\geq 0} U^nK_+ = \bigvee_{n\geq 0} U^n\Big(\bigvee_{m\geq 0} V^mH\Big) = \bigvee_{n\in\mathbb{Z}} U^{*n}H.$$

Thus, we have proved the following Theorem.

Theorem 1.8 (Sz.-Nagy)**.** *Any contraction $T \in \mathcal{B}(H)$ can be dilated to a unitary $U \in \mathcal{B}(K)$. This dilation can be chosen to be minimal, in the sense that*

$$K = \bigvee_{n\in\mathbb{Z}} U^n H. \tag{1.12}$$

If one puts

$$K_+ = \bigvee_{n\geq 0} U^n H, \tag{1.13}$$

then $U_+ = U|K_+$ is the minimal isometric dilation of T. Moreover, any two minimal unitary dilations of T are isomorphic.

Proof. The existence has already been proved, the fact that U_+ is the minimal isometric dilation of T is obvious, and the uniqueness of the minimal unitary dilation is shown in a similar way to the uniqueness of the minimal isometric dilation in Theorem 1.5. □

The second proof of Theorem 1.8 uses the Kolmogorov-von Neumann-Wold decomposition, a structure result for isometric operators.

Theorem 1.9 (Kolmogorov-von Neumann-Wold decomposition)**.** *Let $V \in \mathcal{B}(H)$ be an isometry. Then there is a decomposition $H = H_0 \oplus H_1$ such that H_0 and H_1 reduce V, the restriction $V|H_0$ is unitary, and $V|H_1$ is a unilateral shift.*

Moreover, the decomposition is uniquely determined by

$$H_0 = \bigcap_{n\geq 0} V^n H, \qquad H_1 = M_+(L), \qquad L = H \ominus VH.$$

Proof. First note that L is wandering for V. Indeed, $V^n L \subset VL \perp L$, so $H_1 = M_+(L)$ is well defined. Since

$$L \oplus VL \oplus \cdots \oplus V^n L = (H \ominus VH) \oplus (VH \ominus V^2 H) \oplus (V^n H \ominus V^{n+1} H) = H \ominus V^{n+1} H,$$

we see that $H_0 = H \ominus H_1$. The spaces $V^n H$ form a non-increasing chain and therefore $VH_0 = H_0$. In particular, H_0 is V invariant. The subspace H_1 is also V invariant. Hence, both H_0 and H_1 reduce V. It is clear that $V|H_1$ is a unilateral shift, and $V|H_0$ is unitary, because $VH_0 = H_0$.

To prove the uniqueness of the decomposition, assume that $H = H_0' \oplus H_1'$ is a decomposition such that $VH_0' = H_0'$ and $H_1' = M_+(L')$, where L' is a wandering space for $V|H_1'$. Then

$$L = H \ominus VH = (H_0' \oplus H_1') \ominus (VH_0' \oplus VH_1') = H_1' \ominus VH_1' = L',$$

and this implies $H_0' = H_0$, $H_1' = H_1$. □

To construct a unitary dilation of a contraction T, we first use Theorem 1.5 to construct V an isometric dilation of T. By Theorem 1.9, V is a direct sum $V = V_0 \oplus V_1$ of a unitary and a unilateral shift. Let U_1 be the bilateral shift extension of V_1. Then $U = V_0 \oplus U_1$ is a unitary extension of V, and hence a unitary dilation of T. This unitary dilation is not minimal in general, but if we choose V to be the minimal isometric dilation of T, then one can check that this construction produces the minimal unitary dilation of T.

For the third proof of Theorem 1.8, put $K = \bigoplus_{n\in\mathbb{Z}} H$ and define U by the matrix

$$U = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ \cdots & I & 0 & 0 & 0 & 0 & \cdots & \\ \cdots & 0 & I & 0 & 0 & 0 & 0 & \cdots \\ \cdots & 0 & 0 & D_{T^*} & \check{T} & 0 & 0 & \cdots \\ \cdots & 0 & 0 & -T^* & D_T & 0 & 0 & \cdots \\ \cdots & 0 & 0 & 0 & 0 & I & 0 & \cdots \\ \cdots & 0 & 0 & 0 & 0 & 0 & I & \cdots \\ \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{1.14}$$

(Here the symbol $\check{}$ marks the 00-th entry). Then one can check by direct computation that U is unitary. In doing this, one should use the intertwining relations

$$TD_T = D_{T^*}T, \qquad D_T T^* = T^* D_{T^*}. \tag{1.15}$$

The first relation is true because $TD_T^2 = D_{T^*}^2 T$ (note that D_T is a uniform limit of polynomials of D_T^2), and the second is obtained by taking adjoints. This dilation is not minimal, but if one repeats the construction with $K = (\bigoplus_{n<0} \mathfrak{D}_{T^*}) \oplus \check{H} \oplus (\bigoplus_{n\geq 1} \mathfrak{D}_T)$, where we mark with $\check{}$ the summand in the 0-th position, then the obtained U is minimal. Here one has to use (1.15) to see that U acts correctly in this space.

1.5. The Sz.-Nagy–Foiaș functional model of a $C_{\cdot 0}$ contraction

The geometric study of the minimal unitary dilation of a contraction gives rise to the Sz.-Nagy–Foiaş analytic model of the contraction. Here, for the sake of the simplicity of the exposition, we will construct the functional model for the case of a $C_{\cdot 0}$ contraction, which is somewhat simpler. The statement of the model theorem for a general contraction will be given without proof at the end of the section.

We say that a contraction $T \in \mathcal{B}(H)$ is of class $C_{\cdot 0}$ if $T^{*n} \to 0$ strongly as $n \to \infty$. Let $U \in \mathcal{B}(K)$ be the minimal unitary dilation of a contraction $T \in \mathcal{B}(H)$, and $U_+ \in \mathcal{B}(K_+)$ the minimal isometric dilation of T. We consider U_+ as a restriction of U, as in Theorem 1.8. A moment's thought shows that U_+ is a unilateral shift if and only if T is of class $C_{\cdot 0}$. Indeed, Theorem 1.9 shows that an isometry is a unilateral shift if and only if it is of class $C_{\cdot 0}$. Now, since U_+^* is an extension of T^*, if $n \geq m \geq 0$ and $h \in H$,

$$U_+^{*n} U_+^m h = U_+^{*(n-m)} h = T^{*(n-m)} h.$$

As U_+ is minimal $K_+ = \bigvee_{m\geq 0} U_+^m H$, so this shows that if T is $C_{\cdot 0}$ then U_+ is also $C_{\cdot 0}$. The converse is obvious.

Therefore, the contractions of class $C_{\cdot 0}$ are precisely the ones whose minimal isometric dilation is a unilateral shift (and hence, its minimal unitary dilation is a bilateral shift). In the case of a general contraction, the minimal isometric dilation has a unitary part and this part has to be taken into account in the construction of the model.

The theorem that we are going to prove is the following.

Theorem 1.10 (Sz.-Nagy–Foiaş)**.** *Let T be a $C_{\cdot 0}$ contraction on a separable Hilbert space. Define Θ_T the characteristic function of T by*

$$\Theta_T(z) = [-T + zD_{T^*}(I - zT^*)^{-1}D_T]|\mathfrak{D}_T.$$

Then Θ_T is an inner function and T is unitarily equivalent to the operator $\mathcal{T}$ on $H^2(\mathfrak{D}_{T^*}) \ominus \Theta_T H^2(\mathfrak{D}_T)$ defined by

$$(\mathcal{T}^* u)(z) = \frac{u(z) - u(0)}{z}, \qquad u \in H^2(\mathfrak{D}_{T^*}) \ominus \Theta_T H^2(\mathfrak{D}_T).$$

The minimal isometric dilation of $\mathcal{T}$ is the operator M_z of multiplication by z on the space $H^2(\mathfrak{D}_{T^*})$ and the minimal unitary dilation of $\mathcal{T}$ is the operator $M_{e^{it}}$ of multiplication by e^{it} on the space $L^2(\mathfrak{D}_{T^*})$.

Proof. Assume that $T \in \mathcal{B}(H)$ is a $C_{\cdot 0}$ contraction and that we have constructed its minimal unitary dilation U by using the matrix form (1.14) acting on the space

$$K = (\bigoplus_{n<0} \mathfrak{D}_{T^*}) \oplus \check{H} \oplus (\bigoplus_{n \geq 1} \mathfrak{D}_T)$$

(here and in the sequel we mark with $\check{}$ the 0-th position whenever ambiguity could arise). We also consider the minimal isometric dilation U_+ as the restriction of U to the subspace

$$K_+ = H \oplus (\bigoplus_{n \geq 1} \mathfrak{D}_T).$$

The isometry U_+ is a unilateral shift and its wandering space is $L = K_+ \ominus U_+ K_+ = \ker U_+^*$. Using the matrix form of U_+, we see that

$$L = \ker U_+^* = \{(h, d, 0, \ldots) \in K_+ : T^* h + D_T d = 0\}. \tag{1.16}$$

Now we fix a vector $k = (0, d, 0, \ldots) \in K_+$. Since $K_+ = M_+(L)$, then k can be written as

$$k = \sum_{n \geq 0} U_+^n l_n, \qquad l_n \in L. \tag{1.17}$$

We want to find the expressions in terms of d for the "Fourier coefficients" l_n.

To do this, let P_L be the orthogonal projection onto L. Fix a vector

$$k_0 = (h_0, d_0, 0, \ldots) \in K_+.$$

Now we compute $P_L k_0$. Since $P_L k_0 \in L$, we have

$$P_L k_0 = (h_1, d_1, 0, \ldots), \tag{1.18}$$

for some $h_1 \in H$, $d_1 \in \mathfrak{D}_T$ such that

$$T^* h_1 + D_T d_1 = 0 \tag{1.19}$$

Also, $(I - P_L) k_0 \in K_+ \ominus L = U_+ K_+$. Hence, there is some $k' \in K_+$ such that

$$(I - P_L) k_0 = U_+ k'.$$

Using the matrix form of U_+ and equating the components of the vectors, we get that there is some $h' \in H$ such that

$$h_1 = T h' + h_0, \qquad d_1 = D_T h' + d_0. \tag{1.20}$$

Now we substitute (1.20) into (1.19) to find that

$$0 = T^*h_1 + D_T d_1 = T^*Th' + T^*h_0 + D_T^2 h' + D_T d_0 = h' + T^*h_0 + D_T d_0.$$

This equation determines h' in terms of $k_0 = (h_0, d_0, 0, \ldots)$, and hence, it also determines h_1 and d_1 by (1.20). Therefore, by (1.18), we get

$$P_L(h_0, d_0, 0, \ldots) = (D_{T^*}^2 h_0 - TD_T d_0, T^*Td_0 - D_T T^* h_0, 0, \ldots). \tag{1.21}$$

Returning to the problem of computing the Fourier coefficients l_n of $k = (0, d, 0, \ldots)$, we see that

$$l_0 = P_L k = (-TD_T d, T^*Td, 0, \ldots), \tag{1.22}$$

and for $n \geq 1$

$$l_n = P_L U_+^{*n} k = P_L(T^{*(n-1)} D_T d, 0, \ldots) = (D_{T^*}^2 T^{*(n-1)} D_T d, -D_T T^{*n} D_T d, 0, \ldots). \tag{1.23}$$

Here, the first equality is obtained using the matrix form of U_+.

Since the restriction of U_+ to $K_+ \ominus H$ is a unilateral shift with wandering space $\mathfrak{D}_T$, this computation indeed allows us to give the expression for the Fourier coefficients of any $k \in K_+ \ominus H$. Hence, we can describe the subspace $K_+ \ominus H$ by means of the Fourier coefficients.

Until now, the subspace L has played an important role, but it is unfortunate that it does not have a simple enough expression in terms of the operator T. We will remedy this by showing that

$$U^*L = \cdots \oplus 0 \oplus \mathfrak{D}_{T^*} \oplus \check{0} \oplus 0 \oplus \cdots, \tag{1.24}$$

so that

$$W = U|(\cdots \oplus 0 \oplus \mathfrak{D}_{T^*} \oplus \check{0} \oplus 0 \oplus \cdots)$$

gives an isomorphism

$$W : \mathfrak{D}_{T^*} \to L.$$

The inclusion $\subset$ in (1.24) is obtained by taking an arbitrary $l \in L$, noting that

$$l = (\ldots, 0, \check{h}, d, 0, \ldots), \qquad \text{with } T^*h + D_T d = 0$$

by (1.16), and computing U^*l using the matrix form of U. To show the reverse inclusion, take the vector $k = (\ldots, 0, d, \check{0}, 0, \ldots)$, note that $Uk = (\ldots, 0, \overgroup{D_{T^*}d}, -T^*d, 0, \ldots)$ and check that $Uk \in L$ using (1.15) and (1.16).

In particular, this allows us to identify the Fourier coefficients l_n with elements of $\mathfrak{D}_{T*}$. Using

$$W^*(h, d, 0, \ldots) = D_{T^*}h - Td, \qquad (h, d, 0, \ldots) \in L,$$

which comes from the matrix form of U, and relations (1.15), (1.22) and (1.23), we compute

$$\begin{aligned} W^*l_0 &= -(D_T^* T D_T - TT^*T)d = -Td \\ W^*l_n &= (D_{T^*}^3 T^{*(n-1)} D_T + TD_T T^{*n} D_T)d = D_{T^*} T^{*(n-1)} D_T d, \qquad n \geq 1. \end{aligned} \tag{1.25}$$

This is all the information about the minimal unitary dilation that we need to construct the Sz.-Nagy–Foiaş model. To actually construct the model, we need the L^2, H^2 and H^∞ spaces introduced in Section 1.2.

Applying the results of Section 1.3 about the relation between Hardy spaces and shifts, we see that the operator $Z : K \to L^2(\mathfrak{D}_{T^*})$ defined by

$$Z \sum_{n\in\mathbb{Z}} U^n l_n = \sum_{n\in\mathbb{Z}} e^{int} W^* l_n$$

is a unitary which transforms the operator U into the operator $M_{e^{it}}$, in the sense that $ZU = M_{e^{it}}Z$. The restriction $Z_+ = Z|K_+$ maps K_+ onto the Hardy space $H^2(\mathfrak{D}_{T^*})$ and takes U_+ into the operator M_z. Thus, $L^2(\mathfrak{D}_{T^*})$ and $H^2(\mathfrak{D}_{T^*})$ will be our model spaces, modelling K and K_+ respectively, and $M_{e^{it}}$ and M_z will be our model operators, modelling U and U_+ respectively.

We need to identify which subspace of $H^2(\mathfrak{D}_{T^*})$ is associated with the subspace H (i.e., we need to compute ZH), and that is the reason why we where interested in computing the Fourier coefficients of a vector $k \in K_+ \ominus H$. To present the results (1.25) in a nice form, we will use a contractive analytic function.

First note that, as we have already observed, $U_+|K_+ \ominus H$ is a unilateral shift with wandering space $0 \oplus \mathfrak{D}_T \oplus 0 \oplus \ldots$. Hence, the operator $X : H^2(\mathfrak{D}_T) \to K_+ \ominus H$ defined by

$$X \sum_{n\geq 0} z^n d_n = (0, d_0, d_1, \ldots)$$

is a unitary which transforms the operator M_z in $H^2(\mathfrak{D}_T)$ into the operator $U_+|K_+ \ominus H$ (i.e, $XM_z = U_+X$).

Consider the operator $Q = Z_+X : H^2(\mathfrak{D}_T) \to H^2(\mathfrak{D}_{T^*})$. This operator is an isometry, and its range is precisely $QH^2(\mathfrak{D}_T) = Z_+(K_+ \ominus H)$. Observe that, by the properties of Z_+ and X, we have

$$M_zQ = M_zZ_+X = Z_+U_+X = Z_+XM_z = QM_z.$$

Therefore, by Lemma 1.2, we see that there exists a contractive analytic function $\Theta_T \in H^\infty(\mathfrak{D}_T, \mathfrak{D}_{T^*})$ such that Q is the operator of multiplication by Θ_T. Since Q is an isometry, Θ_T is inner.

To find an explicit formula for Θ_T, we use (1.25). Let $d \in \mathfrak{D}_T$, and denote also by d the function in $H^2(\mathfrak{D}_T)$ which is constantly d. Then, if we put $k = (0, d, 0, \ldots)$ and define l_n by (1.17),

$$\Theta_T d = Z_+Xd = Z_+(0, d, 0, \ldots) = \sum_{n\geq 0} z^n W^* l_n = -Td + \sum_{n\geq 1} z^n D_{T^*} T^{*(n-1)} D_T d.$$

Hence,

$$\Theta_T(z) = \left[-T + \sum_{n\geq 1} z^n D_{T^*} T^{*(n-1)} D_T \right] \Big| \mathfrak{D}_T.$$

This series can be summed to obtain

$$\Theta_T(z) = [-T + zD_{T^*}(I - zT^*)^{-1}D_T]|\mathfrak{D}_T.$$

The function $\Theta_T(z)$ is called the *characteristic function* of T. We have seen that the subspace $K_+ \ominus H$ is modelled by $\Theta_T H^2(\mathfrak{D}_T)$, in the sense that $Z_+|K_+ \ominus H$ is a unitary which maps $K_+ \ominus H$ onto $\Theta_T H^2(\mathfrak{D}_T)$.

Finally, it is also possible to compute the action of the operator T^* on the model space. Noting that $T^* = U_+^*|H$, it suffices to compute the action of U_+^* in the model space $H^2(\mathfrak{D}_{T^*})$. However, U_+ acts as M_z on $H^2(\mathfrak{D}_{T^*})$, so its adjoint is easily seen to act as the operator $u \mapsto z^{-1}(u(z) - u(0))$. □

As we have already commented, for the case of a general contraction T, the model has to take into account the unitary part of the minimal isometric dilation of T. We will only give the statement of the theorem. A contraction $T \in \mathcal{B}(H)$ is called completely non-unitary if no nontrivial subspace of H reduces T to a unitary operator. If T is a general contraction, then it is a direct sum of a completely non-unitary contraction and a unitary, so there is no loss of generality in constructing the model only for completely non-unitary contractions.

Theorem 1.11 (Sz.-Nagy–Foiaş)**.** *Let T be a completely non-unitary contraction on a separable Hilbert space. Let Θ_T be the characteristic function of T, defined as in Theorem 1.10. Put $\Delta(e^{it}) = (I - \Theta_T(e^{it})^*\Theta_T(e^{it}))^{\frac{1}{2}}$ and denote by Δ the operator of multiplication by $\Delta(e^{it})$ on $L^2(\mathfrak{D}_T)$.*

Put

$$\mathcal{K} = L^2(\mathfrak{D}_{T^*}) \oplus \overline{\Delta L^2(\mathfrak{D}_T)}, \qquad \mathcal{K}_+ = H^2(\mathfrak{D}_{T^*}) \oplus \overline{\Delta L^2(\mathfrak{D}_T)},$$

and

$$\mathcal{H} = \mathcal{K}_+ \ominus \{\Theta_T u \oplus \Delta u : u \in H^2(\mathfrak{D}_T)\}.$$

Then T is unitarily equivalent to the operator $\mathcal{T}$ on $\mathcal{H}$ defined by

$$(\mathcal{T}^*(u,v))(z) = \left(\frac{u(z)-u(0)}{z}, e^{-it}v(e^{it})\right), \qquad (u,v) \in \mathcal{H}.$$

The minimal isometric dilation of $\mathcal{T}$ is the operator $M_z \oplus (M_{e^{it}}|\overline{\Delta L^2(\mathfrak{D}_T)})$ acting on the space $\mathcal{K}_+$, and the minimal unitary dilation of $\mathcal{T}$ is the operator $M_{e^{it}} \oplus (M_{e^{it}}|\overline{\Delta L^2(\mathfrak{D}_T)})$ on the space $\mathcal{K}$.

We see that when Θ_T is inner (which happens if and only if T is $C_{\cdot 0}$), then $\Delta = 0$, so the second component in all the model spaces collapses and we recover Theorem 1.10.

1.6. Invariant subspaces

As an application of the Sz.-Nagy–Foiaş model, we will give a characterization of the invariant subspaces of a $C_{0\cdot}$ contraction T in terms of the factorizations of the characteristic function of its adjoint operator. We say that a contraction T is of class $C_{0\cdot}$ if $T^n \to 0$ strongly (i.e., if T^* is $C_{\cdot 0}$). This will also allow us to find the invariant subspaces of $C_{\cdot 0}$ contractions.

Using the functional model for T^*, we can assume that T is the operator M_z^* acting on the space $H^2(\mathscr{Y}) \ominus \Theta H^2(\mathscr{U})$, where $\Theta \in H^\infty(\mathscr{U}, \mathscr{Y})$ is some inner function. Now it is evident that the invariant subspaces of T are precisely those of the form $H^2(\mathscr{Y}) \ominus E$, where E is a subspace invariant for M_z and such that $\Theta H^2(\mathscr{U}) \subset E$. Indeed, a subspace $H^2(\mathscr{Y}) \ominus E$ will be invariant for M_z^* if and only if E is invariant for M_z, and the condition $\Theta H^2(\mathscr{U}) \subset E$ comes from the fact that we are only interested in the invariant subspaces which are contained in $H^2(\mathscr{Y}) \ominus \Theta H^2(\mathscr{U})$.

It is clear that if $\Theta_1 \in H^\infty(\mathscr{U}_1, \mathscr{Y})$ is an inner function, then the subspace $\Theta_1 H^2(\mathscr{U}_1)$ is invariant for M_z. Moreover, if Θ can be factorized as $\Theta = \Theta_1\Theta_2$, where $\Theta_2 \in H^2(\mathscr{U}, \mathscr{U}_1)$ is another inner function, then $\Theta_1 H^2(\mathscr{U}_1) \supset \Theta H^2(\mathscr{U})$. Hence, for every such factorization, we obtain a corresponding invariant subspace for T. What we want to do now is to prove that all the invariant subspaces for T can be obtained in this manner. The invariant subspaces for the operator M_z on $H^2(\mathscr{Y})$ are well known.

Theorem 1.12 (Beurling-Lax-Halmos)**.** *Let $\mathscr{Y}$ be a separable Hilbert space. A subspace $E \subset H^2(\mathscr{Y})$ is invariant for the operator M_z of multiplication by z if and only if E is of the form $E = \Theta H^2(\mathscr{U})$ for some inner function $\Theta \in H^\infty(\mathscr{U}, \mathscr{Y})$.*

Proof. The proof is very similar to the construction of the characteristic function Θ_T that we did in the previous section.

Assume that E is invariant for M_z and put $V = M_z|E$. Then V is an isometry. Since M_z is a unilateral shift, Theorem 1.9 implies that V is a unilateral shift, because

$$\bigcap_{n\geq 0} V^n E \subset \bigcap_{n\geq 0} M_z^n H^2(\mathscr{Y}) = 0.$$

Let $\mathscr{U}$ be the wandering space for V and define $Q : H^2(\mathscr{U}) \to H^2(\mathscr{Y})$ by

$$Q \sum_{n\geq 0} z^n u_n = \sum_{n\geq 0} V^n u_n.$$

Then Q is an isometry and $QH^2(\mathscr{U}) = E$. Moreover, $M_z Q = QM_z$. By Lemma 1.2, there is a $\Theta \in H^\infty(\mathscr{U}, \mathscr{Y})$ such that $Q = \Theta$. This Θ must be inner, because Q is an isometry.

The converse statement is clear. □

Another proof, based on the Lax Theorem about the invariant subspaces of $M_{e^{it}}$ in $L^2(\mathscr{Y})$ can be found in [Nik02b, Corollary 2.1.7]. In the case when $\mathscr{Y}$ has dimension 1, this Theorem is just the Beurling Theorem, which is well known in the theory of Hardy spaces (see, for instance, [MAR07, Corollary 2.3.12] or [Nik02a, Corollary 1.4.1]).

Now we want to see that if $\Theta_1 H^2(\mathscr{U}_1)$ is some invariant subspace for M_z such that $\Theta_1 H^2(\mathscr{U}_1) \supset \Theta H^2(\mathscr{U})$, then we can factorize Θ as $\Theta_1\Theta_2$, where Θ_2 is inner.

Lemma 1.13. *If $\Theta \in H^\infty(\mathscr{U}, \mathscr{Y})$ and $\Theta_1 \in H^\infty(\mathscr{U}_1, \mathscr{Y})$ are inner functions such that $\Theta H^2(\mathscr{U}) \subset \Theta_1 H^2(\mathscr{U}_1)$, then there is an inner function $\Theta_2 \in H^\infty(\mathscr{U}, \mathscr{U}_1)$ such that $\Theta = \Theta_1\Theta_2$.*

Proof. Note that the equation

$$\Theta f = \Theta_1 V f, \qquad f \in H^2(\mathscr{U})$$

defines correctly an isometry $V : H^2(\mathscr{U}) \to H^2(\mathscr{U}_1)$. Also,

$$\Theta_1 V M_z = \Theta M_z = M_z \Theta = M_z \Theta_1 V = \Theta_1 M_z V.$$

Since Θ_1 is an isometry, we get $VM_z = M_z V$. Applying Lemma 1.2 to V, we see that there exists a contractive analytic function $\Theta_2 \in H^\infty(\mathscr{U}, \mathscr{U}_1)$ such that $V = \Theta_2$. Since V is an isometry, Θ_2 must be inner. This proves the Lemma. □

Using this Lemma we can also prove that if $\Theta_1 H^2(\mathscr{U}_1) = \Theta_2 H^2(\mathscr{U}_2)$ and Θ_1, Θ_2 are inner, then $\Theta_1 = \Theta_2 Z$, for some constant unitary $Z \in H^\infty(\mathscr{U}_2, \mathscr{U}_1)$. Indeed, we get from the Lemma inner functions Θ_3 and Θ_4 such that $\Theta_1 = \Theta_2\Theta_3$ and $\Theta_2 = \Theta_1\Theta_4$. We see that $\Theta_1 = \Theta_1\Theta_4\Theta_3$, and since Θ_1 is isometric, $\Theta_4\Theta_3 = I$. In particular, we have $\Theta_4(0)\Theta_3(0) = I$. Similarly, we also get $\Theta_3(0)\Theta_4(0) = I$. Since $\Theta_3(0)$ and $\Theta_4(0)$ are contractions, this implies that they must be unitaries. By Lemma 1.4, Θ_3 and Θ_4 are constant unitaries. The claim now follows, with $Z = \Theta_3$.

To sum up the results of this Section, we have proved the following Theorem.

Theorem 1.14. *If T is a $C_{0\cdot}$ contraction on a separable Hilbert space H, to each invariant subspace $F \subset H$ of T there corresponds a factorization $\Theta_{T^*} = \Theta_1\Theta_2$ of the characteristic function of T^* as the product of two inner functions $\Theta_1 \in H^\infty(\mathscr{U}_1, \mathfrak{D}_T)$ and $\Theta_2 \in H^\infty(\mathfrak{D}_{T^*}, \mathscr{U}_1)$, such that the subspace F is modelled by the subspace $H^2(\mathfrak{D}_T) \ominus \Theta_1 H^2(\mathscr{U}_1)$ in the Sz.-Nagy–Foiaş model of T^* given by Theorem 1.10.*

Conversely, to every such factorization $\Theta_{T^*} = \Theta_1\Theta_2$ as product of inner functions, there corresponds an invariant subspace F of T. Moreover, the identification of invariant subspaces F with inner factors Θ_1 is uniquely determined by F up to a constant unitary factor on the right.

Indeed, from the work done in this section, it is not hard to see how the lattice of invariant subspaces of T and the lattice of factorizations of Θ_{T^*} are isomorphic in some sense (which can be made precise).

It is worthy to mention that, after multiplication by a suitable nonzero scalar, every operator T can be assumed to be a $C_{0\cdot}$ contraction. Hence, in principle, the study of the invariant subspaces could be reduced to the study of factorizations of inner functions. However, this topic is not very well understood. The Theorem is most useful when, for instance, the defect space $\mathfrak{D}_T$ is finite dimensional, and hence the study of factorizations of Θ_{T^*} corresponds to factorizations of matrices.

Noting again that, if $T \in \mathcal{B}(H)$, a subspace $F \subset H$ is invariant for T if and only if $H \ominus E$ is invariant for T^*, it is easy to use the preceding Theorem to obtain the invariant subspaces of a $C_{\cdot 0}$ contraction.

Theorem 1.15. *If T is a $C_{\cdot 0}$ contraction on a separable Hilbert space H, to each invariant subspace $F \subset H$ of T there corresponds a factorization $\Theta_T = \Theta_1\Theta_2$ of the characteristic function of T as the product of two inner functions $\Theta_1 \in H^\infty(\mathscr{U}_1, \mathfrak{D}_{T^*})$ and $\Theta_2 \in H^\infty(\mathfrak{D}_T, \mathscr{U}_1)$, such that the subspace F is modelled by the subspace $\Theta_1 H^2(\mathscr{U}_1) \ominus \Theta_T H^2(\mathfrak{D}_T)$ in the Sz.-Nagy–Foiaş model of T given by Theorem 1.10.*

Conversely, to every such factorization $\Theta_T = \Theta_1\Theta_2$ as product of inner functions, there corresponds an invariant subspace F of T, and Θ_1 is determined by F up to a constant unitary factor on the right.

2. Dilations of several operators

In the preceding chapter, we have considered dilations of a single operator. This chapter deals with the generalization of the concept of dilation to tuples of commuting operators. The key fact in this chapter is that, in contrast to the single operator case, a unitary dilation for a tuple of three or more commuting contractions need not exist. The reasons behind the existence or non-existence of the dilation are not completely understood.

First, we prove the Commutant Lifting Theorem, which is related to the topic of this chapter, and give an application to the Nevanlinna-Pick interpolation problem. Then we give Andô's Theorem, which states that a unitary dilation of a pair of commuting contractions always exists. We treat other related topics, such as dilations of isometries, the theory of regular dilations, and von Neumann's inequality. We also give the main examples of non-existence of a dilation that can be found in the literature. Finally, we give a positive existence result by Lotto.

Many of the results of this section can be found in the books [SNFBK10, FF90]. Also, the book by Agler and McCarthy [AM02] contains a good exposition (see especially Chapter 10).

2.1. The Commutant Lifting Theorem

The Commutant Lifting Theorem deals with the description of the operators which commute with a given contraction, in terms of operators commuting with its minimal isometric dilation. The theorem can also be formulated for operators intertwining two contractions, as we will do. This version is not more general, as the statement involving intertwining can be obtained easily from the one involving commutants. However, it gives a more transparent proof of the theorem.

Theorem 2.1 (Commutant Lifting Theorem)**.** *Let $T \in \mathcal{B}(H)$, $T' \in \mathcal{B}(H')$ be two contractions, and $U_+ \in \mathcal{B}(K_+)$, $U'_+ \in \mathcal{B}(K'_+)$ their respective minimal isometric dilations. If $A \in \mathcal{B}(H', H)$ satisfies the intertwining relationship*

$$TA = AT', \tag{2.1}$$

then there is an operator $B \in \mathcal{B}(K'_+, K_+)$ which is a lifting of A, and is such that

$$U_+B = BU'_+ \tag{2.2}$$

and $\|A\| = \|B\|$.

Proof. We will first deal with the case when $T' = V$ is an isometry, so that $U'_+ = V$. Without loss of generality, we can assume that $\|A\| = 1$ (the case $A = 0$ is trivial). We also assume that U_+ acts on the space $K_+ = H \oplus (\bigoplus_{n\geq 1} \mathfrak{D}_T)$ and is given by the matrix

$$U_+ = \begin{bmatrix} T & 0 & 0 & \cdots \\ D_T & 0 & 0 & \cdots \\ 0 & I & 0 & \cdots \\ 0 & 0 & I & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

(See the proof of Theorem 1.5). An operator $B \in \mathcal{B}(H', K_+)$ which is a lifting of A will have the matrix form

$$B = \begin{bmatrix} A^* & B_1^* & B_2^* & \cdots \end{bmatrix}^*,$$

where B_n are operators in $\mathcal{B}(H', \mathfrak{D}_T)$. Since clearly $\|B\| \geq \|A\| = 1$, the condition $\|A\| = \|B\|$ will be satisfied if and only if B is a contraction.

Writing condition (2.2) in matrix form, using (2.1) and equating matrix components, we see that (2.2) is equivalent to

$$D_T A = B_1 V, \qquad B_n = B_{n+1} V, \qquad n \geq 1. \tag{2.3}$$

Hence, we must construct a sequence of operators B_n which satisfy (2.3) and such that the resulting operator B is a contraction.

To simplify the notation, we define

$$B_0 = D_T A.$$

We will now construct by induction a sequence $\{B_n\}_{n=0}^{\infty}$ which satisfies the following conditions:

$$B_N V = B_{N-1}, \qquad N \geq 1, \tag{I}$$

$$A^*A + \sum_{n=1}^{N} B_n^* B_n \leq I, \qquad N \geq 0, \tag{II}$$

and

$$B_N^* B_N \leq V^* D_N^2 V, \qquad N \geq 0, \tag{III}$$

where

$$D_N = \left(I - A^*A - \sum_{n=1}^{N} B_n^* B_n \right)^{\frac{1}{2}},$$

which is correctly defined by (II). Condition (I) comes directly from (2.3), condition (II) will ensure that

$$B^*B = A^*A + \sum_{n=1}^{\infty} B_n^* B_n \leq I,$$

so that B is a contraction, and condition (III) is auxiliary and will be used to construct B_{N+1} from B_N. Indeed, for $N \geq 1$, (III) can be derived from (I), (II) and the definition of B_0 alone, as we will show. Once we have built such a sequence $\{B_n\}_{n=0}^{\infty}$, the operator B will have all the required properties.

First we check that the conditions are satisfied, for $N = 0$, by the operator $B_0 = D_T A$. Condition (I) is vacuous. Condition (II) amounts to $A^*A \leq I$, which is true because $\|A\| = 1$. To check condition (III), we compute

$$B_0^* B_0 = A^* D_T^2 A = A^*A - A^*T^*TA = A^*A - V^*A^*AV \leq I - V^*A^*AV = V^* D_0^2 V.$$

Here we have used (2.1) and the fact that $V^*V = I$ because V is an isometry.

Now assume that we have constructed operators $B_0, \ldots, B_N$ satisfying (I)–(III). Condition (III) implies that $\|B_N h\| \leq \|D_N V h\|$ for all $h \in H'$. Hence there exists a contraction C_N from the linear manifold $D_N V H'$ into $\mathfrak{D}_T$ such that

$$B_N = C_N D_N V. \tag{2.4}$$

This contraction can be extended by continuity to $\overline{D_N V H'}$ and then to all H' by defining it to be zero on $H' \ominus \overline{D_N V H'}$. This extension, which we will continue to call C_N is still a contraction and satisfies (2.4).

It is clear from (I) that we can now define

$$B_{N+1} = C_N D_N.$$

We must check that with this election of B_{N+1}, conditions (II) and (III) for $N+1$ are satisfied. To check (II), we compute

$$A^*A + \sum_{n=1}^{N+1} B_n^* B_n = I - D_N^2 + B_{N+1}^* B_{N+1} = I - D_N^2 + D_N C_N^* C_N D_N \leq I,$$

where the last inequality is true because C_N is a contraction.

Now we check that condition (III) for $N+1$ is satisfied. As we said before, we will derive this fact from (I), (II) and the definition of B_0 alone, without using the specific form in which we constructed B_{N+1}. We compute

$$V^*(I - D_{N+1}^2)V = V^*A^*AV + \sum_{n=1}^{N+1} V^* B_n^* B_n V = V^*A^*AV + \sum_{n=0}^{N} B_n^* B_n.$$

Here we have used (I) in the last equality. Since

$$V^*A^*AV + B_0^* B_0 = V^*A^*AV + A^* D_T^2 A = V^*A^*AV + A^*A - A^*T^*TA = A^*A,$$

we get $V^*(I - D_{N+1}^2)V = I - D_N^2$. This implies $V^* D_{N+1}^2 V = D_N^2$, because V is an isometry. Finally, we see that (III) for $N+1$ is true because $D_N^2 - B_{N+1}^* B_{N+1} = D_{N+1}^2 \geq 0$ (note that here we used (II) to see that D_{N+1} is correctly defined).

This finishes the proof of the case when T' is an isometry. If T' is not an isometry, note that

$$TAP_{H'} = AT'P_{H'} = AP_{H'}U'_+,$$

because the minimal isometric dilation U'_+ of T' is indeed a lifting of T' (here $P_{H'}$ denotes the orthogonal projection onto H'). Hence, we can apply the previous case to the isometry U'_+ instead of T' and the operator $AP_{H'}$ instead of A to obtain an operator B which is a lifting of $AP_{H'}$ and satisfies (2.2) and $\|B\| = \|AP_{H'}\|$. This operator B will do the job, because $\|AP_{H'}\| = \|A\|$, and since $AP_{H'}$ is a lifting of A, it follows that B is a lifting of A. □

It is now easy to generalize the Theorem to allow for arbitrary isometric dilations of the contractions T and T'. This is useful in the applications.

Corollary 2.2. *Let $T \in \mathcal{B}(H)$, $T' \in \mathcal{B}(H')$ be two contractions, and $V \in \mathcal{B}(K)$, $V' \in \mathcal{B}(K')$ two isometric dilations respectively. If $A \in \mathcal{B}(H', H)$ satisfies the intertwining relationship*

$$TA = AT',$$

then there is an operator $B \in \mathcal{B}(K', K)$ which is a lifting of A and such that

$$VB = BV'$$

and $\|A\| = \|B\|$. In particular, this holds when V and V' are the minimal unitary dilations of T and T' respectively.

Proof. As we already commented on the proof of Theorem 1.5, if we put $K_+ = \bigvee_{n\geq 0} VH$, then the isometry V has the structure

$$V = \begin{bmatrix} U_+ & * \\ 0 & * \end{bmatrix},$$

according to the decomposition $K = K_+ \oplus (K \ominus K_+)$, where U_+ is the minimal isometric dilation of T. The isometry V' has an analogous structure if we define $K'_+ = \bigvee_{n\geq 0} V'H',$. If we denote by $B_+ \in \mathcal{B}(K'_+, K_+)$ the lifting of A obtained in the preceding Theorem, it is easy to check that the operator

$$B = \begin{array}{c c} & \begin{array}{cc} K'_+ & K'\ominus K'_+ \end{array} \\ \begin{array}{c} K_+ \\ K\ominus K_+ \end{array} & \begin{bmatrix} B_+ & 0 \\ 0 & 0 \end{bmatrix} \end{array}$$

satisfies the needed properties. □

2.2. An application: Nevanlinna-Pick interpolation

As an application of the Commutant Lifting Theorem, we will show how it can be used to solve the classical Nevanlinna-Pick interpolation problem. Given a finite number of distinct points $\{\alpha_j\}_{j=1}^n$ in $\mathbb{D}$, and another set of points $\{\beta_j\}_{j=1}^n \subset \mathbb{D}$ (not necessarily distinct), the problem asks to construct (if it is possible) a function $f \in H^\infty(\mathbb{D})$ with $\|f\|_{H^\infty(\mathbb{D})} \leq 1$ and $f(\alpha_j) = \beta_j$, for $1 \leq j \leq n$.

Theorem 2.3. *The Nevanlinna-Pick interpolation problem has a solution if and only if the matrix*

$$\left[\frac{1-\overline{\beta_j}\beta_k}{1-\overline{\alpha_j}\alpha_k}\right]_{j,k=1}^n \tag{2.5}$$

is non-negative.

Proof. Let us start by defining operators $X, Y \in \mathcal{B}(\mathbb{C}^n, H^2(\mathbb{D}))$ by the matrices

$$X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}, \qquad Y = \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \\ \beta_1\alpha_1 & \beta_2\alpha_2 & \cdots & \beta_n\alpha_n \\ \beta_1\alpha_1^2 & \beta_2\alpha_2^2 & \cdots & \beta_n\alpha_n^2 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

(Here we use the orthonormal basis in $H^2(\mathbb{D})$ given by the monomials $1, z, z^2, \ldots$). Note that X and Y are bounded because $|\alpha_j| < 1$.

Given a function $f \in H^\infty(\mathbb{D})$, we define $\widetilde{f} \in H^\infty(\mathbb{D})$ by $\widetilde{f}(z) = \overline{f(\overline{z})}$. We put $Z = M_{\widetilde{f}}^* \in \mathcal{B}(H^2(\mathbb{D}))$, where $M_{\widetilde{f}}$ denotes the operator of multiplication by $\widetilde{f}$ in $H^2(\mathbb{D})$. Observe that if $f(z) = \sum_{n\geq 0} a_n z^n$ is the power series representation of f, then the matrix representation of the operator Z is

$$Z = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 & \cdots \\ 0 & a_0 & a_1 & a_2 & \cdots \\ 0 & 0 & a_0 & a_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{2.6}$$

Hence, we see that

$$ZX = \begin{bmatrix} f(\alpha_1) & f(\alpha_2) & \cdots & f(\alpha_n) \\ \alpha_1 f(\alpha_1) & \alpha_2 f(\alpha_2) & \cdots & \alpha_n f(\alpha_n) \\ \alpha_1^2 f(\alpha_1) & \alpha_2^2 f(\alpha_2) & \cdots & \alpha_n^2 f(\alpha_n) \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

This shows that f satisfies $f(\alpha_j) = \beta_j$, for $1 \leq j \leq n$, if and only if $ZX = Y$.

Also, $\|Z\| = \|M_{\widetilde{f}}\| = \|\widetilde{f}\|_{H^\infty(\mathbb{D})} = \|f\|_{H^\infty(\mathbb{D})}$, so that $\|f\|_{H^\infty(\mathbb{D})} \leq 1$ if and only if Z is a contraction. Therefore, we are interested in the existence of contractive solutions Z of $ZX = Y$ having the form (2.6).

Assume that the interpolation problem has a solution, so that such a Z exists. Then we see that $Y^*Y \leq X^*X$, or equivalently, $X^*X - Y^*Y \geq 0$. A computation reveals that the matrix $X^*X - Y^*Y$ is precisely (2.5).

Conversely, assume that $X^*X \geq Y^*Y$. We will use the Commutant Lifting Theorem to construct a solution of the interpolation problem. We define the diagonal matrix Λ, which has the diagonal $[\alpha_1, \ldots, \alpha_n]$. Note that

$$M_z^* X = X\Lambda, \qquad M_z^* Y = Y\Lambda, \tag{2.7}$$

where M_z is the operator of multiplication in by the independent variable z in $H^2(\mathbb{D})$.

Let us put $H = \overline{XH^2(\mathbb{D})}$. Since, $X^*X \geq Y^*Y$, there is a contraction $C \in \mathcal{B}(H, H^2(\mathbb{D}))$ such that $Y = CX$. We have

$$M_z^* CX = M_z^* Y = Y\Lambda = CX\Lambda = CM_z^* X.$$

Equation (2.7) shows that H is invariant for M_z^*. Hence, we see that

$$M_z^* C = CM_z^* | H.$$

Taking adjoints and noting that $(M_z^*|H)^* = P_H M_z | H$, we get

$$C^* M_z = (P_H M_z | H) C^*.$$

Observe that M_z is an isometric lifting of the contraction $P_H M_z|H$, because H is invariant for M_z^*. Hence, we can apply Corollary 2.2 with $T = P_H M_z|H$, $T' = M_z$, $A = C^*$, $V = M_z$, and $V' = M_z$ to obtain a contraction $B \in \mathcal{B}(H^2(\mathbb{D}))$ which is a lifting of C^* and such that

$$BM_z = M_z B.$$

By Lemma 1.2, there is a function $g \in H^\infty(\mathbb{D})$ with $\|g\|_{H^\infty(\mathbb{D})} \leq 1$ such that $B = M_g$. Since B is a lifting of C^*, then $C = B^*|H$. Hence, we see that $Y = B^*X$. Finally, it is enough to note that if we put $f = \widetilde{g}$, then $Z = B^*$, so that this function f is a solution of the interpolation problem. □

2.3. Extensions, liftings and dilations

The purpose of this section is to generalize the terminology of Section 1.1 to the context of several commuting operators.

We say that a set of operators $T = \{T_j\}_{j \in J} \subset \mathcal{B}(H)$ is a system of commuting operators if all the operators T_j commute, i.e., if $T_j T_k = T_k T_j$ for all $j, k \in J$.

The generalization of extensions and liftings is straightforward. If $T = \{T_j\}_{j \in J} \subset \mathcal{B}(H)$ is a system of commuting operators, we say that $V = \{V_j\}_{j \in J} \subset \mathcal{B}(K)$ is a commuting extension/lifting of this system if the operators V_j commute and each operator V_j is an extension/lifting of the operator T_j.

We say that V is a dilation of T if the operators V_j commute and

$$T_{j_1}^{k_1} \cdots T_{j_n}^{k_n} = P_H V_{j_1}^{k_1} \cdots V_{j_n}^{k_n} | H,$$

for every election of $j_1, \dots, j_n \in J$ and $k_1, \dots, k_n \geq 1$. Note that, in contrast with the case of extensions and liftings, this condition is stronger than asking each operator V_j to be a dilation of T_j. However, it is easy to check that if V is either an extension or a lifting of T, then it is also a dilation.

We say that the extension/lifting/dilation V is isometric/unitary if each operator $V_j \in V$ is isometric/unitary.

If $U = \{U_j\}_{j \in J} \subset \mathcal{B}(K)$ is a dilation of $T = \{T_j\}_{j \in J} \subset \mathcal{B}(H)$, applying Lemma 1.1 to $\mathcal{A} = \mathbb{C}[z_j; j \in J]$, the algebra of polynomials in variables z_j, $j \in J$, and the map φ given by $\varphi(p) = p(\{U_j\})$, we see that the space K decomposes as $K = H_2 \oplus H \oplus (K \ominus H_1)$, and that each operator U_j has the structure

$$U_j = \begin{bmatrix} * & * & * \\ 0 & T_j & * \\ 0 & 0 & * \end{bmatrix}$$

with respect to this decomposition. Hence, if we define $K_+ = H_2 \oplus H$, then K_+ is invariant for every U_j and the operators $V_j = U_j | K_+$ are a lifting of T.

If U was a unitary dilation of T, this construction produces an isometric lifting V of T. Conversely, in Section 2.5 we will show how every system of commuting isometries can be extended to a system of commuting unitaries. Therefore, the problems of finding an isometric lifting of a given system of contractions T and a unitary dilation of the same system T can be thought to be equivalent.

2.4. Dilations of a pair of operators: Andô's theorem

Andô's Theorem was the first achievement in the theory of simultaneous dilations of contractions. It shows that every pair of commuting contractions has a commuting isometric lifting and a commuting unitary dilation. The usual proofs of the theorem limit to show the existence of the isometric lifting. It is known that an arbitrary system of commuting isometries can be dilated to a commuting system of unitaries, so the existence of the unitary dilation follows from this fact. We will prove the theorem about the unitary dilation of a system of isometries in Section 2.5 below.

Theorem 2.4 (Andô's Theorem)**.** *Every pair of commuting contractions $T_1, T_2 \in \mathcal{B}(H)$ has a commuting isometric lifting V_1, V_2 and a commuting unitary dilation U_1, U_2.*

We will give several proofs of Andô's Theorem. The first one shows that it can be obtained as a corollary of the Commutant Lifting Theorem. It is remarkable that with this proof one can construct the unitary dilation by the same means as the isometric lifting.

Proof using the Commutant Lifting Theorem. Since $T_1 T_2 = T_2 T_1$, we can use the Commutant Lifting Theorem to obtain a contractive lifting B of T_2 such that $U_+ B = B U_+$, where $U_+ \in \mathcal{B}(K_+)$ is the minimal isometric dilation of T_1. Let $V_2 \in \mathcal{B}(K)$ be the minimal isometric dilation of B.

Applying the Commutant Lifting Theorem again, we get a contractive lifting V_1 of U_+ such that $V_1V_2 = V_2V_1$. Note that V_1 is a lifting of T_1 and V_2 is a lifting of T_2. It remains to see that V_1 is an isometry.

Observe that $V_1 \in \mathcal{B}(K)$ is a contractive lifting of the isometry U_+. It follows that K_+ is invariant for V_1 and $V_1|K_+ = U_+$ (note that $P_{K\ominus K_+}V_1|K_+ = 0$ because V_1 is contractive and $P_{K_+}V_1|K_+ = U_+$ is an isometry). Now, if $k \in K_+$ and $n \geq 0$,

$$\|V_1V_2^nk\| = \|V_2^nV_1k\| = \|V_1k\| = \|U_+k\| = \|k\| = \|V_2^nk\|.$$

This shows that $D_{V_1}V_2^nK_+ = 0$, where D_{V_1} is the defect operator of V_1. Since V_2 is the minimal isometric dilation of B, we have $K = \bigvee_{n\geq 0} V_2^nK_+$, so this implies that $D_{V_1} = 0$. Hence, V_1 is an isometry and the first part of the Theorem is proved.

To prove the existence of the commuting unitary dilation, one follows the same reasoning, replacing U_+ by a unitary dilation U of T_1, V_2 by the minimal unitary dilation U_2 of B, using Corollary 2.2 instead of Theorem 2.1, and denoting by U_1 the lifting obtained from U in the second application of the Commutant Lifting Theorem. Then one has to check that U_1 is a unitary. An argument similar to the one applied to V_1 suffices (one should prove that $D_{U_1} = D_{U_1^*} = 0$). $\square$

The Commutant Lifting Theorem can also be obtained as a Corollary of Andô's Theorem, and we will now show this.

Proof of Theorem 2.1 using Theorem 2.4. We can assume without loss of generality that $\|A\| = 1$. Put

$$T_1 = \begin{bmatrix} T & 0 \\ 0 & T' \end{bmatrix}, \qquad T_2 = \begin{bmatrix} 0 & A \\ 0 & 0 \end{bmatrix},$$

and note that T_1 and T_2 are commuting contractions. By Theorem 2.4, there is a commuting isometric lifting V_1, V_2 of T_1, T_2.

Since T_1 is a lifting of T, we see that V_1 is an isometric lifting of T_1. By Lemma 1.6, it follows that if we put $K_+ = \bigvee_{n\geq 0} V_1^nH$, then K_+ reduces V_1 and $U_+ = V_1|K_+$ is the minimal isometric dilation of T. Similarly, if $K'_+ = \bigvee_{n\geq 0} V_1^nH'$, then $U'_+ = V_1|K'_+$ is the minimal isometric dilation of T'.

We define $B = P_{K_+}V_2|K'_+$. Then B is a contraction. Since V_2 is a lifting of T_2, we have $P_{H\oplus H'}V_2 = T_2P_{H\oplus H'}$. Hence,

$$P_HB = P_HV_2|K'_+ = P_HT_2P_{H\oplus H'} = AP_{H'}.$$

This shows that B is a lifting of A. As a consequence, $1 = \|A\| \leq \|B\| \leq 1$, so that $\|B\| = \|A\|$.

Since K_+ reduces V_1, we have $V_1P_{K_+} = P_{K_+}V_1$. Using this, we compute

$$U_+B = V_1P_{K_+}V_2|K'_+ = P_{K_+}V_1V_2|K'_+ = P_{K_+}V_2V_1|K'_+ = P_{K_+}V_2U'_+ = BU'_+.$$

This shows that B satisfies all the needed properties. $\square$

Now we will give two independent proofs of Andô's Theorem. The main idea of both of them is similar: to construct certain liftings W_1 and W_2 of T_1 and T_2 respectively and define $V_1 = UW_1$, $V_2 = W_2U^*$, where U is some unitary that has to be chosen to make V_1 and V_2 commute. However, the construction of the isometric liftings and the unitary in both proofs is quite different.

First independent proof of Theorem 2.4. The proof proceeds in several steps.

First we find isometric liftings W_1, W_2 of T_1, T_2 which act on the same space K. Let $W_j' \in \mathcal{B}(H \oplus K_j)$ be the minimal isometric lifting of T_j, for $j = 1, 2$. We define $K = H \oplus K_1 \oplus K_2$ and the operators W_j by $W_1 = W_1' \oplus I_{K_2}$ and $W_2 = I_{K_1} \oplus W_2'$. It is clear that W_j is an isometric lifting of T_j.

Now we find isometric liftings X_1, X_2 acting on the same space and such that X_1X_2 is unitarily equivalent to X_2X_1 by an unitary operator which is the identity on H. We will use the symbol $\cong$ to denote unitary equivalence by a unitary operator which is the identity on H. Note that both W_1W_2 and W_2W_1 are isometric liftings of $T_1T_2 = T_2T_1$. Indeed,

$$P_H W_1 W_2 = T_1 P_H W_2 = T_1 T_2 P_H,$$

so that W_1W_2 is a lifting of T_1T_2, and similarly, W_2W_1 is also a lifting of T_1T_2. Let W_0 be the minimal isometric dilation of T_1T_2. Using Lemma 1.6, we see that

$$W_1W_2 \cong W_0 \oplus W_{12}, \qquad W_2W_1 \cong W_0 \oplus W_{21},$$

where W_{12} and W_{21} are isometries acting on the spaces H_{12} and H_{21} respectively. Now we define the operators X_1 and X_2 on the space $K \oplus [\bigoplus_{n\geq 1}(H_{12} \oplus H_{21})]$ by

$$X_1 = W_1 \oplus [\bigoplus_{n\geq 1}(I_{H_{12}} \oplus I_{H_{21}})], \qquad X_2 = W_2 \oplus [\bigoplus_{n\geq 1}(W_{12} \oplus W_{21})],$$

It is clear that X_1 and X_2 are isometric liftings of T_1 and T_2 respectively. Now we compute

$$X_1X_2 = W_1W_2 \oplus [\bigoplus_{n\geq 1}(W_{12} \oplus W_{21})] \cong W_0 \oplus W_{12} \oplus [\bigoplus_{n\geq 1}(W_{12} \oplus W_{21})],$$

$$X_2X_1 = W_2W_1 \oplus [\bigoplus_{n\geq 1}(W_{12} \oplus W_{21})] \cong W_0 \oplus W_{21} \oplus [\bigoplus_{n\geq 1}(W_{12} \oplus W_{21})].$$

It follows that $X_1X_2 \cong X_2X_1$.

By the previous step, there is a unitary U which is the identity on H and such that $X_1X_2 = UX_2X_1U^*$. We put $V_1 = X_1U^*$, $V_2 = UX_2$. We see that V_1, V_2 are commuting isometries. Moreover, since U is the identity on H, we have $P_HU = P_HU^* = P_H$, so we see that

$$P_H V_1 = P_H X_1 U^* = T_1 P_H U^* = T_1 P_H,$$

which shows that V_1 is a lifting of T_1. Similarly, V_2 is a lifting of T_2.

As we have remarked before the statement of Andô's Theorem, the existence of U_1, U_2 now follows from Theorem 2.5 below. □

Second independent proof of Theorem 2.4. We put $K = \bigoplus_{n\geq 0} H$, embed H into K by $h \mapsto (h, 0, \ldots)$ and define the operators $W_1, W_2 \in \mathcal{B}(K)$ by

$$W_j(h_0, h_1, h_2, \ldots) = (T_j h_0, D_{T_j} h_0, 0, h_1, h_2, \ldots), \qquad j = 1, 2.$$

These operators are isometric, but they do not commute in general.

We define the space $G = H \oplus H \oplus H \oplus H$. By the identification

$$(h_0, h_1, h2, \ldots) = (h_0, (h_1, h_2, h_3, h_4), (h_5, h_6, h_7, h_8), \ldots)$$

we have

$$K = H \oplus (\bigoplus_{n\geq 1} G).$$

Let U be an unitary operator on G to be determined later. We define the unitary operator X on K by

$$X(h_0, h_1, h_2, \ldots) = (h_0, U(h_1, \ldots, h_4), U(h_5, \ldots, h_8), \ldots).$$

We define $V_1 = XW_1$, $V_2 = W_2X^*$ and try to find a U such that V_1 and V_2 commute. Note that this will suffice, because V_1 and V_2 are isometries and we can check that they are liftings of T_1 and T_2 respectively in the same way as we did in the first proof of Theorem 2.4.

We first compute V_1V_2 and V_2V_1.

$$\begin{aligned} V_1V_2(h_0, h_1, \ldots) &= XW_1W_2X^*(h_0, h_1, \ldots) = XW_1W_2(h_0, U^*(h_1, \ldots, h_4), \ldots) \\ &= XW_1(T_2h_0, D_{T_2}h_0, 0, U^*(h_1, \ldots, h_4)) \\ &= X(T_1T_2h_0, D_{T_1}T_2h_0, 0, D_{T_2}h_0, 0, U^*(h_1, \ldots, h_4)) \\ &= (T_1T_2h_0, U(D_{T_1}T_2h_0, 0, D_{T_2}h_0, 0), (h_1, \ldots, h_4), \ldots). \end{aligned}$$

$$\begin{aligned} V_2V_1(h_0, h_1, \ldots) &= W_2W_1(h_0, h_1, \ldots) = W_2(T_1h_0, D_{T_1}h_0, 0, h_1, \ldots) \\ &= (T_2T_1h_0, D_{T_2}T_1h_0, 0, D_{T_1}h_0, 0, h_1, \ldots). \end{aligned}$$

Hence, the operator U must satisfy

$$U(D_{T_1}T_2h_0, 0, D_{T_2}h_0, 0) = (D_{T_2}T_1h_0, 0, D_{T_1}h_0, 0). \tag{2.8}$$

We see that

$$\begin{aligned} \|D_{T_1}T_2h_0\|^2 + \|D_{T_2}h_0\|^2 &= \|T_2h_0\|^2 - \|T_1T_2h_0\|^2 + \|h_0\|^2 - \|T_2h_0\|^2 \\ &= \|h_0\|^2 - \|T_1T_2h_0\|^2. \end{aligned}$$

Similarly,

$$\|D_{T_2}T_1h_0\|^2 + \|D_{T_1}h_0\|^2 = \|h_0\|^2 - \|T_2T_1h_0\|^2 = \|D_{T_1}T_2h_0\|^2 + \|D_{T_2}h_0\|^2.$$

This shows that (2.8) defines an isometry U from the linear manifold $L_1 = D_{T_1}T_2H \oplus 0 \oplus D_{T_2}H \oplus 0$ onto the linear manifold $L_2 = D_{T_2}T_1H \oplus 0 \oplus D_{T_1}H \oplus 0$. By continuity, it extends to an isometry from $\overline{L_1}$ onto $\overline{L_2}$. It remains to see that it can be extended to an isometry from G onto G. This is equivalent to the condition $\dim G \ominus \overline{L_1} = \dim G \ominus \overline{L_2}$. When the dimension of H is finite, then the dimension of G is also finite and this condition clearly holds. When H is infinite-dimensional,

$$\dim H = \dim G \geq \dim G \ominus \overline{L_j} \geq \dim H, \qquad j = 1, 2,$$

because $G \ominus \overline{L_j}$ contains the subspace $0 \oplus H \oplus 0 \oplus 0$. This shows that the dimensions of $G \ominus \overline{L_1}$ and $G \ominus \overline{L_2}$ coincide, and therefore, finishes the proof. □

2.5. Dilations of a system of isometries

The existence of a unitary dilation for a system of commuting isometries is one of the easiest positive results for tuples of operators. The key fact in obtaining the dilation is a procedure which, when applied to the system of isometries, decreases the number of non-unitary isometries in the system. By iterating this procedure several times, one finally obtains a system composed exclusively of unitaries.

Theorem 2.5. *Any system of commuting isometries $V = \{V_j\}_{j\in J}$ has an extension to a commuting system of unitaries $V = \{U_j\}_{j\in J}$.*

Proof. When the system is finite, it suffices to apply the following Lemma 2.6 a finite number of times. When the system is infinite, one could use also apply Lemma 2.6 a transfinite number of times, but we will also give a proof that does not require transfinite induction using the theory of regular dilations in Section 2.7. □

Lemma 2.6. *If $\{V\} \cup \{V_j\}_{j\in J}$ is a system of commuting isometries on H, then there is a commuting system $\{U\} \cup \{W_j\}_{j\in J}$ of operators on K such that U is unitary extension of V, each W_j is an isometric extension of V_j, and if V_j was unitary then W_j is still unitary.*

Proof. Let $U \in \mathcal{B}(K)$ be the minimal unitary extension of the isometry V, so that $K = \bigvee_{n\in\mathbb{Z}} U^n H$. If $n, m \in \mathbb{Z}$, $n \geq m$, and $h, h' \in H$, we have

$$\begin{aligned}\langle U^n V_j h, U^m V_j h'\rangle &= \langle U^{n-m} V_j h, V_j h'\rangle = \langle V^{n-m} V_j h, V_j h'\rangle = \langle V_j V^{n-m} h, V_j h'\rangle \\ &= \langle V^{n-m} h, h'\rangle = \langle U^n h, U^m h'\rangle,\end{aligned}$$

because U is an extension of V, the operators V and V_j commute, and V_j is an isometry. This shows that

$$W_j U^n h = U^n V_j h, \qquad n \in \mathbb{Z},\ h \in H,$$

defines an isometry on K which extends V_j. If V_j was unitary, then $V_j H = H$, so we see that $W_j K = K$, and therefore W_j is unitary.

It remains to see that all the operators U, W_j commute. If $n \in \mathbb{Z}$ and $h \in H$, then

$$U W_j U^n h = U^{n+1} V_j h = W_j U^{n+1} h = W_j U U^n h,$$

so that U and W_j commute. If $j, j' \in J$, then

$$W_j W_{j'} U^n h = U^n V_j V_{j'} h = U^n V_{j'} V_j h = W_{j'} W_j U^n h,$$

so that W_j and $W_{j'}$ commute. □

2.6. Von Neumann's inequality

The most basic form of von Neumann's inequality states that if $T \in \mathcal{B}(H)$ is a contraction, then

$$\|p(T)\| \leq \|p\|_\infty, \tag{2.9}$$

where $p \in \mathbb{C}[z]$ is any complex polynomial and $\|p\|_\infty = \sup_{z\in\mathbb{D}} |p(z)|$. The proof of this inequality using dilations is very easy. Let $U \in \mathcal{B}(K)$ be a unitary dilation of T and observe that $p(T) = P_H p(U)|H$ for $p \in \mathbb{C}[z]$. Therefore,

$$\|p(T)\| = \|P_H p(U)|H\| \leq \|p(U)\| \leq \|p\|_\infty.$$

Here the last inequality comes from the spectral theorem (see Appendix A.2).

There is also a version of the inequality for matrix-valued polynomials. This version will be important because of its relation with the existence of a dilation. We will denote by M_s the ring of $s \times s$ complex matrices. If $A = [a_{jk}]_{j,k=1}^s \in M_s$ is a matrix and $T \in \mathcal{B}(H)$ is an operator, we let H^s be the direct sum of s copies of H and define the operator $A \otimes T \in \mathcal{B}(H^s)$ by the matrix $[a_{jk}T]_{j,k=1}^s$. This allows us to define $p(T)$ for $p \in M_s[z]$, a matrix valued polynomial. If $p(z) = \sum_{j=1}^n A_j z^j$, we put $p(T) = \sum_{j=1}^n A_j \otimes T^j$.

In this context, von Neumann's inequality is expressed again as (2.9). Now, $p \in M_s[z]$ is a matrix-valued polynomial, $\|p(T)\|$ is the operator norm of $p(T)$ in $\mathcal{B}(H^s)$, and $\|p\|_\infty = \sup_{z\in\mathbb{D}} \|p(z)\|$, where $\|p(z)\|$ is the norm of the matrix $p(z)$ seen as an operator in $\mathcal{B}(\mathbb{C}^s)$. The proof of the matrix-valued inequality is more or less the same as the scalar-valued one. One just has to take a bit of care with the notation involving matrices. If $p \in M_s[z]$, it can also be written as $p(z) = [p_{jk}(z)]_{j,k=1}^s$, where $p_{jk} \in \mathbb{C}[z]$ are scalar-valued polynomials. Then,

$$\begin{aligned}\|p(T)\| &= \|[p_{jk}(T)]_{j,k=1}^s\| = \|[P_H p_{jk}(U)|H]_{j,k=1}^s\| \\ &\le \|[p_{jk}(U)]_{j,k=1}^s\| = \|p(U)\| \le \|p\|_\infty.\end{aligned}$$

Once again, the last inequality comes from the spectral theorem, which is easily seen to work for matrix-valued functions.

One can also ask whether the following generalization of the inequality (2.9) holds for tuples of commuting contractions $T_1, \dots, T_n \in \mathcal{B}(H)$:

$$\|p(T_1, \dots, T_n)\| \le \|p\|_\infty. \tag{2.10}$$

Here, depending on whether we are considering the scalar-valued inequality or the matrix-valued inequality, either $p \in \mathbb{C}[z_1, \dots, z_n]$ will be a scalar-valued polynomial in n variables or $p \in M_s[z_1, \dots, z_n]$ will be a matrix-valued polynomial in n variables. The norm $\|p\|_\infty$ is defined by taking the supremum of $\|p(z_1, \dots, z_n)\|$ when $(z_1, \dots, z_n)$ ranges over the polydisk $\mathbb{D}^n$ (the cartesian product of n copies of $\mathbb{D}$). This inequality (2.10) is also called von Neumann's inequality.

Whenever the tuple $T_1, \dots, T_n$ has a unitary dilation $U_1, \dots, U_n \in \mathcal{B}(K)$, von Neumann's inequality holds for this tuple of contractions. The proof is very similar to the case of a single contraction, because the key fact we need to use is $p(T_1, \dots, T_n) = P_H p(U_1, \dots, U_n)|H$, which holds because of the properties of the dilation. Thus, if $T_1, \dots, T_n$ has a unitary dilation, both the scalar and the matrix-valued inequalities hold for this tuple.

If the tuple $T_1, \dots, T_n$ does not have a unitary dilation, then (2.10) may or may not hold. Indeed, both the scalar and the matrix-valued version may fail, or it could happen that the scalar version holds but the matrix-valued version fails (the matrix-valued version is obviously stronger). Examples of this will be mentioned in Section 2.8. However, if the matrix-valued version holds, then a unitary dilation exists. This is a consequence of a theorem due to Arveson which uses the theory of C^*-algebras. In the remaining part of this section, we will present the material needed to give this result.

A complex algebra $\mathcal{A}$, together with a norm $\|\cdot\|$ on $\mathcal{A}$, is called a Banach algebra if it is a Banach space with respect to this norm, and the norm satisfies $\|xy\| \le \|x\|\|y\|$ for all $x, y \in \mathcal{A}$. If the algebra has a unit e, then it is required that $\|e\| = 1$.

A C^*-algebra $\mathcal{A}$ is a Banach algebra together with a map $x \mapsto x^*$, $x \in \mathcal{A}$, which is conjugate-linear and satisfies $x^{**} = x$ and $(xy)^* = y^*x^*$, and the so called C^* identity: $\|x^*x\| = \|x\|\|x^*\|$. The map $*$ is usually called the involution of the C^*-algebra.

A subalgebra $\mathcal{B}$ of a C^*-algebra $\mathcal{A}$ is called closed if it is closed with respect to the norm, and is called selfadjoint if $x^* \in \mathcal{B}$ for all $x \in \mathcal{B}$.

The canonical example of a C^*-algebra is $\mathcal{B}(H)$, for every Hilbert space H. Here the involution $*$ is the usual operator adjoint. Indeed, every C^*-algebra is isomorphic to a closed selfadjoint subalgebra of some $\mathcal{B}(H)$. Another important example of C^*-algebra is the space $\mathcal{C}(K)$ of continuous complex-valued functions on some compact set K, together with the norm $\|f\| = \sup_{x\in K} |f(x)|$ and the involution $f^*(x) = \overline{f(x)}$.

An algebra is called unital if it has a unit, and a map between two unital algebras is called unital if it takes the unit of the first algebra to the unit of the second one.

A representation of a C^*-algebra $\mathcal{A}$ is a map $\pi : \mathcal{A} \to \mathcal{B}(H)$ for some Hilbert space H such that π is a continuous homomorphism satisfying $\pi(x^*) = \pi(x)^*$ for every $x \in \mathcal{A}$.

If $\mathcal{A}$ is a C^*-algebra, we will denote by $M_s(\mathcal{A})$ the space of $s \times s$ matrices whose entries are elements of $\mathcal{A}$. It turns out that there is a unique norm in $M_s(\mathcal{A})$ that turns it into a C^*-algebra. The easiest way to think of this norm is to view $\mathcal{A}$ as a closed selfadjoint subalgebra of $\mathcal{B}(H)$ and note that $M_s(\mathcal{B}(H)) \cong \mathcal{B}(H^s)$.

Suppose that we are given a linear submanifold $\mathcal{B}$ of a C^*-algebra $\mathcal{A}$, another C^*-algebra $\mathcal{C}$, and a linear map $\theta : \mathcal{B} \to \mathcal{C}$. We can consider $M_s(\mathcal{B})$, the linear submanifold of $M_s(\mathcal{A})$ formed by the matrices with entries on $\mathcal{B}$. The map θ induces a linear map $\theta_s : M_s(\mathcal{B}) \to M_s(\mathcal{C})$ by $\theta_s([x_{jk}]_{j,k=1}^s) = [\theta(x_{jk})]_{j,k=1}^s$. The map θ is called a complete contraction if θ_s is a contraction for each $s = 1, 2, \dots$.

Now we are ready to formulate the theorem about C^*-algebras that we are going to use.

Theorem 2.7. *Let $\mathcal{A}$ be a unital C^*-algebra, $\mathcal{B}$ a subalgebra (not necessarily closed or selfadjoint) of $\mathcal{A}$. Let $\theta : \mathcal{B} \to \mathcal{B}(H)$ be a unital homomorphism. Then the following are equivalent:*

(i) There is a Hilbert space K containing H and a unital representation $\pi : \mathcal{A} \to \mathcal{B}(K)$ such that

$$\theta(x) = P_H \pi(x)|H, \qquad \forall x \in \mathcal{A}.$$

(ii) The map θ is a complete contraction.

This Theorem is part of [AM02, Corollary 14.16], and this Corollary 14.16 is a direct consequence of the theorem of Arveson on the extension of completely positive maps.

Now we will use the Theorem to prove the following result.

Theorem 2.8. *Suppose that $T_1, \dots, T_n \in \mathcal{B}(H)$ is a tuple of commuting contractions which satisfies the von Neumann inequality*

$$\|p(T_1, \dots, T_n)\| \le \|p\|_\infty,$$

for every matrix-valued polynomial $p \in M_s[z_1, \dots, z_n]$, $s = 1, 2, \dots$. Then the tuple has a commuting unitary dilation $U_1, \dots, U_n \in \mathcal{B}(K)$.

Proof. We put $\mathcal{A} = \mathcal{C}(\mathbb{T}^n)$ the C^*-algebra of continuous functions on the n-torus $\mathbb{T}^n$, which is the product of n copies of $\mathbb{T} = \partial\mathbb{D}$. Consider $\mathcal{B} = \mathbb{C}[z_1, \dots, z_n]$ the algebra of polynomials in n variables. By the maximum modulus principle, we have

$$\|p\|_\infty = \sup_{z\in\mathbb{D}^n} |p(z)| = \sup_{z\in\mathbb{T}^n} |p(z)| = \|p\|_{\mathcal{A}}, \qquad p \in \mathcal{B}.$$

This shows that $\mathcal{B}$ is indeed a subalgebra of $\mathcal{A}$ and that the induced norm in $\mathcal{B}$ is just the norm $\|\cdot\|_\infty$.

Define the map $\theta : \mathcal{B} \to \mathcal{B}(H)$ by $\theta(p) = p(T_1, \dots, T_n)$. Then, the map

$$\theta_s : M_s(\mathbb{C}[z_1, \dots, z_n]) \to M_s(\mathcal{B}(H)),$$

if one takes into account the isomorphisms $M_s(\mathbb{C}[z_1, \dots, z_n]) \cong M_s[z_1, \dots, z_n]$ and $M_s(\mathcal{B}(H)) \cong \mathcal{B}(H^s)$, can be regarded as the map $\theta_s(p) = p(T_1, \dots, T_n)$, where $p \in M_s[z_1, \dots, z_n]$ is a matrix-valued polynomial. Therefore, the matrix-valued von Neumann's inequality is equivalent to θ being a complete contraction.

Since θ is a complete contraction, we get a representation π of $\mathcal{A}$ into $\mathcal{B}(K)$ as in Theorem 2.7. We define $U_j = \pi(z_j)$. Note that $z_j^* z_j = z_j z_j^* = 1$ in $\mathcal{A}$, because in $\mathbb{T}^n$ we have $|z_j| = 1$. Since π is unital,

$$U_j^* U_j = \pi(z_j)^* \pi(z_j) = \pi(z_j^* z_j) = \pi(1) = I.$$

Similarly, $U_j U_j^* = I$. This shows that U_j are unitary. Also,

$$U_j U_k = \pi(z_j)\pi(z_k) = \pi(z_j z_k) = \pi(z_k z_j) = \pi(z_k)\pi(z_j) = U_k U_j,$$

so that U_j commute.

Finally, we have to check that $U_1, \dots, U_n$ is a dilation of $T_1, \dots, T_n$, but this follows from the property $\theta(x) = P_H \pi(x)|H$ by letting x be a monomial $x = z_1^{k_1} \cdots z_n^{k_n}$. □

2.7. Regular dilations

Regular dilations are a special kind of unitary dilations which are more well behaved. For instance, the conditions under which a system of contractions has a regular dilation is completely understood. Also, there is a notion of minimality for regular dilations, and the minimal regular dilation is unique.

To give the definition of a regular dilation, we first need a bit of notation. Let $T = \{T_j\}_{j \in J} \subset \mathcal{B}(H)$ be a system of commuting contractions and $U = \{U_j\}_{j \in J} \subset \mathcal{B}(K)$, with $K \supset H$, a system of commuting unitaries. We denote by $\mathbb{Z}_c^J$ the set of functions taking J to $\mathbb{Z}$ which have finite support (i.e., which vanish except on a finite number of elements of J). Given a function $n \in \mathbb{Z}_c^J$, we say that $n \geq 0$ if $n(j) \geq 0$ for all $j \in J$.

We define, for $n \in \mathbb{Z}_c^J$,

$$U^n = \prod_{j \in J} U_j^{n(j)}.$$

This product is well defined because $n(j) = 0$ except for a finite number of indices $j \in J$. If $n \geq 0$, we can also define

$$T^n = \prod_{j \in J} T_j^{n(j)}.$$

We also define the functions $n^+, n^- \in \mathbb{Z}_c^J$ by $n^+(j) = \max\{n(j), 0\}$, and $n^-(j) = \max\{-n(j), 0\}$.

We say that U is a regular dilation of T if

$$(T^{n^-})^* T^{n^+} = P_H U^n | H, \qquad \forall n \in \mathbb{Z}_c^J. \tag{2.11}$$

Note that if we restrict (2.11) to only those $n \in \mathbb{Z}_c^J$ such that $n \geq 0$, we obtain the usual condition of a dilation. Hence, a regular dilation is a stronger concept.

To formulate the theorem about the existence of regular dilations, we need just another bit of notation. If K is a finite subset of J, we denote by χ_K the characteristic or indicator function of K. This means that $\chi_K(j) = 1$ if $j \in K$, and $\chi_K(j) = 0$ otherwise. Note that $\chi_K \in \mathbb{Z}_c^J$. We denote by $|K|$ the number of elements of K.

Theorem 2.9. *Let $T = \{T_j\}_{j\in J}$ be a system of commuting contractions on some Hilbert space H. The system T has a regular unitary dilation $U = \{U_j\}_{j\in J}$ in some larger Hilbert space $K \supset H$ if and only if*

$$S(J_0) = \sum_{K\subset J_0} (-1)^{|K|} (T^{\chi_K})^* (T^{\chi_K}) \geq 0 \tag{2.12}$$

for every finite subset $J_0 \subset J$.

Moreover, the dilation U can be taken to be minimal, in the sense that

$$K = \bigvee_{n\in\mathbb{Z}_c^J} U^n H. \tag{2.13}$$

The minimal regular unitary dilation is unique up to isomorphism.

The proof of the Theorem is a bit involved and uses the theory of positive functions defined on groups and some lengthy computations involving combinatorics of subsets. We refer the reader to [SNFBK10, Section I.9] for a good exposition of the topic.

However, the part about minimality and uniqueness is very easy to obtain. Note that if U is any regular unitary dilation of T and one defines $\widehat{K} = \bigvee_{n\in\mathbb{Z}_c^J} U^n H$, then $\widehat{K}$ reduces all the operators U_j and the system $\{U_j|\widehat{K}\}_{j\in J}$ is also a regular dilation of T. If U is minimal, (2.11) implies that if $h, h' \in H$ and $n, n' \in \mathbb{Z}_c^J$, then the inner product

$$\langle U^n h, U^{n'} h'\rangle = \langle U^{n-n'} h, h'\rangle = \langle (T^{(n-n')^-})^* T^{(n-n')^+} h, h'\rangle$$

is completely determined by the system T. This implies that any two minimal regular dilations must be isomorphic.

Now we would like to investigate sufficient conditions under which the condition (2.12) appearing in the Theorem holds. First we show that if the system T contains an isometry T_{j_0}, then $S(J_0) = 0$ for any finite subset J_0 containing j_0.

To see this, let J_0 be any such subset and take a $K \subset J_0$. Then K can be partitioned as $K = K_0 \cup K_1$, where $K_0 \subset \{j_0\}$ and $K_1 \subset K \setminus \{j_0\}$. Also, any such an election of K_0 and K_1 gives a subset K by $K = K_0 \cup K_1$. We have

$$(T^{\chi_K})^*(T^{\chi_K}) = (T^{\chi_{K_1}})^*(T^{\chi_{K_0}})^* T^{\chi_{K_0}} T^{\chi_{K_1}} = (T^{\chi_{K_1}})^* T^{\chi_{K_1}},$$

because $T^{\chi_{K_0}} = I$ whenever $K_0 = \emptyset$, and $T^{\chi_{K_0}}$ is the isometry T_{j_0} whenever $K_0 = \{j_0\}$. Using this decomposition to split the sum in (2.12), we get

$$S(J_0) = \sum_{K_1\subset J_0\setminus\{j_0\}} (-1)^{|K_1|} \left[\sum_{K_0\subset\{j_0\}} (-1)^{|K_0|} \right] (T^{\chi_{K_1}})^* T^{\chi_{K_1}}.$$

It suffices to observe that the expression in square brackets is 0.

We say that two operators A and B doubly commute if A commutes with both B and B^* (which implies that B commutes with A and A^*). Assume that J_d is a subset of J_0 such that T_j and T_k doubly commute whenever $j \in J_d$ and $k \in J_0$, $j \neq k$. Put $J_0' = J_0 \setminus J_d$. Then, if $S(J_0') \geq 0$, we also have $S(J_0) \geq 0$.

The proof of this fact just involves computing

$$\begin{aligned} S(J_0) &= \sum_{K_d\subset J_d}\sum_{K'\subset J_0'}(-1)^{|K_d|+|K'|}(T^{\chi_{K_d}})^*(T^{\chi_{K'}})^*T^{\chi_{K'}}T^{\chi_{K_d}} \\ &= \sum_{K'\subset J_0'}(-1)^{|K'|}(T^{\chi_{K'}})^*T^{\chi_{K'}}\sum_{K_d\subset J_d}(-1)^{|K_d|}\prod_{j\in K_d}T_j^*T_j \\ &= \sum_{K'\subset J_0'}(-1)^{|K'|}(T^{\chi_{K'}})^*T^{\chi_{K'}}\prod_{j\in J_d}(I-T_j^*T_j) \\ &= S(J_0')\prod_{j\in J_d}(I-T_j^*T_j). \end{aligned}$$

Here we have used that the operators with indices in J_d doubly commute with the rest to rearrange them. Now, since each factor $I-T_j^*T_j$, $j\in J_d$, commutes with each other and with $S(J_0')$, we see that if $S(J_0')\geq 0$, then $S(J_0)\geq 0$ also.

Another important fact is that if $\sum_{j\in J_0}\|T_j\|^2\leq 1$, then $S(J_0)\geq 0$. To prove this, we will put $J_0=\{j_1,\dots,j_r\}$ and write T_n instead of T_{j_n} for brevity. We define, for $0\leq k\leq r$ and $h\in H$, the quantity

$$a_k(h)=\sum_{\substack{K\subset J_0\\|K|=k}}\|T^{\chi_K}h\|^2.$$

For $1\leq k\leq r$, we have

$$\begin{aligned} a_k(h) &= \sum_{1\leq n_1<\ldots<n_k\leq r}\|T_{n_1}\cdots T_{n_k}h\|^2 \leq \sum_{1\leq n_1<\ldots<n_k\leq r}\|T_{n_k}\|^2\|T_{n_1}\cdots T_{n_{k-1}}h\|^2 \\ &= \sum_{1\leq n_1<\ldots<n_{k-1}\leq r}\|T_{n_1}\cdots T_{n_{k-1}}h\|^2\sum_{n_{k-1}<n_k\leq r}\|T_{n_k}\|^2 \\ &\leq \sum_{1\leq n_1<\ldots<n_{k-1}\leq r}\|T_{n_1}\cdots T_{n_{k-1}}h\|^2 = a_{k-1}(h). \end{aligned}$$

Therefore,

$$\begin{aligned} \langle S(J_0)h,h\rangle &= \sum_{K\subset J_0}(-1)^{|K|}\|T^{\chi_K}h\|^2=\sum_{k=0}^{r}(-1)^k a_k(h)\geq a_0(h)-a_1(h) \\ &= \|h\|^2-\sum_{n=1}^{r}\|T_nh\|^2\geq\left(1-\sum_{n=1}^{r}\|T_n\|^2\right)\|h\|^2\geq 0. \end{aligned}$$

To sum up, we have proved the three following facts about the positivity of $S(J_0)$:

1. If T_{j_0} is an isometry and $j_0\in J_0$, then $S(J_0)=0$.

2. Assume that $J_d\subset J_0$ and T_j doubly commutes with T_k whenever $j\in J_d$ and $k\in J_0$, $j\neq k$. If $S(J_0\setminus J_d)\geq 0$, then $S(J_0)\geq 0$ also.

3. If $\sum_{j\in J_0}\|T_j\|^2\leq 1$, then $S(J_0)\geq 0$.

The following Theorem is a straightforward consequence of these three facts.

Theorem 2.10. *Let $\mathcal{T} = \{T_j\}_{j\in J}$ be a system of commuting contractions on some Hilbert space H. Delete the isometries from $\mathcal{T}$ and call the remaining system $\mathcal{T}_1$. Delete from $\mathcal{T}_1$ those operators which doubly commute with every other operator, and call the remaining system $\mathcal{T}_2$. If $\mathcal{T}_2$ has a regular unitary dilation, then so does $\mathcal{T}$.*

In particular, $\mathcal{T}$ has a regular unitary dilation in any of the three following particular cases:

(a) $\mathcal{T}$ consists only of isometries.

(b) $\mathcal{T}$ consists only of doubly commuting contractions.

(c) $\mathcal{T}$ is countable and satisfies $\sum_{j\in J} \|T_j\|^2 \leq 1$.

In the case (a) of this theorem, we obtain a unitary dilation $\{U_j\}$ of a system of commuting isometries $\{V_j\}$. It is easy to check that such a dilation must be indeed an extension (the proof is the same as for the case of only one operator). Therefore, we have just obtained another proof of Theorem 2.5.

Continuing with the topic of unitary extensions of a system of isometries, the theory of regular dilations can also be used to obtain the following result regarding their uniqueness.

Proposition 2.11. *Let $V = \{V_j\}_{j\in J} \subset \mathcal{B}(H)$ be a system of commuting isometries. Every unitary dilation $U = \{U_j\}_{j\in J} \subset \mathcal{B}(K)$ of such a system V is also a regular dilation. In particular, if U is minimal, in the sense that it satisfies (2.13), then U is unique up to isomorphism.*

Proof. As we have already mentioned, U is indeed an extension of V, which means that $V_j = U_j|H$. Therefore, U_j^* is a lifting of V_j^*, so that $P_H U_j^* = V_j^* P_H$. Using this, we see that, for indices $j_1, \dots, j_n \in J$ and positive numbers $k_1, \dots, k_n, l_1, \dots, l_n \geq 0$, we have

$$P_H U_{j_1}^{*k_1} \cdots U_{j_n}^{*k_n} U_{j_1}^{l_1} \cdots U_{j_n}^{l_n} | H = V_{j_1}^{*k_1} \cdots V_{j_n}^{*k_n} V_{j_1}^{l_1} \cdots V_{j_n}^{l_n}.$$

This implies that the dilation is regular. □

2.8. Non-existence of dilations

In this section we will give several counterexamples that show that, for three or more commuting contractions, a unitary dilation does not always exist. The usual way to give such a counterexample is using operators acting on a finite dimensional space (hence, represented as matrices). Then there are several ways to prove that their dilation does not exist: to assume that a unitary dilation exists and get to a contradiction, to show that they do not satisfy the scalar-valued von Neumann's inequality for a given polynomial, or to show that they do not satisfy the matrix-valued von Neumann's inequality for a given polynomial.

The first counterexample was given by Parrott in [Par70]. It was essentially as follows. Let H_0 be a Hilbert space and choose three unitaries $A_1, A_2, A_3 \in \mathcal{B}(H_0)$ such that

$$A_1 A_2^* A_3 \neq A_3 A_2^* A_1. \tag{2.14}$$

For instance, take $A_2 = I$ and choose A_1 and A_3 to be two non-commuting unitaries.

Now we put $H = H_0 \oplus H_0$ and define the contractions $T_1, T_2, T_3 \in \mathcal{B}(H)$ by

$$T_j = \begin{bmatrix} 0 & 0 \\ A_j & 0 \end{bmatrix}, \qquad j = 1, 2, 3.$$

We see that if $j \neq k$, then $T_jT_k = T_kT_j = 0$, so (T_1, T_2, T_3) is a tuple of commuting contractions on H. Assume that there are commuting unitaries U_1, U_2, U_3 acting on a larger Hilbert space $K \supset H$ and such that $T_j = P_H U_j|H$ (this will certainly happen if (T_1, T_2, T_3) has a unitary dilation).

We compute

$$P_H U_j(h, 0) = T_j(h, 0) = (0, A_j h), \qquad h \in H_0.$$

Since $\|U_j(h, 0)\| = \|(h, 0)\|$ and $\|(0, A_j h)\| = \|h\|$, we see that $U_j(h, 0) = (0, A_j h)$. Now we can compute

$$\begin{aligned} U_l U_k^* U_j(h, 0) &= U_l U_k^*(0, A_j h) = U_l U_k^*(0, A_k A_k^* A_j h) = U_l U_k^* U_k(A_k^* A_j h, 0) \\ &= U_l(A_k^* A_j h, 0) = (0, A_l A_k^* A_j h). \end{aligned}$$

Since U_j are commuting unitaries, we have $U_1 U_2^* U_3 = U_3 U_2^* U_1$, which implies $A_1 A_2^* A_3 = A_3 A_2^* A_1$. This contradicts (2.14).

Later, Kaijser and Varopoulos in an addendum to the paper [Var74], and Crabb and Davie in [CD75], gave independently examples of commuting contractions on a finite Hilbert space for which the scalar-valued von Neumann's inequality fails. It is worthy to mention that the operators constructed in the counterexample of Parrott do satisfy the scalar von Neumann's inequality.

The example of Kaijser and Varopoulos is given in a five-dimensional Hilbert space H with orthonormal basis $\{e, f_1, f_2, f_3, h\}$. The commuting contractions $T_1, T_2, T_3 \in \mathcal{B}(H)$ are defined by

$$T_j e = f_j, \qquad T_j f_k = a_{jk} h, \qquad T_j h = 0,$$

where $a_{jj} = 1/\sqrt{3}$ and $a_{jk} = -1/\sqrt{3}$ if $j \neq k$. Then, they consider the polynomial

$$p(z_1, z_2, z_3) = \sum_{j,k=1}^{3} a_{jk} z_j z_k$$

and show that $\|p(T_1, T_2, T_3)\| \geq 3$ but $\|p\|_\infty = 5/\sqrt{3}$.

The example of Crabb and Davie is given in an eight-dimensional Hilbert space H with orthonormal basis $\{e, f_1, f_2, f_3, g_1, g_2, g_3, h\}$. The commuting contraction $T_1, T_2, T_3 \in \mathcal{B}(H)$ are defined by

$$T_j e = f_j, \qquad T_j f_j = -g_j,$$

for $j \neq k$,

$$T_j f_k = g_l,$$

where $l \neq j$ and $l \neq k$, and

$$T_j g_k = \delta_{jk} h, \qquad T_j h = 0.$$

Then, they put

$$p(z_1, z_2, z_3) = z_1 z_2 z_3 - z_1^3 - z_2^3 - z_3^3$$

and show that $\|p(T_1, T_2, T_3)\| \geq 4$ but $\|p\|_\infty < 4$.

These examples use contractions which are not diagonalizable. In [LS94], Lotto and Steger construct an example consisting of diagonalizable contractions by doing a perturbation of the example of Kaijser and Varopoulos.

Recently, Choi and Davidson have given in [CD13] four 3×3 commuting contractive matrices which fail to satisfy the matrix-valued von Neumann's inequality, and therefore, they do not have a unitary dilation. However, the matrices satisfy the scalar-valued von Neumann's inequality. The question of whether one can give a similar example with only three 3×3 matrices remains open. It is also open whether the scalar-valued von Neumann's inequality holds for 3×3 contractions.

2.9. A dilation existence result for diagonalizable contractions

The purpose of this section is to prove the following theorem.

Theorem 2.12 (Lotto). *Assume that $T_1, \ldots, T_n$ is a tuple of commuting, diagonalizable contractions on a finite-dimensional Hilbert space H and that no nontrivial subspace of H reduces all the contractions T_j. If there exists a diagonalizable contraction $X \in \mathcal{B}(H)$ which commutes with every T_j and such that $I - X^*X$ has rank 1, then the matrix-valued von Neumann's inequality holds for $T_1, \ldots, T_n$.*

This theorem first appeared in [Lot94] for the scalar-valued von Neumann's inequality. The statement involving the matrix-valued von Neumann's inequality does not seem to appear in the literature, but it is obtained by literally the same proof as the scalar-valued case. The importance of the matrix-valued case is that, together with Theorem 2.8, it shows that if $T_1, \ldots, T_n$ satisfies the hypothesis of the Theorem, then $T_1, \ldots, T_n$ has a unitary dilation.

The assumption that no notrivial subspace reduces all the contractions T_j is not restrictive. Indeed, if $M \subset H$ is a subspace which reduces every T_j, then von Neumann's inequality for $T_1, \ldots, T_n$ will hold if and only if it holds for both the tuples $T_1|M, \ldots, T_n|M$ and $T_1|H \ominus M, \ldots, T_n|H \ominus M$.

If T_j satisfy the hypothesis of the theorem, then they can be simultaneously diagonalized in a basis $\{v_1, \ldots, v_N\}$ of eigenvectors. For $w \in \mathbb{C}^N$, we will denote by D_w the operator defined by $D_w v_j = w_j v_j$. We see that each operator T_j is of the form D_w for an appropriate choice of w. First we need to compute when such an operator D_w is a contraction.

Lemma 2.13. *The operator D_w is a contraction if and only if the matrix*

$$\left[(1 - w_j \overline{w_k})\langle v_j, v_k\rangle\right]_{j,k=1}^{N} \tag{2.15}$$

is non-negative.

Proof. It suffices to observe that (2.15) is the matrix of the operator $I - D_w^* D_w$ in the (non-orthonormal) basis $\{v_1, \ldots, v_n\}$, so that (2.15) will be non-negative if and only if $I - D_w^* D_w \geq 0$.

To check that (2.15) is the matrix of $I - D_w^* D_w$, we compute

$$\langle (I - D_w^* D_w) v_j, v_k\rangle = \langle v_j, v_k\rangle - \langle D_w v_j, D_w v_k\rangle = (1 - w_j \overline{w_k})\langle v_j, v_k\rangle.$$

□

We need another technical lemma where we use the hypothesis concerning nontrivial reducing subspaces.

Lemma 2.14. *If D_w is a contraction, then either D_w is a scalar multiple of the identity or $|w_j| < 1$ for $j = 1, \ldots, N$.*

Proof. It is clear that $|w_j| \leq 1$. Assume that $|w_{j_0}| = 1$ for some j_0. By the preceding lemma, the matrix (2.15) is non-negative, therefore any submatrix of that matrix is also non-negative.

Since no nontrivial subspace of H reduces all the contractions T_j, there must be an eigenvector v_{k_0} such that $\langle v_{j_0}, v_{k_0}\rangle \neq 0$. Otherwise, the subspace generated by v_{j_0} would reduce all the contractions T_j.

The submatrix of (2.15) formed by the j_0-th and k_0-th rows and columns is

$$\begin{bmatrix} 0 & (1 - w_{j_0}\overline{w_{k_0}})\langle v_{j_0}, v_{k_0}\rangle \\ (1 - w_{k_0}\overline{w_{j_0}})\langle v_{k_0}, v_{j_0}\rangle & (1 - |w_{k_0}|^2)\|v_{k_0}\|^2. \end{bmatrix}$$

The determinant of this matrix is $-|(1-w_{j_0}\overline{w_{k_0}})\langle v_{j_0}, v_{k_0}\rangle|^2$, and must be non-negative because the matrix is non-negative. Hence, we get $1-w_{j_0}\overline{w_{k_0}}=0$, which implies $w_{k_0}=w_{j_0}$.

Iterating this argument for w_{k_0} instead of w_{j_0}, we see that $w_k = w_{j_0}$ for any k such that v_k can be linked to v_{j_0} by a chain of eigenvectors $v_{j_0}, v_{j_1}, \ldots, v_{j_r} = w_k$ such that $\langle v_{j_l}, v_{j_{l+1}}\rangle \neq 0$. However, the set of all such eigenvectors v_k must be $\{v_1, \ldots, v_n\}$, because their linear span is a nonempty subspace which reduces all the contractions T_j. □

Now we consider the contraction X in the statement of the theorem. Since it commutes with all T_j, it is also of the form D_z for some $z \in \mathbb{C}^N$. Moreover, since $I - D_z^* D_z$ has rank 1, then D_z cannot be a scalar multiple of the identity. By the preceding Lemma, we see that $|z_j| < 1$ for all j. It follows that D_z is a C_{00} contraction (note that, for instance, its spectral radius is less than 1).

Now we can apply the Sz.-Nagy–Foiaş theory to D_z. First note that the defect space $\mathfrak{D}_{D_z}$ is one-dimensional. By applying Theorem 1.10 to D_z^*, we see that its characteristic function $\Theta_{D_z^*}$ is an inner function in $H^\infty(\mathfrak{D}_{D_z^*}, \mathfrak{D}_{D_z})$. This implies that $\mathfrak{D}_{D_z^*}$ is also one-dimensional (recall that $\Theta(e^{it})$ is an isometry from $\mathfrak{D}_{D_z^*}$ into $\mathfrak{D}_{D_z}$ a.e. in $\mathbb{T}$). Hence, the Sz.-Nagy–Foias model for D_z is constructed in the scalar-valued $H^2(\mathbb{D})$ space.

We apply Theorem 1.10 to D_z to see that it is unitarily equivalent to the compression operator $P_K M_z|K$, where K is some subspace of $H^2(\mathbb{D})$, invariant for M_z^*. The remark now is that this gives the operator D_z an $H^\infty(\mathbb{D})$ functional calculus, which can be defined easily in terms of the model operator. If $f \in H^\infty(\mathbb{D})$, we can define

$$f(P_K M_z|K) = P_K f(M_z)|K = P_K M_f|K.$$

This means that for $g \in K$, we put $f(P_K M_z|M)g = P_K(fg)$. Then, it is clear that $P_K M_z|K$ satisfies von Neumann's inequality for $f \in H^\infty(\mathbb{D})$:

$$\|f(P_K M_z|M)\| \leq \|f\|_{H^\infty(\mathbb{D})}.$$

A similar construction can be done for $f \in M_s(H^\infty(\mathbb{D}))$, i.e., when f is an $s \times s$ matrix with $H^\infty(\mathbb{D})$ entries, and the corresponding von Neumann's inequality also holds.

This allows us to define an $M_s(H^\infty(\mathbb{D}))$ functional calculus for the operator D_z such that it satisfies von Neumann's inequality:

$$\|f(D_z)\| \leq \|f\|_\infty, \qquad f \in M_s(H^\infty(\mathbb{D})). \tag{2.16}$$

Here $\|f\|_\infty = \sup_{z\in\mathbb{D}} \|f(z)\|$. It is easy to check that when f is scalar-valued, $f(D_z) = D_{f(z)}$, where $f(z) = (f(z_1), \ldots, f(z_n))$.

Now we give the main Theorem which allows one to deduce Theorem 2.12 from the functional calculus of D_z.

Theorem 2.15. *Assume that D_z is a contraction such that $I - D_z^* D_z$ has rank 1. Then the following are equivalent for a vector $w \in \mathbb{C}^N$:*

(i) D_w is a contraction.

(ii) The matrix

$$\left[\frac{1-w_j\overline{w_k}}{1-z_j\overline{z_k}}\right]_{j,k=1}^N$$

is non-negative.

(iii) There is a function $f \in H^\infty(\mathbb{D})$ with $\|f\|_{H^\infty(\mathbb{D})} \le 1$ and $f(z_j) = w_j$.

(iv) There is a function $f \in H^\infty(\mathbb{D})$ with $\|f\|_{H^\infty(\mathbb{D})} \le 1$ and $f(D_z) = D_w$.

Proof. Assume that (i) holds. By Lemma 2.13 the matrix $[(1 - z_j\overline{z_k})\langle v_j, v_k\rangle]_{j,k=1}^N$ is non-negative and has rank 1 (because it is the matrix of $I - D_z^* D_z$). Hence, there is a vector $c \in \mathbb{C}^N$ such that

$$(1 - z_j\overline{z_k})\langle v_j, v_k\rangle = c_j\overline{c_k}.$$

This shows that no c_j can vanish, because $|z_j| < 1$ by Lemma 2.14.

Now, if C is the diagonal matrix having the vector c as its diagonal, we see that

$$\left[(1 - w_j\overline{w_k})\langle v_j, v_k\rangle\right]_{j,k=1}^N = C\left[\frac{1 - w_j\overline{w_k}}{1 - z_j\overline{z_k}}\right]_{j,k=1}^N C^*.$$

Since the matrix on the left hand side is non-negative by Lemma 2.13 and C is invertible, this proves (ii).

If (ii) holds, then we get (iii) by the Nevanlinna-Pick interpolation Theorem (see Section 2.2). That (iii) implies (iv) is just a consequence of the functional calculus for D_z. Finally, (iv) implies (i) by von Neumann's inequality (2.16). $\square$

Proof of Theorem 2.12. We see that under the hypothesis of the Theorem, there are functions $f_1, \dots, f_n \in H^\infty(\mathbb{D})$ with $\|f_j\|_{H^\infty(\mathbb{D})} \le 1$ and $T_j = f_j(D_z)$. If $p \in M_s[z_1, \dots, z_n]$ is a matrix-valued polynomial,

$$p(T_1, \dots, T_n) = [p \circ (f_1, \dots, f_n)](D_z).$$

We just have to note that $[p \circ (f_1, \dots, f_n)]$ is in $M_s(H^\infty(\mathbb{D}))$ and $\|p \circ (f_1, \dots, f_n)\|_\infty \le \|p\|_\infty$, so it suffices to use (2.16). $\square$

As concluding remarks, Lotto also proves that the hypothesis of Theorem 2.12 are always satisfied for diagonalizable 2×2 contractions, and gives a practical way of checking them for diagonalizable 3×3 contractions. Since every tuple of 2×2 or 3×3 matrices can be perturbed to a commuting tuple of diagonalizable matrices (see [Lot94, Lemma 10]), this implies that a unitary dilation always exists for tuples of commuting 2×2 contractions. The case of 2×2 contractions was previously studied by Drury in [Dru83], where he proved that they satisfied the scalar-valued von Neumann's inequality.

3. Livšic-Vinnikov theory

This chapter contains an introductory exposition of the Livšic-Vinnikov theory of commuting non-selfadjoint operators. As we have commented in the Preface, it is a different approach to the theory of commuting operators. Its main idea is, roughly speaking, to embed the tuple of operators in a structure depending on some auxiliary matrices which characterize, in a certain sense, the behaviour of the operators. These matrices can be used to assign an algebraic curve to the tuple, thus giving a connection with Algebraic Geometry.

The exposition of this chapter is organized in the following way. First we treat the theory of colligations of a single operator. Then, we discuss its generalization to several operators, but some difficulties arise. This motivates the definition of vessel, which allows one to solve these difficulties. Finally, we treat the topic of vessels of two operators, which have a richer theory.

The main monograph about this theory is the book by Livšic, Kravitsky, Marcus and Vinnikov [LKMV95]. Some interesting expository papers are [Vin98, BV03].

3.1. Colligations of a single operator

Definition (Colligation of a single operator)**.** Given H a Hilbert space, E a finite dimensional Hilbert space, operators $\Phi \in \mathcal{B}(H, E)$ and $A \in \mathcal{B}(H)$, and a selfadjoint operator $\sigma \in \mathcal{B}(E)$, the tuple $\mathcal{C} = (A; H, \Phi, E; \sigma)$ is called a colligation if the relation

$$\frac{1}{i}(A - A^*) = \Phi^* \sigma \Phi \tag{3.1}$$

holds.

The space H is called the inner space, the space E is called the outer space, the operator Φ is said to be the window of the colligation, and σ is called the rate. From (3.1) we see that the operator A must have an imaginary part $\operatorname{Im} A = (A - A^*)/2i$ of finite rank and that the operator σ somehow models $\operatorname{Im} A$. The colligation is said to be *strict* if $\Phi H = E$ and $\ker \sigma = 0$. This condition has the meaning that the outer space E is not too large. Although it seems quite natural, it is often too restrictive. For instance, the projection of a strict colligation (to be defined below) need not be strict, or a strict colligation with one-dimensional inner space must have $\dim E = 1$.

Every operator A with finite-dimensional imaginary part can be embedded in a strict colligation by putting $E = (A - A^*)H$, $\Phi = P_E$, the orthogonal projection onto E, and $\sigma = (A - A^*)/i|E$. It turns out that, in some aspects, colligations are a more useful framework than the operators alone to study this class of non-selfadjoint operators. Hence, embedding a given operator in a colligation is usually the first step in its study.

Another common way of embedding an operator A in a colligation is to define E in the same way as above and then put $\Phi = (|(A - A^*)/i||E)^{\frac{1}{2}}$, $\sigma = \operatorname{sign}(A - A^*)/i|E$. Note that here one is using the Borel (or continuous) functional calculus for the selfadjoint operator $(A - A^*)/i$ (see Appendix A.2). This embedding has the advantage that the operator σ is rather simple. However, it cannot be generalized to tuples of several operators.

The concepts of decomposition and coupling allow us to respectively break a colligation into two smaller colligations and to join two colligations into a larger one. By coupling colligations several times, one can build up a colligation from simple colligations, such as, for instance, colligations having one-dimensional inner space.

Let H'' be a subspace invariant for A and put $H' = H \ominus H''$. Then we define $A' = P_{H'}A|H'$, $A'' = A|H''$, $\Phi' = \Phi|H$ and $\Phi'' = \Phi|H''$. It is now easy to check that the tuples $\mathcal{C}' = (A'; H', \Phi', E; \sigma)$ and $\mathcal{C}'' = (A''; H'', \Phi'', E; \sigma)$ are colligations. These colligations are called the decomposition of $\mathcal{C}$ with respect to H'', and $\mathcal{C}'$, $\mathcal{C}''$ are called the projections of $\mathcal{C}$ onto H' and H'' respectively. Multiplying (3.1) by $P_{H''}$ on the left, restricting to H', and using $P_{H''}A^*|H' = 0$ (which is true because H' is invariant for A^*), we get

$$\frac{1}{i}P_{H''}A|H' = P_{H''}\Phi^*\sigma\Phi|H' = \Phi''^*\sigma\Phi'.$$

This means that A has the form

$$A = \begin{bmatrix} A' & 0 \\ i\Phi''^*\sigma\Phi' & A'' \end{bmatrix} \tag{3.2}$$

with respect to the decomposition $H = H' \oplus H''$.

Conversely, given colligations $\mathcal{C}' = (A'; H', \Phi', E; \sigma)$ and $\mathcal{C}'' = (A''; H'', \Phi'', E; \sigma)$ with the same rate σ, one can form the coupling $\mathcal{C}' \vee \mathcal{C}'' = (A; H, \Phi, E; \sigma)$ by puting $H = H' \oplus H''$, $\Phi = [\Phi', \Phi'']$, and defining A by (3.2). Then it is easy to check that the copuling $\mathcal{C}' \vee \mathcal{C}''$ is a colligation. Moreover, the space H'' is invariant for A and the decomposition of the coupling with respect to H'' is precisely the colligations $\mathcal{C}'$ and $\mathcal{C}''$.

A colligation has a system theoretical interpretation that allows us to use techniques from the control theory and the theory of partial differential equations to study operators. Consider the dynamical system

$$\begin{aligned} i\frac{df}{dt} + Af &= \Phi^*\sigma u, \\ v &= u - i\Phi f, \end{aligned} \tag{3.3}$$

where t ranges over an interval in $\mathbb{R}$ (finite or infinite). Here $u(t) \in E$ is the input of the system, $f(t) \in H$ is the state, and $v(t) \in E$ is the output of the system.

This system satisfies the law of conservation of energy

$$\frac{d}{dt}\langle f, f\rangle = \langle \sigma u, u\rangle - \langle \sigma v, v\rangle. \tag{3.4}$$

We understand the (indefinite) quadratic form given by σ as a measure of the energy at the input and output, so that (3.4) just says that the variation of the internal energy of the system $\|f\|^2$ amounts just to the energy added at the input and the energy extracted at the output.

To prove (3.4), we first use (3.3) to obtain

$$\frac{d}{dt}\langle f, f\rangle = 2\,\mathrm{Re}\langle \frac{df}{dt}, f\rangle = 2\,\mathrm{Re}\langle iAf - i\Phi^*\sigma u, f\rangle.$$

Now we use $iA = iA^* - \Phi^*\sigma\Phi$, which comes from (3.1) and $i\Phi f = u - v$, which comes from (3.3) to get

$$2\,\mathrm{Re}\langle iAf, f\rangle = \mathrm{Re}\langle iAf + iA^*f - \Phi^*\sigma\Phi f, f\rangle = -\langle \sigma\Phi f, \Phi f\rangle = -\langle \sigma(u-v), u-v\rangle,$$

where the second equality holds because $A + A^*$ is selfadjoint. Also,

$$2\,\mathrm{Re}\langle -i\Phi^*\sigma u, f\rangle = 2\,\mathrm{Re}\langle \sigma u, i\Phi f\rangle = 2\,\mathrm{Re}\langle \sigma u, u - v\rangle.$$

Putting together these two equations, we finally get

$$\begin{aligned}\frac{d}{dt}\langle f, f\rangle &= 2\,\mathrm{Re}\langle iAf - i\Phi^*\sigma u, f\rangle = -\langle \sigma(u-v), u-v\rangle + 2\,\mathrm{Re}\langle \sigma u, u-v\rangle \\ &= \langle \sigma u, u\rangle - \langle \sigma v, v\rangle.\end{aligned}$$

The coupling of colligations has a very natural interpretation in terms of systems. Let $\mathcal{C}'$ and $\mathcal{C}''$ be two colligations with the same rate, as above, and let f', u', v' and f'', u'', v'' denote the state, input and output of the system associated with each colligation. We can cascade connect these two systems by feeding the output of the first system into the input of the second system. This means that we form a new system with input u and output v and put $u' = u$, $u'' = v'$ and $v = v''$. The system we obtain has the equations

$$\begin{aligned} i\frac{df'}{dt} + A'f' &= \Phi'^*\sigma u, \\ i\frac{df''}{dt} + A''f'' &= \Phi''^*\sigma(u - i\Phi' f'), \\ v &= u - i\Phi' f' - i\Phi'' f''. \end{aligned} \tag{3.5}$$

Thus, we see that if we put $H = H' \oplus H''$, $\Phi = [\Phi', \Phi'']$ and define A by (3.2), then the system (3.5) is equivalent to (3.3) (with the identifications $f' = P_{H'}f$, $f'' = P_{H''}f$). This shows that the colligation which corresponds to the cascade connection of the systems is just the coupling $\mathcal{C}' \vee \mathcal{C}''$.

Another important tool in the study of colligations is the characteristic function. It is a $\mathcal{B}(E)$-valued analytic function on $\widehat{\mathbb{C}} \setminus \sigma(A)$ defined by

$$S(\lambda) = I - i\Phi(A - \lambda I)^{-1}\Phi^*\sigma. \tag{3.6}$$

The motivation for the characteristic function comes from the system interpretation: it is the transfer function of the system. If we assume that the state, input an output of the system are monochromatic waves of the same complex frequency $\lambda \in \mathbb{C}$, i.e.,

$$u(t) = u_0 e^{i\lambda t}, \qquad f(t) = f_0 e^{i\lambda t}, \qquad v(t) = v_0 e^{i\lambda t}, \tag{3.7}$$

then we define $S(\lambda)$ by the input to output relation: $v_0 = S(\lambda)u_0$.

To get formula (3.6), we plug in the expressions (3.7) into the system (3.3) and cancel out the factor $e^{i\lambda t}$ to get

$$\begin{aligned} -\lambda f_0 + Af_0 &= \Phi^*\sigma u_0 \\ v_0 &= u_0 - i\Phi f_0. \end{aligned}$$

This shows that $v_0 = u_0 - i\Phi(A - \lambda I)^{-1}\Phi^*\sigma u_0$ if $\lambda \notin \sigma(A)$.

From the system interpretation, it is clear that if $\mathcal{C} = \mathcal{C}' \vee \mathcal{C}''$ is the coupling of two colligations, its transfer function can be obtained as $S(\lambda) = S''(\lambda)S'(\lambda)$, where $S'(\lambda)$ and $S''(\lambda)$ are the characteristic functions of $\mathcal{C}'$ and $\mathcal{C}''$ respectively. Also, every A-invariant subspace H'' produces a decomposition of of the colligation, and hence, a factorization of the characteristic function.

This gives a link between the theory of factorization of analytic functions and the study of the invariant subspaces.

The characteristic function contains all the relevant information about the colligation. To make this notion precise, we need to introduce the unitary equivalence of colligations and to define the principal subspace. Two colligations $\mathcal{C} = (A; H, \Phi, E; \sigma)$ and $\mathcal{C}' = (A'; H', \Phi', E; \sigma)$ with the same rate are called unitarily equivalent if there is a unitary $U \in \mathcal{B}(H, H')$ such that $A' = UAU^*$ and $\Phi' = \Phi U^*$. The principal subspace of the colligation $\mathcal{C}$ is the space

$$\widehat{H} = \bigvee_{k\geq 0} A^k \Phi^* E = \bigvee_{k\geq 0} A^{*k} \Phi^* E.$$

Here, the second equality is not obvious, but it is not difficult to prove using (3.1) (see [LKMV95, Lemma 3.4.2]). Hence, the principal subspace reduces A. A colligation is called irreducible if $\widehat{H} = H$. We say that $\mathcal{C}$ and $\mathcal{C}'$ are unitarily equivalent on their principal subspace if the colligations $(A|\widehat{H}; \widehat{H}, \Phi|\widehat{H}, E; \sigma)$ and $(A'|\widehat{H'}; \widehat{H'}, \Phi'|\widehat{H'}, E; \sigma)$ are unitarily equivalent.

If $\sigma \neq 0$, the characteristic function determines the colligation up to unitary equivalence on the principal subspace. This means that if $\mathcal{C}$ and $\mathcal{C}'$ are two colligations which have the same characteristic function, then the two colligations are unitarily equivalent on their principal subspace. The proof of this result can be found in [LKMV95, Theorem 3.4.4].

3.2. Colligations of several commuting operators

Now we will try to generalize the concepts and results of the preceding section for a tuple $(A_1, \dots, A_n)$ of commuting operators. Some of the results will generalize with little changes, but when trying to generalize some others, we will find problems. These problems arise because, in some sense, the colligation does not contain enough information about the interplay of the operators A_k. This will motivate the concept of *vessel*, which will be introduced in the next section.

Definition (Colligation of several operators)**.** Given H a Hilbert space, E a finite-dimensional Hilbert space, $\Phi \in \mathcal{B}(H, E)$, selfadjoint operators $\sigma_k \in \mathcal{B}(E)$, $k = 1, \dots, n$, and operators $A_k \in \mathcal{B}(H)$, $k = 1, \dots, n$, we say that the tuple $\mathcal{C} = (A_k; H, \Phi, E; \sigma_k)$ is a colligation if

$$\frac{1}{i}(A_k - A_k^*) = \Phi^* \sigma_k \Phi, \qquad k = 1, \dots, n. \tag{3.8}$$

The colligation is called commutative if the operators A_k commute (i.e., if $A_j A_k = A_k A_j$). It is said to be strict if $\Phi H = E$ and $\bigcap_k \ker \sigma_k = 0$.

As in the previous section, the operators σ_k are called the rates of the colligation. Every tuple of commuting operators A_k with finite-dimensional imaginary parts can be embedded into a strict commutative colligation by defining $E = \bigvee_k (A_k - A_k^*)H$, $\Phi = P_E$ and $\sigma_k = (A_k - A_k^*)/i|E$.

Now we try to generalize the decomposition and coupling of colligations. Let $\mathcal{C}$ be a commutative colligation. If H'' is a joint invariant subspace for the operators $A_1, \dots, A_n$ (meaning that it is invariant for each operator A_k), then we can proceed as in the previous section and define $H' = H \ominus H''$, $A_k' = P_{H'} A_k | H'$, $A_k'' = A_k | H''$, $\Phi' = \Phi | H'$, and $\Phi'' = \Phi | H''$. Then the tuples $\mathcal{C}' = (A_k'; H', \Phi', E; \sigma_k)$ and $\mathcal{C}'' = (A_k''; H'', \Phi'', E; \sigma_k)$ are commutative colligations. Note that since the operators A_k commute, the operators A_k' also commute, and similarly, A_k'' commute. As before, the operators A_k have the form

$$A_k = \begin{bmatrix} A_k' & 0 \\ i\Phi''^* \sigma_k \Phi' & A_k'' \end{bmatrix}. \tag{3.9}$$

Now assume that we have commutative colligations $\mathcal{C}' = (A'_k; H', \Phi', E; \sigma_k)$ and $\mathcal{C}'' = (A''_k; H'', \Phi'', E; \sigma_k)$ with the same rates. We want to define their coupling as before. We put $H = H' \oplus H''$, $\Phi = [\Phi', \Phi'']$ and define A_k by (3.9). However, now we run into problems: in general, the operators A_k need not commute, even though A'_k commute and A''_k commute. Indeed, we see that A_k will commute if and only if the compatibility conditions

$$\Phi''^* \sigma_k \Phi' A'_j + A''_k \Phi''^* \sigma_j \Phi' = \Phi''^* \sigma_j \Phi' A'_k + A''_j \Phi''^* \sigma_k \Phi', \qquad j, k = 1, \dots, n \tag{3.10}$$

are satisfied.

Hence, we see that the coupling of two commutative colligations is not always a commutative colligation. The conditions (3.10) under which the coupling is commutative are not very useful, because they are difficult to understand. Hence, we cannot build commutative colligations by coupling several simple commutative colligations.

The dynamical system associated with a colligation $\mathcal{C}$ is

$$\begin{aligned} i\frac{\partial f}{\partial t_k} + A_k f &= \Phi^* \sigma_k u, \\ v &= u - i\Phi f. \end{aligned} \tag{3.11}$$

Now the state, input and output are functions defined on $\mathbb{R}^n$. There are two interpretations of this system. The first one is the *vector field interpretation*: given an input vector field $u(t)$ on $\mathbb{R}^n$, and an initial state $f_0 \in H$, find, if possible, vector fields $f(t)$ and $v(t)$ on $\mathbb{R}^n$ satisfying (3.11) and $f(0) = f_0$. The second one is the *curve interpretation*: given a piecewise smooth curve L on $\mathbb{R}^n$ parametrized by $(t_1(\tau), \dots, t_n(\tau))$, $\tau \in \mathbb{R}$, an input vector field $u(\tau)$ along L and an initial state $f_0 \in H$, find functions of the parameter $f = f(\tau)$ and $v = v(\tau)$ satisfying the system

$$\begin{aligned} i\frac{df}{d\tau} + \sum_{k=1}^{n} \frac{\partial t_k}{\partial \tau} A_k f &= \Phi^* \sum_{k=1}^{n} \frac{\partial t_k}{\partial \tau} \sigma_k u, \\ v &= u - i\Phi f, \end{aligned} \tag{3.12}$$

and $f(0) = f_0$. This system is obtained by restricting (3.11) to the curve L and writing everything as a function of the parameter τ.

The system (3.11) can be thought as a system having n independent temporal variables. One can also think of one of the variables, say t_1, representing time and the remaining $n-1$ variables representing space, so that (3.11) models a continuous of interacting temporal systems distributed in space.

In general, the system is overdetermined, and will not be consistent. Given a vector field $u(t)$ on $\mathbb{R}^n$, we will say that the system obtained is consistent if the vector field interpretation has a solution for any $f_0 \in H$. This is equivalent to the following condition using the curve interpretation: for any initial condition $f_0 \in H$, and any parametrized curve $(t_1(\tau), \dots, t_n(\tau))$ such that $(t_1(0), \dots, t_n(0)) = 0$, if $f(\tau)$ and $v(\tau)$ are the solutions of (3.12), then the values $f(1)$ and $v(1)$ depend only on the initial condition f_0 and the point $p = (t_1(1), \dots, t_n(1))$, but not on the curve in question which joins 0 and p.

One can see (in [LKMV95, Theorem 3.2.1], for instance) that the necessary and sufficient conditions for the system to be consistent are the compatibility conditions which arise from the equality of the mixed partial derivatives

$$\frac{\partial^2 f}{\partial t_j \partial t_k} = \frac{\partial^2 f}{\partial t_k \partial t_j}.$$

When the system is given the zero input (i.e., $u \equiv 0$), we see that the system is consistent because the operators A_k commute. However, for an arbitrary input u, the system is not going to be consistent in general. Let us deduce what are the conditions on u for the equality of the mixed partial derivatives.

We compute

$$\frac{\partial^2 f}{\partial t_k \partial t_j} = \frac{\partial}{\partial t_k}(iA_j f - i\Phi^* \sigma_j u) = iA_j(iA_k f - i\Phi^* \sigma_k u) - i\Phi^* \sigma_j \frac{\partial u}{\partial t_k}.$$

Hence, since A_k commute, we see that the equality of the mixed partials is equivalent to the conditions

$$\left(\Phi^* \sigma_j \frac{\partial}{\partial t_k} - \Phi^* \sigma_k \frac{\partial}{\partial t_j} + iA_j \Phi^* \sigma_k - iA_k \Phi^* \sigma_j\right) u = 0, \qquad j, k = 1, \dots, n. \tag{3.13}$$

Once again, this condition is complicated, so we do not obtain a satisfactory interpretation of commutative colligations in terms of systems.

However, even if the system is not consistent, we can always consider the curve interpretation with $L = L_\xi$ the straight line given by $(\xi_1 \tau, \dots, \xi_n \tau)$, where $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ is a fixed direction. Then the system (3.12) that we obtain is the system associated to the single operator colligation $\mathcal{C}_\xi = (\xi A; H, \Phi, E; \xi\sigma)$, where we abuse notation a bit and write

$$\xi A = \sum_{k=1}^n \xi_k A_k, \qquad \xi\sigma = \sum_{k=1}^n \xi_k \sigma_k.$$

Hence, applying the results of the preceding section, we obtain the conservation law

$$\sum_{k=1}^n \xi_k \frac{\partial}{\partial t_k} \langle f, f \rangle = \langle \xi\sigma u, u \rangle - \langle \xi\sigma v, v \rangle. \tag{3.14}$$

This means that the energy of the system is conserved along any direction ξ, when one takes into account the energy added at the input and the energy extracted at the output, measured with the indefinite quadratic form $\xi\sigma$.

The theory of the characteristic function generalizes pretty well to commutative colligations. The analogue here is the complete characteristic function $S(\xi, z)$, where $\xi \in \mathbb{C}^n$, and $z \in \mathbb{C}$. It is a function analytic on (ξ, z) outside the set where $z \in \sigma(\xi A)$, and is defined by

$$S(\xi, z) = I - i\Phi(\xi A - zI)^{-1}\Phi^* \xi\sigma.$$

When $\xi \in \mathbb{R}^n$, this has the clear interpretation of the characteristic function of the single operator colligation $\mathcal{C}_\xi = (\xi A; H, \Phi, E; \xi\sigma)$, so it has a system theoretical interpretation. When $\xi \notin \mathbb{R}^n$, one can just think of $S(\xi, z)$ as an analytic continuation.

When the colligation $\mathcal{C}$ decomposes into colligations $\mathcal{C}'$ and $\mathcal{C}''$, we obtain a factorization of the complete characteristic function $S(\xi, z) = S''(\xi, z)S'(\xi, z)$, where $S'(\xi, z)$ and $S''(\xi, z)$ are the complete characteristic functions of $\mathcal{C}'$ and $\mathcal{C}''$ respectively. To prove this fact, we first consider the case when $\xi \in \mathbb{R}^n$. In this case, $S(\xi, z)$ is the characteristic function of the single operator colligation $\mathcal{C}_\xi = (\xi A; H, \Phi, E; \xi\sigma)$, which decomposes into colligations $\mathcal{C}'_\xi = (\xi A'; H', \Phi', E; \xi\sigma)$ and $\mathcal{C}''_\xi = (\xi A''; H'', \Phi', E; \xi\sigma)$. Hence, the factorization follows from the theory of single operators colligations. The general case is obtained by analytic continuation.

The principal subspace of the colligation $\mathcal{C}$ is defined as the subspace

$$\widehat{H} = \bigvee_{k_1,\dots,k_n\geq 0} A_1^{k_1}\cdots A_n^{k_n}H = \bigvee_{k_1,\dots,k_n\geq 0} A_1^{*k_1}\cdots A_n^{*k_n}H.$$

Unitary equivalence of colligations is defined as in the previous section. Then, if the non-degeneracy condition

$$\det \xi\sigma \not\equiv 0 \tag{3.15}$$

holds, the complete characteristic function $S(\xi, z)$ determines the colligation up to unitary equivalence on the principal subspace.

3.3. Vessels of several commuting operators

In the previous section, we have seen how some important aspects of the theory of single operator colligations cannot be satisfactorily generalized to several commuting operators. The main problem is that we encountered conditions such as (3.10) and (3.13) which do not seem very natural and are also difficult to check because they involve the operators A_k, which in general act on an infinite-dimensional space. We would like to reformulate these conditions in terms of operators acting on the finite-dimensional outer space.

The key point here is that the rates σ_k of the colligation do not capture all the information we need, because they only model the imaginary parts of the operators A_k. In particular, they do not capture any information about the interplay of the operators A_k. It turns out that what is needed is another tuple of operators which reflect in some sense the relation between the operators A_k and A_j^*.

To motivate the definition of this new tuple, assume that we have a commutative strict colligation $\mathcal{C} = (A_k; H, \Phi, E; \sigma_k)$. First note that

$$\frac{1}{i}(A_kA_j^* - A_jA_k^*) = \frac{1}{i}[(A_k - A_k^*)A_j^* - (A_j - A_j^*)A_k^*], \tag{3.16}$$

and

$$\frac{1}{i}(A_j^*A_k - A_k^*A_j) = \frac{1}{i}[(A_k - A_k^*)A_j - (A_j - A_j^*)A_k]. \tag{3.17}$$

Now (3.8) shows that the range of the operators on the left hand side of these equalities is contained in $G = \Phi^*E$. It is easy to see that since the colligation is strict, the operators $\Phi|G$ and Φ^* are one-to-one maps taking G onto E and E onto G respectively (see [LKMV95, Proposition 2.1.2]). Hence, if we define operators γ_{kj}^{in} and γ_{kj}^{out} on E by

$$\begin{aligned}\gamma_{kj}^{\text{in}} &= (\Phi^*)^{-1}\left[\frac{1}{i}(A_kA_j^* - A_jA_k^*)\right](\Phi|G)^{-1},\\ \gamma_{kj}^{\text{out}} &= (\Phi^*)^{-1}\left[\frac{1}{i}(A_j^*A_k - A_kA_j^*)\right](\Phi|G)^{-1},\end{aligned}$$

then we get

$$\begin{aligned}\frac{1}{i}(A_kA_j^* - A_jA_k^*) &= \Phi^*\gamma_{kj}^{\text{in}}\Phi,\\ \frac{1}{i}(A_j^*A_k - A_k^*A_j) &= \Phi^*\gamma_{kj}^{\text{out}}\Phi,\end{aligned} \tag{3.18}$$

Hence, the operators γ_{kj}^{in} and γ_{kj}^{out} are operators on the outer space which model the operators on the left hand side of equalities (3.18). The "in" and "out" labels will make sense when we get to discuss the system theoretical interpretation. Note that γ_{kj}^{in} and γ_{kj}^{out} are selfadjoint and satisfy $\gamma_{kj}^{\text{in}} = -\gamma_{jk}^{\text{in}}$ and $\gamma_{kj}^{\text{out}} = -\gamma_{jk}^{\text{out}}$.

Using equations (3.8), (3.16) and (3.17), we obtain

$$\begin{aligned}\Phi^*(\sigma_k \Phi A_j^* - \sigma_j \Phi A_k^*) &= \Phi^* \gamma_{kj}^{\text{in}} \Phi,\\ \Phi^*(\sigma_k \Phi A_j - \sigma_j \Phi A_k) &= \Phi^* \gamma_{kj}^{\text{out}} \Phi.\end{aligned}$$

Since the colligation is strict, Φ^* can be cancelled to yield

$$\begin{aligned}\sigma_k \Phi A_j^* - \sigma_j \Phi A_k^* &= \gamma_{kj}^{\text{in}} \Phi,\\ \sigma_k \Phi A_j - \sigma_j \Phi A_k &= \gamma_{kj}^{\text{out}} \Phi.\end{aligned} \tag{3.19}$$

Subtracting these two inequalities, using (3.8) and cancelling out the factor Φ on the right (again, this can be done because the colligation is strict), we also get

$$\gamma_{kj}^{\text{out}} = \gamma_{kj}^{\text{in}} + i(\sigma_k \Phi\Phi^* \sigma_j - \sigma_j \Phi\Phi^* \sigma_k). \tag{3.20}$$

As we have already mentioned, the strictness of a commutative colligation is generally too restrictive. However, the identities (3.19) and (3.20) are all we need to give good generalizations of the concepts in Section 3.1. Hence, this motivates the definition of an *operator vessel*.

Definition (Operator vessel)**.** Suppose that we are given H a Hilbert space, E a finite-dimensional Hilbert space, $\Phi \in \mathcal{B}(H, E)$, a tuple of commuting operators $A_k \in \mathcal{B}(H)$, $k = 1, \dots, n$, selfadjoint operators $\sigma_k \in \mathcal{B}(E)$, and selfadjoint operators $\gamma_{kj}^{\text{in}}, \gamma_{kj}^{\text{out}} \in \mathcal{B}(E)$ satisfying $\gamma_{kj}^{\text{in}} = -\gamma_{jk}^{\text{in}}$ and $\gamma_{kj}^{\text{out}} = -\gamma_{jk}^{\text{out}}$. We say that the tuple $\mathcal{V} = (A_k; H, \Phi, E; \sigma_k, \gamma_{kj}^{\text{in}}, \gamma_{kj}^{\text{out}})$ is a (commutative) vessel if the following conditions are satisfied:

$$\frac{1}{i}(A_k - A_k^*) = \Phi^* \sigma_k \Phi, \tag{3.21}$$

$$\sigma_k \Phi A_j^* - \sigma_j \Phi A_k^* = \gamma_{kj}^{\text{in}} \Phi, \tag{3.22}$$

$$\gamma_{kj}^{\text{out}} = \gamma_{kj}^{\text{in}} + i(\sigma_k \Phi\Phi^* \sigma_j - \sigma_j \Phi\Phi^* \sigma_k), \tag{3.23}$$

$$\sigma_k \Phi A_j - \sigma_j \Phi A_k = \gamma_{kj}^{\text{out}} \Phi. \tag{3.24}$$

The operators γ_{kj}^{in} and γ_{kj}^{out} are called the *gyrations* of the vessel. It is easy to see that conditions (3.24) follow from (3.21)–(3.23), and that conditions (3.22) follow from (3.21), (3.23) and (3.24). Conditions (3.23) are called the *linkage* conditions.

The discussion above shows that any strict commutative colligation can be embedded in a vessel. In particular, any tuple of commuting operators A_k with finite-dimensional imaginary parts can be embedded in a vessel. As we will show now, vessels are the appropriate tool for studying commuting operators with finite-dimensional imaginary parts.

Let us first discuss the system theoretical interpretation, because this will later motivate the decomposition and coupling of vessels. If $\mathcal{V} = (A_k; H, \Phi, E; \sigma_k, \gamma_{kj}^{\text{in}}, \gamma_{kj}^{\text{out}})$ is a vessel, we assign to it the system (3.11), as we did with colligations in the preceding section. The problem is to rewrite the compatibility conditions (3.13) using the gyrations of the vessel. Taking adjoints in (3.22), we see that (3.13) rewrites as

$$\Phi^* \left(\sigma_j \frac{\partial}{\partial t_k} - \sigma_k \frac{\partial}{\partial t_j} + i\gamma_{kj}^{\text{in}} \right) u = 0, \qquad j, k = 1, \dots, n. \tag{3.25}$$

Hence, we see that an input vector field u makes the system (3.11) compatible if an only u satisfies the compatibility conditions (3.25). This explains the label "in" in the gyrations γ_{kj}^{in}.

Now we will see that the output v of the system satisfies similar compatibility conditions. Since $u = v + i\Phi f$, we get

$$\Phi^* \left(\sigma_j \frac{\partial}{\partial t_k} - \sigma_k \frac{\partial}{\partial t_j} + i\gamma_{kj}^{\text{in}}\right)(v + i\Phi f) = 0.$$

Now we compute

$$\begin{aligned}\Phi^*\left(\sigma_j \frac{\partial}{\partial t_k} - \sigma_k \frac{\partial}{\partial t_j} + i\gamma_{kj}^{\text{in}}\right)(i\Phi f) = \\ &= \Phi^*\sigma_j\Phi(\Phi^*\sigma_k u - A_k f) - \Phi^*\sigma_k\Phi(\Phi^*\sigma_j u - A_j f) - \Phi^*\gamma_{kj}^{\text{in}}\Phi f \\ &= i\Phi^*\gamma_{kj}^{\text{out}} u - i\Phi^*\gamma_{kj}^{\text{in}} u + \Phi^*\gamma_{kj}^{\text{out}}\Phi f - \Phi^*\gamma_{kj}^{\text{in}}\Phi f \\ &= i\Phi^*(\gamma_{kj}^{\text{out}} - \gamma_{kj}^{\text{in}})v.\end{aligned}$$

Here we have used the system (3.11) in the first equality, the vessel conditions (3.23) and (3.24) in the second equality and the identity $v = u - i\Phi f$ in the third equality. Hence, we see that the output of the system satisfies the compatibility conditions

$$\Phi^* \left(\sigma_j \frac{\partial}{\partial t_k} - \sigma_k \frac{\partial}{\partial t_j} + i\gamma_{kj}^{\text{out}}\right) v = 0, \qquad j, k = 1, \dots, n. \tag{3.26}$$

If we drop Φ^* in (3.25), we obtain the conditions

$$\left(\sigma_j \frac{\partial}{\partial t_k} - \sigma_k \frac{\partial}{\partial t_j} + i\gamma_{kj}^{\text{in}}\right) u = 0, \qquad j, k = 1, \dots, n. \tag{3.27}$$

These conditions are sufficient for the input u to make the system compatible, and are also necessary when Φ^* is injective. The important aspect of these conditions is that they are written entirely in terms of operators on E, so they can be checked using matrices. If the input u satisfies (3.27), arguing as above we see that the output v satisfies the compatibility conditions

$$\left(\sigma_j \frac{\partial}{\partial t_k} - \sigma_k \frac{\partial}{\partial t_j} + i\gamma_{kj}^{\text{out}}\right) v = 0, \qquad j, k = 1, \dots, n. \tag{3.28}$$

With this in mind, the system theoretical interpretation of the vessel $\mathcal{V}$ is the system (3.11) together with the compatibility conditions (3.25) and (3.26) (or (3.27) and (3.28)) at the input and output respectively. Indeed, it is more usual to take (3.27) and (3.28) as the compatibility conditions.

Now we turn to the analysis of the decomposition and coupling of vessels. The coupling of vessels is actually easier to deduce, because it corresponds to cascade connection of systems, as in the case of single operator colligations. Let $\mathcal{V}' = (A_k'; H', \Phi', E; \sigma_k; \gamma_{kj}^{\text{in}\,\prime}, \gamma_{kj}^{\text{out}\,\prime})$ and $\mathcal{V}'' = (A_k''; H'', \Phi'', E; \sigma_k; \gamma_{kj}^{\text{in}\,\prime\prime}, \gamma_{kj}^{\text{out}\,\prime\prime})$ be two vessels with the same rates. The important thing if one wishes to cascade connect their corresponding systems is that whenever we feed some compatible input u' into the first system, then the output v' it produces should be a compatible input for the second system. This imposes the matching conditions

$$\gamma_{kj}^{\text{out}\,\prime} = \gamma_{kj}^{\text{in}\,\prime\prime}. \tag{3.29}$$

If the matching conditions hold, then we define H, Φ and A_k as we did for colligations and put

$$\gamma_{kj}^{\text{in}} = \gamma_{kj}^{\text{in}\,\prime}, \qquad \gamma_{kj}^{\text{out}} = \gamma_{kj}^{\text{out}\,\prime\prime},$$

which again is natural in view of the system interpretation: the compatibility conditions for the input of the cascaded system are those at the input of the first system, and the compatibility conditions for the output of the cascaded system are those at the output of the second system.

Now we should check that the coupling $\mathcal{V} = \mathcal{V}' \vee \mathcal{V}'' = (A_k; H, \Phi, E; \sigma_k, \gamma_{kj}^{\text{in}}, \gamma_{kj}^{\text{out}})$ is indeed a vessel. In the preceding section we saw that the operators A_k commute if and only if (3.10) is satisfied. Rewriting these conditions in terms of the gyrations of the vessels, we get

$$\Phi''^* \gamma_{kj}^{\text{out}\,\prime} \Phi' = \Phi''^* \gamma_{kj}^{\text{in}\,\prime\prime} \Phi'. \tag{3.30}$$

These conditions hold, because of the matching conditions (3.29). Moreover, if we start with strict colligations $\mathcal{C}'$ and $\mathcal{C}''$ and embed them into vessels $\mathcal{V}'$ and $\mathcal{V}''$ respectively, then the operators Φ''^* and Φ' in (3.30) can be cancelled. Hence, we see that, in this case, the matching conditions (3.29) are also necessary for the coupling $\mathcal{C}' \vee \mathcal{C}''$ to be a commutative colligation.

We also need to check the vessel conditions (3.21)–(3.24) for the coupling $\mathcal{V} = \mathcal{V}' \vee \mathcal{V}''$. Conditions (3.21) are checked in the same way as for a single operator colligation. Conditions (3.23) hold because

$$\begin{aligned}\gamma_{kj}^{\text{out}} - \gamma_{kj}^{\text{in}} &= \gamma_{kj}^{\text{out}\,\prime\prime} - \gamma_{kj}^{\text{in}\,\prime\prime} + \gamma_{kj}^{\text{out}\,\prime} - \gamma_{kj}^{\text{in}\,\prime} \\ &= i(\sigma_k \Phi'' \Phi''^* \sigma_j - \sigma_j \Phi'' \Phi''^* \sigma_k) + i(\sigma_k \Phi' \Phi'^* \sigma_j - \sigma_j \Phi' \Phi'^* \sigma_k) \\ &= i(\sigma_k \Phi \Phi^* \sigma_j - \sigma_j \Phi \Phi^* \sigma_k).\end{aligned}$$

Conditions (3.24) restricted to H'' are true because H'' is joint invariant for A_k, so they are equivalent to the corresponding conditions for $\mathcal{V}''$. To check the conditions restricted to H', we compute

$$\begin{aligned}(\sigma_k \Phi A_j - \sigma_j \Phi A_k)|H' &= \sigma_k(\Phi' A_j' + i\Phi''\Phi''^* \sigma_j \Phi') - \sigma_j(\Phi' A_k' + i\Phi''\Phi''^* \sigma_k \Phi') \\ &= \gamma_{kj}^{\text{out}\,\prime} \Phi' + (\gamma_{kj}^{\text{out}\,\prime\prime} - \gamma_{kj}^{\text{in}\,\prime\prime})\Phi' = \gamma_{kj}^{\text{out}} \Phi | H'.\end{aligned}$$

Finally, conditions (3.22) follow from (3.21), (3.23) and (3.24).

The decomposition of a vessel can also be understood in a similar way. Assume that $\mathcal{V} = (A_k; H, \Phi, E; \sigma_k, \gamma_{kj}^{\text{in}}, \gamma_{kj}^{\text{out}})$ is a vessel, H'' is a joint invariant subspace of the operators A_k and $H' = H \ominus H''$. We have seen in the preceding section how the colligation $(A_k; H, \Phi, E; \sigma_k)$ can be decomposed with respect to H'' into two colligations. To decompose $\mathcal{V}$, we define the operators A_k', A_k'', Φ' and Φ'' in the same way.

We also need to define the gyrations $\gamma_{kj}^{\text{in}\,\prime}$, $\gamma_{kj}^{\text{in}\,\prime\prime}$, $\gamma_{kj}^{\text{out}\,\prime}$ and $\gamma_{kj}^{\text{out}\,\prime\prime}$. To do this, we imagine that we break the system associated with $\mathcal{V}$ into two cascade connected systems, which will be associated with the vessels $\mathcal{V}'$ and $\mathcal{V}''$. Then it is apparent that we should define $\gamma_{kj}^{\text{in}\,\prime} = \gamma_{kj}^{\text{in}}$ and $\gamma_{kj}^{\text{out}\,\prime\prime} = \gamma_{kj}^{\text{out}\,\prime\prime}$. The system interpretation does not tell us how to define $\gamma_{kj}^{\text{out}\,\prime}$ and $\gamma_{kj}^{\text{in}\,\prime\prime}$. It only tells us that they should match: $\gamma_{kj}^{\text{out}\,\prime} = \gamma_{kj}^{\text{in}\,\prime\prime}$. However, the linkage conditions (3.23) for $\mathcal{V}'$ and $\mathcal{V}''$ allow us to define $\gamma_{kj}^{\text{out}\,\prime}$ and $\gamma_{kj}^{\text{in}\,\prime\prime}$ in terms of $\gamma_{kj}^{\text{in}\,\prime}$ and $\gamma_{kj}^{\text{out}\,\prime\prime}$. We put

$$\begin{aligned}\gamma_{kj}^{\text{out}\,\prime} &= \gamma_{kj}^{\text{in}\,\prime} + i(\sigma_k \Phi' \Phi'^* \sigma_j - \sigma_j \Phi' \Phi'^* \sigma_k), \\ \gamma_{kj}^{\text{in}\,\prime\prime} &= \gamma_{kj}^{\text{out}\,\prime\prime} - i(\sigma_k \Phi'' \Phi''^* \sigma_j - \sigma_j \Phi'' \Phi''^* \sigma_k).\end{aligned}$$

A simple computation shows that $\gamma_{kj}^{\text{out}\,\prime} = \gamma_{kj}^{\text{in}\,\prime\prime}$.

Hence, the decomposition of $\mathcal{V}$ with respect to H'' is formed by the vessels $\mathcal{V}' = (A_k'; H', \Phi', E; \sigma_k, \gamma_{kj}^{\text{in}\,\prime}, \gamma_{kj}^{\text{out}\,\prime})$ and $\mathcal{V}'' = (A_k''; H', \Phi', E; \sigma_k, \gamma_{kj}^{\text{in}\,\prime\prime}, \gamma_{kj}^{\text{out}\,\prime\prime})$. Routine checks show that $\mathcal{V}'$ and $\mathcal{V}''$ are indeed vessels and that $\mathcal{V}$ is the coupling $\mathcal{V} = \mathcal{V}' \vee \mathcal{V}''$.

If $\mathcal{V} = (A_k; H, \Phi, E; \sigma_k; \gamma_{kj}^{\text{in}}, \gamma_{kj}^{\text{out}})$ is a vessel, then the tuple $(A_k; H, \Phi, E; \sigma_k)$ is a commutative colligation. Hence, the concepts of complete characteristic function, principal subspace and irreducibility given in the previous section, can also be applied to vessels.

Another construction that vessels allow is that of the *discriminant varieties*, which allow one to give a generalization of the Cayley-Hamilton Theorem, and also give a link between algebraic geometry and operator theory. We define the input discriminant ideal $\mathcal{J}^{\text{in}}$ as the ideal in $\mathbb{C}[z_1, \dots, z_n]$ generated by the polynomials of the form

$$\det\left(\sum_{j,k=1}^{n} \Gamma^{jk}(z_j\sigma_k - z_k\sigma_j + \gamma_{jk}^{\text{in}})\right),$$

where $\Gamma^{jk} = -\Gamma^{kj}$ can be arbitrary operators on the outer space E. Similarly, the output discriminant ideal $\mathcal{J}^{\text{out}}$ is defined by replacing the input gyrations γ_{jk}^{in} by the output gyrations γ_{jk}^{out}. The input and output discriminant varieties, D^{in} and D^{out} respectively, are defined as the algebraic varieties in $\mathbb{C}^n$ associated with these ideals. This means that

$$D^{\text{in}} = \{z \in \mathbb{C}^n : p(z) = 0, \forall p \in \mathcal{J}^{\text{in}}\}, \qquad D^{\text{out}} = \{z \in \mathbb{C}^n : p(z) = 0, \forall p \in \mathcal{J}^{\text{out}}\}.$$

We can also define, for $z \in \mathbb{C}^n$, the following subspaces of E:

$$E^{\text{in}}(z) = \bigcap_{j,k=1}^{n} \ker(z_j\sigma_k - z_k\sigma_j + \gamma_{jk}^{\text{in}}), \qquad E^{\text{in}}(z) = \bigcap_{j,k=1}^{n} \ker(z_j\sigma_k - z_k\sigma_j + \gamma_{jk}^{\text{out}}).$$

Their connection with the discriminant varieties is that $z \in D^{\text{in}}$ if and only if $E^{\text{in}}(z) \neq 0$, and analogously for the output discriminant variety (see [LKMV95, Proposition 4.1.3]).

Now we can give the statement of the generalized Cayley-Hamilton theorem. For the proof, see [LKMV95, Theorem 4.1.2].

Theorem 3.1 (Generalized Cayley-Hamilton)**.** *Let $\mathcal{V} = (A_k; H, \Phi, E; \sigma_k, \gamma_{kj}^{in}, \gamma_{kj}^{out})$ be an irreducible vessel, and $p^{in}(z) \in \mathcal{J}^{in}$, $p^{out}(z) \in \mathcal{J}^{out}$ arbitrary polynomials in the input and output discriminant ideals of the vessel. Then the operators $A_1, \dots, A_n$ satisfy the algebraic equations*

$$p^{in}(A_1^*, \dots, A_n^*) = 0, \qquad p^{out}(A_1, \dots, A_n) = 0.$$

In the next section we will give the statement of this theorem for vessels of two operators and show how the classical Cayley-Hamilton theorem can be derived from it.

Another construction that can be done with vessels is that of the *adjoint vessel*. If $\mathcal{V} = (A_k; H, \Phi, E; \sigma_k, \gamma_{kj}^{\text{in}}, \gamma_{kj}^{\text{out}})$ is a vessel, then it is easy to check that $\mathcal{V}^* = (A_k^*; H, -\Phi, E; -\sigma_k, -\gamma_{kj}^{\text{out}}, -\gamma_{kj}^{\text{in}})$ is also a vessel. This vessel is called the adjoint vessel. It is also easy to see that when passing from a vessel $\mathcal{V}$ to its adjoint vessel $\mathcal{V}^*$, the input and output discriminant varieties D^{in} and D^{out} interchange, and so do the subspaces $E^{\text{in}}(z)$ and $E^{\text{out}}(z)$. There is an interesting interpretation of the adjoint vessel in terms of the system theoretical interpretation: it corresponds to running the system backwards, i.e., interchanging the input and the output of the system. The proof of this fact is a simple computation (see [LKMV95, Proposition 3.3.1]).

3.4. Vessels of two commuting operators

For a vessel of two operators all the results of the preceding section can be applied. Moreover, one can make some simplifications that allow to develop the theory further.

First note that since $\gamma_{12}^{\mathrm{in}} = -\gamma_{21}^{\mathrm{in}}$, there is essentially only one input gyration. We will write $\gamma^{\mathrm{in}} = \gamma_{12}^{\mathrm{in}}$. Analogously, there is essentially only one output gyration and we will write $\gamma^{\mathrm{out}} = \gamma_{12}^{\mathrm{out}}$. Also, the input discriminant ideal $\mathcal{J}^{\mathrm{in}}$ is principal, which means that it is generated by a single polynomial. This happens because

$$\det\left(\Gamma^{12}(z_1\sigma_2 - z_2\sigma_1 + \gamma_{12}^{\mathrm{in}}) + \Gamma^{21}(z_2\sigma_1 - z_1\sigma_2 + \gamma_{21}^{\mathrm{in}})\right) = \det(2\Gamma^{12})\det(z_1\sigma_2 - z_2\sigma_1 + \gamma^{\mathrm{in}}).$$

Hence, the polynomial $\det(z_1\sigma_2 - z_2\sigma_1 + \gamma^{\mathrm{in}})$ generates the ideal $\mathcal{J}^{\mathrm{in}}$. This polynomial is called the input discriminant polynomial. A similar thing happens with the output discriminant ideal: it is generated by the polynomial $\det(z_1\sigma_2 - z_2\sigma_1 + \gamma^{\mathrm{out}})$, which is called the output discriminant polynomial.

Indeed, it turns out that the input and output discriminant polynomials are equal:

$$\det(z_1\sigma_2 - z_2\sigma_1 + \gamma^{\mathrm{in}}) = \det(z_1\sigma_2 - z_2\sigma_1 + \gamma^{\mathrm{out}}).$$

A proof of this equality is given in [LKMV95, Corollary 4.2.2]. We will denote by $\Delta(z_1, z_2)$ this polynomial, and we will call it the *discriminant polynomial*. This implies that the input and output discriminant varieties coincide, so we will write $D = D^{\mathrm{in}} = D^{\mathrm{out}}$. The variety D is either an algebraic curve in $\mathbb{C}^2$ or all of $\mathbb{C}^2$ (this second case is considered to be degenerate), so it will usually be called the *discriminant curve* of the vessel.

For vessels of more than two operators, the input and output discriminant varieties are distinct in general. However, they may only differ by a finite number of isolated points. It is believed that this points may be related to some pathologies involving commuting tuples of more than two operators, such as the failure of von Neumann's inequality or the non-existence of a dilation. A discussion of this fact is included in [LKMV95, Section 7.2].

With these observations, the generalized Cayley-Hamilton theorem becomes the following Theorem.

Theorem 3.2 (Generalized Cayley-Hamilton)**.** *Let $\mathcal{V} = (A_1, A_2; H, \Phi, E; \sigma_1, \sigma_2, \gamma^{in}, \gamma^{out})$ be an irreducible two operator vessel, and $\Delta(z_1, z_2)$ its discriminant polynomial. Then the operators A_1, A_2 satisfy the algebraic equations*

$$\Delta(A_1, A_2) = 0, \qquad \Delta(A_1^*, A_2^*) = 0.$$

One can derive the classical Cayley-Hamilton theorem from this Theorem. Indeed, let A be an operator on a finite-dimensional Hilbert space H of dimension m. We put $A_1 = A$, $A_2 = iI$ and embed these two operators in a vessel following the procedure that we have explained above. The space E is defined by

$$E = \frac{1}{i}(A_1 - A_1^*)H + \frac{1}{i}(A_2 - A_2^*)H = H.$$

Here the second equality comes from $A_2 - A_2^* = 2iI$. Hence, $\Phi = P_E = I$. The rates are defined by

$$\sigma_1 = \frac{1}{i}(A - A^*), \qquad \sigma_2 = 2I,$$

and the gyrations can be computed as

$$\gamma^{\mathrm{in}} = \gamma_{12}^{\mathrm{in}} = \frac{1}{i}(A_1A_2^* - A_2A_1^*) = -(A + A^*),$$
$$\gamma^{\mathrm{out}} = \gamma_{12}^{\mathrm{out}} = \frac{1}{i}(A_2^*A_1 - A_1^*A_2) = -(A + A^*).$$

Hence, the discriminant polynomial is

$$\Delta(z_1, z_2) = \det\left(z_1 2\widetilde{I} + z_2 i(\widetilde{A} - \widetilde{A}^*) - (\widetilde{A} + \widetilde{A}^*)\right),$$

where $\widetilde{I}$ and $\widetilde{A}$ are the corresponding matrices. Hence, the conclusion of the theorem corresponds to

$$0 = \Delta(A_1, A_2) = \det\left(2A\widetilde{I} - I(\widetilde{A} - \widetilde{A}^*) - I(\widetilde{A} + \widetilde{A}^*)\right) = 2^m \det(A\widetilde{I} - I\widetilde{A}).$$

Recall that the classical Cayley-Hamilton theorem states that the polynomial

$$p(z) = \det(z\widetilde{I} - \widetilde{A})$$

satisfies $p(A) = 0$, but this is just the identity

$$\det(A\widetilde{I} - I\widetilde{A}) = 0.$$

An important tool in the study of two operator vessels is the joint characteristic function. We have seen that one can define a complete characteristic function $S(\xi, z)$, $\xi \in \mathbb{C}^n$, $z \in \mathbb{C}$ for commutative colligations, and therefore also for vessels. In the case of a two operator vessel, the complete characteristic function has the form

$$S(\xi_1, \xi_2, z) = I - i\Phi(\xi_1 A_1 + \xi_2 A_2 - zI)^{-1}\Phi^*(\xi_1\sigma_1 + \xi_2\sigma_2).$$

This is a function of three complex variables, but because of homogeneity, it can be thought of as a function of two independent complex variables. We have also seen that for single operator colligations, there is a relation between the factorizations of its characteristic function and the invariant subspaces of the operator in the colligation. However, functions of two complex variables do not admit a good factorization theory, so the complete characteristic function of a two operator vessel is not a good analogue to study invariant subspaces.

The good analogue of the characteristic function for two operator vessels is the *joint characteristic function* $\widehat{S}(z)$. If $z = (z_1, z_2) \in D$ is a point on the discriminant curve, then the operator

$$S(\xi_1, \xi_2, \xi_1 z_1 + \xi_2 z_2)|E^{\text{in}}(z)$$

does not depend on the election of $(\xi_1, \xi_2) \in \mathbb{C}^2$ as long as

$$\xi_1 z_1 + \xi_2 z_2 \notin \sigma(\xi_1 A_1 + \xi_2 A_2). \tag{3.31}$$

Moreover, this operator maps $E^{\text{in}}(z)$ into $E^{\text{out}}(z)$. A proof of these facts can be seen in [LKMV95, Theorem 4.3.1]. We will also give later a proof based on the system theoretical interpretation. Hence, we can define the joint characteristic function

$$\widehat{S}(z) : E^{\text{in}}(z) \to E^{\text{out}}(z),$$

for all $z = (z_1, z_2) \in D$ for which there exists $(\xi_1, \xi_2) \in \mathbb{C}^2$ such that (3.31) holds, by

$$\widehat{S}(z) = S(\xi_1, \xi_2, \xi_1 z_1 + \xi_2 z_2)|E^{\text{in}}(z).$$

One can think of $E^{\text{in}}(z)$ and $E^{\text{out}}(z)$ as vector bundles on the algebraic curve D. More precisely, they are vector bundles on the desingularization of D (see Section 4.2 for an introduction to the desingularization of an algebraic curve in a slightly different context). Hence, $\widehat{S}$ plays the role of a bundle map on an algebraic curve, so it is a function of one independent complex variable, and admits a good factorization theory.

The joint characteristic function has also an interpretation as a transfer function of the associated system. We take a double frequency $\lambda = (\lambda_1, \lambda_2) \in \mathbb{C}^2$ and assume that the input, state and output of the system are waves with this frequency:

$$u(t_1, t_2) = u_0 e^{i\lambda_1 t_1 + i\lambda_2 t_2}, \quad f(t_1, t_2) = f_0 e^{i\lambda_1 t_1 + i\lambda_2 t_2}, \quad v(t_1, t_2) = v_0 e^{i\lambda_1 t_1 + i\lambda_2 t_2}.$$

Then the input and output compatibility conditions (3.27) and (3.28) are

$$(\lambda_1 \sigma_2 - \lambda_2 \sigma_1 + \gamma^{\text{in}})u_0 = 0, \qquad (\lambda_1 \sigma_2 - \lambda_2 \sigma_1 + \gamma^{\text{out}})v_0 = 0.$$

This means that $\lambda \in D$, $u_0 \in E^{\text{in}}(\lambda)$ and $v_0 \in E^{\text{out}}(\lambda)$. If we integrate the system along any temporal straight line $(\xi_1 \tau, \xi_2 \tau)$, and apply the result obtained for the transfer function of a single operator colligation, we get

$$v_0 = S(\xi_1, \xi_2; \xi_1 \lambda_1 + \xi_2 \lambda_2)u_0.$$

This proves that $S(\xi_1, \xi_2; \xi_1 z_1 + \xi_2 z_2)|E^{\text{in}}(z)$ does not depend on ξ, that it maps $E^{\text{in}}(\lambda)$ into $E^{\text{out}}(\lambda)$ and that

$$v_0 = \widehat{S}(\lambda)u_0.$$

Finally, another important fact of the joint characteristic function is that the complete characteristic function can be recovered from it by the so called *restoration formula*. Assume that $\dim E = m$ and that the discriminant curve D has degree m (this is a non-degeneracy condition). Fix a line $\xi_1 y_1 + \xi_2 y_2 = z$ in $\mathbb{C}^2$ and assume that the line intersects D in m different points $p_1, \dots, p_m$. Then the space E decomposes in direct sum as

$$E = E(p_1) \dotplus \cdots \dotplus E(p_m).$$

Let $P(p_j, \xi_1, \xi_2, z)$ be the projection onto $E(p_j)$ according to this decomposition. The restoration formula allows one to recover the complete characteristic function by

$$S(\xi_1, \xi_2, z) = \sum_{j=1}^{m} \widehat{S}(p_j) P(p_j, \xi_1, \xi_2, z).$$

A proof of this fact can be found in [LKMV95, Section 10.3].

4. Separating structures

This chapter contains the original part of this work. As we have stated in the Preface, our objective is to construct a structure, called separating structure, which can be compressed to a vessel and to which a vessel could be dilated.

First we deal with the topic of operator pools, which are structures built around two commuting selfadjoint operators and which have some resemblance to vessels. In particular, a discriminant curve can be assigned to them in the same way. Then we pass to separating structures, which are the principal construction of this chapter. First we deal with the affine case, to show which properties are just a consequence of linear algebra, and then we treat the orthogonal case and show how an orthogonal separating structure produces a pool. Then we prove that under some mild conditions, the discriminant curve of a separating structure is separated (see the Preface). Finally, we give the definition of the generalized compression and show its application to separating structures.

4.1. Operator pools

Definition. Let K be a Hilbert space, M a finite-dimensional Hilbert space, $\Phi : K \to M$ an operator, and A_1, A_2 two commuting selfadjoint operators on K. The tuple

$$Б = (A_1, A_2; K, \Phi, M; \sigma_1, \sigma_2, \gamma)$$

is called an operator pool if σ_j, γ are selfadjoint operators on M such that the following three term relationship holds:

$$\sigma_2 \Phi A_1 - \sigma_1 \Phi A_2 + \gamma \Phi = 0. \tag{4.1}$$

The operators σ_j are called rates and the operator γ is called gyration, as in the case of vessels (see Chapter 3). We define the principal subspace of the pool Б as

$$\widehat{K} = \bigvee_{k_1, k_2 \geq 0} A_1^{k_1} A_2^{k_2} \Phi^* M. \tag{4.2}$$

We say that the pool is irreducible if $K = \widehat{K}$. Typically, we will be considering irreducible pools.

Given a direction $\xi = (\xi_1, \xi_2) \in \mathbb{C}^2$, we will say that the direction ξ is nondegenerate if $\xi_1 \overline{\xi_2} \notin \mathbb{R}$. We will denote by Ξ the set of all nondegenerate directions:

$$\Xi = \{\xi = (\xi_1, \xi_2) : \xi_1 \overline{\xi_2} \notin \mathbb{R}\}.$$

For every fixed nondegenerate direction $\xi = (\xi_1, \xi_2) \in \Xi$, the operator

$$N_\xi = \xi_1 A_1 + \xi_2 A_2 \tag{4.3}$$

is normal and (4.1) is equivalent to

$$\alpha_\xi^* \Phi N_\xi + \alpha_\xi \Phi N_\xi^* + \gamma_\xi \Phi = 0, \tag{4.4}$$

where

$$\alpha_\xi = i(\xi_1\sigma_1 + \xi_2\sigma_2), \qquad \gamma_\xi = -2\,\mathrm{Im}(\overline{\xi_1}\xi_2)\gamma. \tag{4.5}$$

Now, we will construct a functional model for the pool Б using an L^2 space of M-valued functions. Let E_ξ be the spectral measure of N_ξ (see Appendix A.2). It is easy to see that

$$\widehat{K} = \bigvee_{k_1,k_2\geq 0} N_\xi^{k_1} N_\xi^{*k_2}\Phi^* M = \bigvee_{\substack{\Omega\subset\mathbb{C}\\ \Omega \text{ Borel}}} E_\xi(\Omega)\Phi^* M. \tag{4.6}$$

Here, the first equality is a direct consequence of (4.3). The second equality is true because the polynomials in z and $\overline{z}$ are uniformly dense in $\mathcal{C}(\sigma(N_\xi))$, and the operator $E_\xi(\Omega)$ is strong limit of operators $g_n(N_\xi)$, where $g_n \in \mathcal{C}(\sigma(N_\xi))$, by the properties of the spectral measure.

We consider the non-negative matrix-valued measure e_ξ given by

$$e_\xi(\Omega) = \Phi E_\xi(\Omega)\Phi^*, \qquad \Omega \subset \mathbb{C}. \tag{4.7}$$

Next, we define the space $L^2(e_\xi)$ of Borel functions $\mathbb{C} \to M$ with the scalar product

$$\langle f, g\rangle_{L^2(e_\xi)} = \int_{\mathbb{C}} \langle de_\xi(u) f(u), g(u)\rangle.$$

After factoring by the set $\{f : \|f\|_{L^2(e_\xi)} = 0\}$, it becomes a Hilbert space. We have $f = 0$ in $L^2(e_\xi)$ if and only if $de_\xi(u)f(u) = 0$ a.e. $u \in \mathbb{C}$.

Recall that for every bounded Borel function g on $\mathbb{C}$, we can define the operator $g(N_\xi)$ by means of the spectral functional calculus (see Appendix A.2). This allows us to construct a unitary $W_\xi : \widehat{K} \to L^2(e_\xi)$.

Proposition 4.1. *If* Б *is an irreducible pool, the operator W_ξ given by*

$$W_\xi g(N_\xi)\Phi^* m = g(\cdot)m$$

for $m \in M$ and g an arbitrary Borel function is well defined and extends by continuity to a unitary $W_\xi : \widehat{K} \to L^2(e_\xi)$. It also satisfies

$$(W_\xi N_\xi W_\xi^* h)(u) = uh(u), \qquad (W_\xi N_\xi^* W_\xi^* h)(u) = \overline{u}h(u) \tag{4.8}$$

and

$$\Phi W_\xi^* h = \int_{\mathbb{C}} de_\xi(u) h(u), \tag{4.9}$$

for every $h \in L^2(de_\xi)$.

Proof. First compute, for g, h Borel functions and $m, n \in M$,

$$\begin{aligned}\langle g(N_\xi)\Phi^* m, h(N_\xi)\Phi^* n\rangle &= \int_{\mathbb{C}} (\overline{h}g)(u)\langle dE_\xi(u)\Phi^* m, \Phi^* n\rangle = \int_{\mathbb{C}} (\overline{h}g)(u)\langle de_\xi(u)m, n\rangle \\ &= \langle g(\cdot)m, h(\cdot)n\rangle_{L^2(de_\xi)} \\ &= \langle W_\xi g(N_\xi)\Phi^* m, W_\xi h(N_\xi)\Phi^* n\rangle_{L^2(de_\xi)}.\end{aligned}$$

Using (4.6), since $\{g(\cdot)m : m \in M,\ g \text{ bounded Borel}\}$ spans $L^2(e_\xi)$, we see that W_ξ continues to a unitary.

To prove equations (4.8) and (4.9), observe that they are trivial for $h = g(\cdot)m$, with $m \in M$ and g bounded Borel, so they are also true for a general $h \in L^2(e_\xi)$ by continuity. □

The following Proposition will play an important role in the next section, because it will motivate the definition of the discriminant curve of the pool.

Proposition 4.2. *The following relation holds:*

$$(u\alpha_\xi^* + \overline{u}\alpha_\xi + \gamma_\xi)de_\xi(u) \equiv 0. \tag{4.10}$$

Proof. Multiply (4.4) by $W_\xi^* h$ on the right and use (4.8) and (4.9) to obtain

$$\int_{\mathbb{C}} (u\alpha_\xi^* + \overline{u}\alpha_\xi + \gamma_\xi)de_\xi(u)h(u) = 0.$$

Since this is true for every $h \in L^2(e_\xi)$, the Proposition follows. □

4.2. The discriminant curve

The affine algebraic curve

$$X_{\text{aff}} = \{(x_1, x_2) \in \mathbb{C}^2 : \det(x_1\sigma_2 - x_2\sigma_1 + \gamma) = 0\} \tag{4.11}$$

is called the discriminant curve of the pool. The discriminant curve is a real algebraic curve, equipped with the involution $*$ which sends $p = (x_1, x_2)$ to $p^* = (\overline{x_1}, \overline{x_2})$. The real part of the curve is

$$X_{\text{aff},\mathbb{R}} = \{p \in X_{\text{aff}} : p = p^*\} = X_{\text{aff}} \cap \mathbb{R}^2.$$

For a nondegenerate direction $\xi = (\xi_1, \xi_2) \in \Xi$, we introduce the coordinates (z_ξ, w_ξ) in $\mathbb{C}^2$. If $p = (x_1, x_2) \in \mathbb{C}^2$, we put

$$z_\xi(p) = \xi_1 x_1 + \xi_2 x_2, \qquad w_\xi(p) = \overline{\xi_1} x_1 + \overline{\xi_2} x_2.$$

Then we see that in these coordinates, the equation of X_{aff} rewrites as

$$\det(z_\xi \alpha_\xi^* + w_\xi \alpha_\xi + \gamma_\xi) = 0. \tag{4.12}$$

The involution $*$ can be written in these coordinates as $(z_\xi, w_\xi)^* = (\overline{w_\xi}, \overline{z_\xi})$. Using this and Proposition 4.2, we see that

$$\operatorname{supp} e_\xi \subset z_\xi(X_{\text{aff},\mathbb{R}}). \tag{4.13}$$

Moreover, if the pool Б is irreducible, we also have

$$\sigma(N_\xi) = \operatorname{supp} E_\xi = \operatorname{supp} e_\xi,$$

so we get

$$\sigma(N_\xi) \subset z_\xi(X_{\text{aff},\mathbb{R}}).$$

We will always assume that X_{aff} is a curve of full degree $\dim M$. It is easy to see that this happens if and only if

$$\det(x_1\sigma_2 - x_2\sigma_1) \not\equiv 0. \tag{4.14}$$

In this case, α_ξ is invertible for a general direction $\xi \in \mathbb{C}^2$ (just put $x_2 = -\xi_1$, $x_1 = \xi_2$ in (4.14)). This non-degeneracy condition (4.14) was already considered in the context of vessels in (3.15).

Whenever α_ξ is invertible, we define the operators

$$\Sigma_\xi = -\alpha_\xi^{-1}\alpha_\xi^*, \qquad D_\xi = -\alpha_\xi^{-1}\gamma_\xi. \tag{4.15}$$

The equation of X_{aff} can be rewritten as

$$\det(z_\xi \Sigma_\xi + D_\xi - w_\xi) = 0$$

(here and in the sequel we use the notation λ for the matrix λI). This means that $(z_\xi, w_\xi) \in X_{\text{aff}}$ if and only if $w_\xi \in \sigma(z_\xi \Sigma_\xi + D_\xi)$. We use this to define a projection-valued function Q_ξ on X_{aff}. If $p \in X_{\text{aff}}$, we put

$$Q_\xi(p) = \Pi_{w_\xi(p)}(z_\xi(p)\Sigma_\xi + D_\xi)$$

the Riesz projection of the matrix $z_\xi(p)\Sigma_\xi + D_\xi$ associated to the eigenvalue $w_\xi(p)$ (see Appendix A.1). By the properties of the Riesz projections, we get, for every $z_0 \in \mathbb{C}$, the direct sum decomposition

$$M = \sum_{\substack{p \in X_{\text{aff}} \\ z_\xi(p) = z_0}} Q_\xi(p) M. \tag{4.16}$$

We will also consider the projectivization X of the affine curve X_{aff}. We use projective coordinates $(\zeta_1 : \zeta_2 : \zeta_3)$ in $\mathbb{CP}^2$ and embed $\mathbb{C}^2$ in $\mathbb{CP}^2$ by

$$x_1 = \frac{\zeta_1}{\zeta_3}, \qquad x_2 = \frac{\zeta_2}{\zeta_3}.$$

The line $\zeta_3 = 0$ is the line at infinity. It will play an important role in the sequel. Since X_{aff} has degree $\dim M$, the projective curve X is

$$X = \{(\zeta_1 : \zeta_2 : \zeta_3) \in \mathbb{CP}^2 : \det(\zeta_1 \sigma_2 - \zeta_2 \sigma_1 + \zeta_3 \gamma) = 0\}.$$

The involution $*$ extends to $\mathbb{CP}^2$ by $(\zeta_1 : \zeta_2 : \zeta_3)^* = (\overline{\zeta_1} : \overline{\zeta_2} : \overline{\zeta_3})$, the curve X is a real projective curve, and its real part $X_{\mathbb{R}}$ is the set of points of X fixed by the involution.

If $\xi = (\xi_1, \xi_2) \in \Xi$ is a nondegenerate direction, we define the projective coordinates in $\mathbb{CP}^2$

$$\eta_{\xi,1} = \xi_1 \zeta_1 + \xi_2 \zeta_2, \qquad \eta_{\xi,2} = \overline{\xi_1}\zeta_1 + \overline{\xi_2}\zeta_2, \qquad \eta_{\xi,3} = \zeta_3.$$

In these coordinates, the equation of X is

$$\det(\eta_{\xi,1}\alpha_\xi^* + \eta_{\xi,2}\alpha_\xi + \eta_{\xi,3}\gamma_\xi) = 0.$$

The functions z_ξ and w_ξ extend to meromorphic functions on $\mathbb{CP}^2$ by

$$z_\xi = \frac{\eta_{\xi,1}}{\eta_{\xi,3}}, \qquad w_\xi = \frac{\eta_{\xi,2}}{\eta_{\xi,3}}.$$

We define the points at infinity of X by

$$X_\infty = \{p \in X : \zeta_3(p) = 0\}.$$

By the Fundamental Theorem of Algebra (or Bézout's Theorem about the number of intersections of two projective curves), X_∞ is a set of $\dim M$ points counting multiplicities. Indeed, for a general direction $\xi \in \mathbb{C}^2$, we can rewrite the equation of X as

$$\det(\eta_{\xi,1}\Sigma_\xi + \eta_{\xi,3}D_\xi - \eta_{\xi,2}) = 0.$$

From this, it follows that a point at infinity $p \in \mathbb{CP}^2$ with $\zeta_3(p) = 0$ is in X_∞ if and only if $(w_\xi / z_\xi)(p) \in \sigma(\Sigma_\xi)$.

Now we will construct $\widehat{X}$ the blow-up or desingularization of X. Assume that the polynomial $\det(x_1\sigma_2 - x_2\sigma_1 + \gamma)$ decomposes in irreducible factors over $\mathbb{C}[x_1, x_2]$ as

$$\det(x_1\sigma_2 - x_2\sigma_1 + \gamma) = \prod_{j=1}^{J} p_j(x_1, x_2)^{m_j},$$

where all the polynomials $p_j(x_1, x_2)$ are distinct. Then we say that X has J components, $p_j(x_1, x_2) = 0$ is the (affine) equation of the j-th component X_j, and m_j is the multiplicity of X_j. An affine point $p \in X_{\text{aff}}$ will be called regular if

$$\left.\frac{\partial}{\partial x_k}\right|_p \left(\prod_{j=1}^{J} p_j(x_1, x_2)\right) \neq 0$$

for $k = 1$ or $k = 2$. The set of regular points, which will be denoted by X_0, has a natural Riemann surface structure, and $X \setminus X_0$ is a finite collection of points. The blow-up $\widehat{X}$ is a compact Riemann surface with J connected components $\widehat{X}_j$ and a surjective continuous map $\pi_X : \widehat{X} \to X$ such that the restriction $\pi_X|\pi_X^{-1}(X_0)$ is an isomorphism of Riemann surfaces. The blow-up can be constructed by gluing a finite number of points to X_0 (see, for instance, [Mir95, Section III.2]). We put $\widehat{X}_0 = \pi_X^{-1}(X_0)$ and observe that $\widehat{X} \setminus \widehat{X}_0$ is finite.

The meromorphic functions z_ξ, w_ξ in $\mathbb{CP}^2$ induce meromorphic functions in $\widehat{X}$, which we will denote by the same letters:

$$z_\xi(p) = z_\xi(\pi_X(p)), \qquad w_\xi(p) = w_\xi(\pi_X(p)), \qquad p \in \widehat{X}.$$

The involution $*$ maps X_0 onto X_0, so it induces an antianalytic involution in $\widehat{X}$ (which we will also call $*$) by $\pi_X(p^*) = \pi_X(p)^*$ for $p \in X_0$, and then extending $*$ to all of $\widehat{X}$ by continuity. The real part of $\widehat{X}$, denoted by $\widehat{X}_\mathbb{R}$, is the set of points fixed by the involution. By definition, $\pi_X(\widehat{X}_\mathbb{R} \cap \widehat{X}_0) = X_\mathbb{R} \cap X_0$. However, the set $X_\mathbb{R}$ might be larger that $\pi_X(\widehat{X}_\mathbb{R})$ (although it will only differ by a finite number of points). Indeed, if $p \in X_\mathbb{R}$, then $*$ permutes the points in the fibre $\pi_X^{-1}(\{p\})$, but these points are not necessarily fixed by $*$ if the fibre has more than one point.

We define the points at infinity of $\widehat{X}$ by $\widehat{X}_\infty = \pi_X^{-1}(X_\infty)$. Note that every connected component $\widehat{X}_j$ contains points of $\widehat{X}_\infty$ (indeed degree(p_j) points).

The function Q_ξ induces a projection-valued meromorphic function on $\widehat{X}$ (which we will also denote by Q_ξ) defined by $Q_\xi(p) = Q_\xi(\pi_X(p))$ for $p \in \widehat{X}_0$.

4.3. Affine separating structures

Definition. Let N be an operator on a Hilbert space K. An affine separating structure for N is a direct sum decomposition

$$K = H_{0,-} \dotplus M_- \dotplus M_+ \dotplus H_{0,+} \tag{4.17}$$

such that the channel space

$$M = M_- + M_+$$

is finite dimensional and

$$NH_{0,-} \subset H_-, \quad NH_- \subset H_- + M_+, \quad NH_+ \subset H_+ + M_-, \quad NH_{0,+} \subset H_+, \tag{4.18}$$

where

$$H_- = H_{0,-} + M_-, \qquad H_+ = H_{0,+} + M_+.$$

According to the decomposition (4.17), we can write

$$N = \begin{bmatrix} * & \widetilde{R_{-2}} & 0 & 0 \\ \widetilde{T_{-1}} & \Lambda_{-1} & R_{-1} & 0 \\ 0 & T_0 & \Lambda_0 & \widetilde{R_0} \\ 0 & 0 & \widetilde{T_1} & * \end{bmatrix}. \tag{4.19}$$

The decomposition (4.17) also produces the dual decomposition

$$K = H'_{0,-} \dotplus M'_- \dotplus M'_+ \dotplus H'_{0,+}. \tag{4.20}$$

Here we put

$$H'_{0,-} = (M_- + M_+ + H_{0,+})^\perp$$

and analogously for the other subspaces:

$$M'_- = (H_{0,-} + M_+ + H_{0,+})^\perp, \qquad M'_+ = (H_{0,-} + M_- + H_{0,+})^\perp,$$
$$H'_{0,+} = (H_{0,-} + M_- + M_+)^\perp.$$

We can make the duality identifications $H'_{0,-} \cong (H_{0,-})^*$, and so on.

We denote by $P_{H_{0,-}}$, P_{M_-}, etc. the parallel projections corresponding to the summands in (4.17). The projections corresponding to (4.20) are $P_{H'_{0,-}} = P^*_{H_{0,-}}$, etc. We also define the channel operators

$$P_M = P_{M_-} + P_{M_+}, \qquad P_{M'} = P^*_M = P_{M'_-} + P_{M'_+},$$

the parallel projections

$$P_- = P_{H_-}, \qquad P_+ = P_{H_+},$$

and the corresponding parallel projections for the dual decomposition.

We define s, the compression of N to M:

$$s = P_M N | M.$$

We also define the operator $\alpha : M \to M$ by

$$P_+ N - N P_+ = \alpha P_M. \tag{4.21}$$

It is easy to check that α is well defined and

$$\alpha = \begin{bmatrix} 0 & -R_{-1} \\ T_0 & 0 \end{bmatrix}. \tag{4.22}$$

Now we can define the mosaic ν and the almost diagonalizing transform V by

$$\nu(z) = P_M (N - z)^{-1} P_+ (N - z) | M, \qquad z \notin \sigma(N). \tag{4.23}$$

$$(Vx)(z) = P_M (N - z)^{-1} x, \qquad x \in K, \ z \notin \sigma(N). \tag{4.24}$$

An affine separating structure is called *pure* if

$$K = \bigvee_{z \notin \sigma(N^*)} (N^* - z)^{-1} M.$$

It is easy to see that a structure is pure if and only if its associated almost diagonalizing transform V is injective. In this case, the transform V gives an analytic model for the structure.

If $\omega = (K, N, H_{0,-}, M_-, M_+, H_{0,+})$ is an affine separating structure, then

$$\omega^* = (K, N^*, H'_{0,-}, M'_-, M'_+, H'_{0,+}) \tag{4.25}$$

is also an affine separating structure, called the dual structure. We denote by ν_* and V_* the mosaic and the almost diagonalizing transform assigned to the dual structure.

Example 4.3. This example concerns the relation between separating structures and subnormal operators, and it will be continued by giving additional examples throughout the rest of this chapter. A subnormal operator $S \in \mathcal{B}(H)$ is, by definition, an operator having an extension (see Section 1.1) to a normal operator $N \in \mathcal{B}(K)$, with $K \supset H$. We say that S is pure if no nontrivial subspace of H reduces S to a normal operator. The normal extension N is called minimal if $K = \bigvee_{n\geq 0} N^{*n}H$. Every subnormal operator has a minimal normal extension. The subnormal operator S is said to be of finite type if its self-commutator $C = S^*S - SS^*$ has finite rank. See [Con91] for a treatment of the theory of subnormal operators.

A pure subnormal operator of finite type S and its minimal normal extension N can be used to construct a separating structure in the following way. The operator N and the space K of the structure will be the minimal normal extension N and the space on which it acts. The space H_+ is going to be H, the subspace on which S acts, and H_- is going to be $K \ominus H$. To do this, we start by putting $M_+ = CH$, which has finite dimension because S is of finite type. Then we put $H_{0,+} = H \ominus CH$.

The operator N has the structure

$$N = \begin{bmatrix} S'^* & 0 \\ X & S \end{bmatrix} \tag{4.26}$$

according to the decomposition $K = (K \ominus H) \oplus H$. The operator S' is pure subnormal and is called the dual of S. Using the fact that N is normal, we get the equalities

$$XX^* = S^*S - SS^* = C, \tag{4.27}$$
$$X^*X = S'^*S' - S'S'^* = C', \tag{4.28}$$

where C' is the self-commutator of S'. We note that

$$X(K \ominus H) = X\overline{X^*H} = CH = M_+.$$

Here the first equality comes from the fact that $\ker X = (K \ominus H) \ominus \overline{X^*H}$. We see that

$$C'(K \ominus H) = X^*M_+,$$

so that C' has finite rank and S' is pure subnormal of finite type.

We can define $M_- = C'(K \ominus H)$ and $H_{0,-} = (K \ominus H) \ominus M_-$. Now we have to check that conditions (4.18) are satisfied. We have

$$NH_- \subset S'^*(K \ominus H) + X(K \ominus H) \subset H_- + M_+.$$

The inclusion $NH_+ \subset H_+ + M_-$ is trivial. Indeed, $NH_+ \subset H_+$, which shows that $NH_{0,+} \subset H_+$ is also trivial. It remains to show that $NH_{0,-} \subset H_-$.

We have

$$\ker X^* = H \ominus X(K \ominus H) = H \ominus M_+ = H_{0,+}. \tag{4.29}$$

This implies $M_- = X^* M_+ = X^* H_+$. Hence, $\ker X = H_{0,-}$. It follows that

$$NH_{0,-} \subset S'^* H_{0,-} + XH_{0,-} = S'^* H_{0,-} \subset H_-.$$

Hence, we see that $\omega = (K, N, H_{0,-}, M_-, M_+, H_{0,+})$ is a separating structure. Moreover, the decomposition $K = H_{0,-} \oplus M_- \oplus M_+ \oplus H_{0,+}$ is orthogonal. Thus, ω is a particular kind of separating structure, called orthogonal separating structure, which will be introduced in the next section.

From what we have done, it also follows that X maps M_- onto M_+ and X^* maps M_+ onto M_-. Hence, the dimensions of M_- and M_+ coincide and $T_0 = P_{M_+} N|M_- = X|M_-$ is an isomorphism from M_- to M_+. This will play a role later. Also, note that the operator R_{-1} from (4.19) is 0.

Another important fact about subnormal operators is that $M_+ = CH$ is invariant for S^* (this is easy to prove; see [Con91, Section II.3, exercises 6 and 7]). The operator $(S^*|M_+)^*$ is just the operator Λ_0 in (4.19). The pair of operators (C, Λ_0) determines the subnormal operator S. The associated algebraic curve constructed in [Yak98a] is given in terms of this pair (there, the operator Λ_0 is denoted just by Λ). ♠

The following Theorem lists the main properties of the almost diagonalizing transform and the mosaic.

Theorem 4.4. *Suppose that ω is an affine separating structure. Denote by VK the image of the almost diagonalizing transform V, endowed with the Hilbert space structure inherited from K (i.e., the unique structure that makes $V : K \ominus \ker V \to VK$ a unitary). Then the following statements are true:*

(i) VK is a reproducing kernel Hilbert space of M-valued holomorphic functions on $\Omega = \widehat{\mathbb{C}} \backslash \sigma(N)$ which vanish at ∞.

(ii) ν is a holomorphic projection-valued function on Ω. The operator P_ν defined by $(P_\nu f)(z) = \nu(z) f(z)$ for $f \in VK$ is a parallel projection on VK.

(iii) V almost diagonalizes N:

$$(VNx)(z) = z(Vx)(z) - [z(Vx)(z)]|_{z=\infty}.$$

(iv) V transforms the resolvent operator $(N-z)^{-1}$ into the operator $f(u) \mapsto (f(u)-f(z))/(u-z)$:

$$(V(N-z)^{-1}x)(u) = \frac{(Vx)(u) - (Vx)(z)}{u-z}.$$

(v) V transforms P_+ into the operator P_ν, i.e., $VP_+ = P_\nu V$.

(vi) We have the following formula for the mosaic:

$$\nu(z) = P_M (N-z)^{-1} \alpha + P_{M_+}.$$

In particular, $\nu(\infty) = P_{M_+}$ and

$$\nu(z)m = (V\alpha m)(z) + P_{M_+} m, \qquad m \in M.$$

(vii) We have

$$(V(N-z)^{-1}\alpha m)(u) = \frac{\nu(u)-\nu(z)}{u-z}m, \qquad m \in M.$$

Proof. Statement (i) is clear from the definition of the transform V. To prove (ii), we first observe that

$$P_+(N-z)P_M = P_+(N-z)(I-P_{H_{0,+}}) = P_+(N-z) - (N-z)P_{H_{0,+}}.$$

Hence,

$$\begin{aligned} P_M(N-z)^{-1}P_+(N-z)P_M(N-z)^{-1} &= P_M(N-z)^{-1}P_+ - P_M P_{H_{0,+}}(N-z)^{-1} \\ &= P_M(N-z)^{-1}P_+. \end{aligned} \tag{4.30}$$

Multiplying this equality by $x \in K$ on the right, we get (v). Also, multiplying by $P_+(N-z)|M$ on the right, we get $\nu^2(z) = \nu(z)$. This implies that ν is projection-valued. The operator P_ν is bounded by (v), and hence, it is a parallel projection.

To prove (iii) we observe that

$$(VNx)(z) = P_M(N-z)^{-1}Nx = P_M x + z(Vx)(z).$$

It is easy to check from the definition that $[z(Vx)(z)]|z=\infty = -P_M x$.

Part (iv) is obtained directly from the definition of V. To check (vi), we compute

$$\begin{aligned} \nu(z) &= P_M(N-z)^{-1}P_+(N-z)|M \\ &= P_M(N-z)^{-1}[P_+(N-z)-(N-z)P_+]|M + P_M(N-z)^{-1}(N-z)P_+|M \\ &= P_M(N-z)^{-1}\alpha + P_{M_+}. \end{aligned}$$

The remaining of (vi) is obvious and (vii) is a direct consequence of (iv) and (vii). $\square$

Now we will do a brief geometric study of the mosaic ν. Define, for $z \notin \sigma(N)$, the following two subspaces of M:

$$\widetilde{F}(z) = P_M(N-z)^{-1}H_+, \qquad \widetilde{G}(z) = P_M(N-z)^{-1}H_-. \tag{4.31}$$

Proposition 4.5. *The space M decomposes in direct sum as*

$$M = \widetilde{F}(z) \dotplus \widetilde{G}(z)$$

for every $z \notin \sigma(N)$. The operator $\nu(z)$ is the projection onto $\widetilde{F}(z)$ parallel to $\widetilde{G}(z)$.

Proof. If $m \in M$, put $h_- = P_-(N-z)m$, and $h_+ = P_+(N-z)m$. Then $h_- \in H_-$ and $h_+ \in H_+$. It follows that $m_- = P_M(N-z)^{-1}h_- \in \widetilde{G}(z)$, and $m_+ = P_M(N-z)^{-1}h_+ \in \widetilde{F}(z)$. Moreover, $m_- + m_+ = m$. This shows that $M = \widetilde{F}(z) + \widetilde{G}(z)$.

We must check that $\widetilde{F}(z) \cap \widetilde{G}(z) = 0$. To do this, take $m \in \widetilde{F}(z)$ and write $m = P_M(N-z)^{-1}h_+$, where $h_+ \in H_+$. Multiplying (4.30) by h_+ on the right and using formula (4.23) for the mosaic $\nu(z)$, we get $\nu(z)m = m$. Similarly, we see that if $m \in \widetilde{G}(z)$, then $\nu(z)m = 0$. This shows that $\widetilde{F}(z) \cap \widetilde{G}(z) = 0$ and that $\nu(z)$ is the projection onto $\widetilde{F}(z)$ parallel to $\widetilde{G}(z)$. $\square$

An alternative proof of this Proposition can be given by observing that $\widetilde{F}(z) = V(z)H_-$ and $\widetilde{G}(z) = V(z)H_+$. Then it is enough to use Theorem 4.4 (v).

Here and in the sequel we will use the notation 1 for the identity matrix on a finite-dimensional Hilbert space. If we put

$$E_0^2(\nu) = \{v \in VK : v(z) \in \nu(z)M = \widetilde{F}(z)\}, \tag{4.32}$$
$$E_0^2(1-\nu) = \{v \in VK : v(z) \in (1-\nu(z))M = \widetilde{G}(z)\}, \tag{4.33}$$

then Theorem 4.4 (v) proves that

$$VH_- = E_0^2(1-\nu), \qquad VH_+ = E_0^2(\nu).$$

Example 4.6. A mosaic function $\mu(z)$ for a subnormal operator S appears in [Yak98a]. This function is holomorphic on $\widehat{\mathbb{C}}\setminus\sigma(N)$ and its values are parallel projections on M_+. In the notation of separating structures, it is defined by

$$\mu(z) = \alpha P_M(N-z)^{-1}|M_+,$$

where N is the minimal normal extension of S. Using Theorem 4.4 (vi), we see that

$$\alpha\nu(z) = \mu(z)\alpha, \tag{4.34}$$

because $R_{-1} = 0$ in the case of a subnormal operator (see Example 4.3).

The spaces $\mu(z)M_+$ and $(1-\mu(z))M_+$ play an important role in [Yak98b]. Using (4.34), we see that

$$\mu(z)M_+ = \alpha\widetilde{F}(z), \qquad (1-\mu(z))M_+ = \alpha\widetilde{G}(z).$$

Also, an almost diagonalizing transform $\widetilde{U}$ appears in [Yak98b]. It is defined by

$$(\widetilde{U}x)(z) = P_+NP_-(N-z)^{-1}x,$$

and plays a similar role to V. For instance, $\widetilde{U}$ transforms the projection P_+ into the operator of multiplication by $\mu(z)$ (c.f. Theorem 4.4 (v)). We see that

$$(\widetilde{U}x)(z) = \alpha(Vx)(z).$$

Hence, the operator α, which in the subnormal case maps M onto M_+, can be used to pass from many of our constructions to the analogue constructions in [Yak98a, Yak98b]. ♠

The following two Lemmas are not used later but could be helpful to keep in mind. To interpret these Lemmas, one should know that in the case considered in the following sections, the operator α typically will be invertible.

Lemma 4.7. *Assume that ω is pure. Then $\nu(z)m$ is constant if and only if $\alpha m = 0$.*

Proof. If $\nu(z)m$ is constant, then Theorem 4.4 (vii) shows that $V(N-z)^{-1}\alpha m \equiv 0$, so that $\alpha m = 0$, because ω is pure, and hence, V is injective. Conversely, if $\alpha m = 0$, Theorem 4.4 (vi) shows that $\nu(z)m = P_{M_+}m$. □

Lemma 4.8. *There is the following relation between the mosaic ν of the structure ω and the mosaic ν_* of the dual structure ω^*:*

$$(1-\nu_*^*(\overline{z}))\alpha = \alpha\nu(z).$$

Proof. Recall that the dual structure ω^* is defined by (4.25). We will denote by α_* the operator defined by (4.21) with the dual structure ω^* in place of ω, i.e., the operator $\alpha_* : M' \to M'$ defined by

$$P'_+ N^* - N^* P'_+ = \alpha_* P_{M'}.$$

We see that $\alpha_* = -\alpha^*$.

By Theorem 4.4 (vi),

$$\nu_*(z) = P_{M'_+} - P_{M'}(N^* - z)^{-1}\alpha^*.$$

We get

$$\begin{aligned}\nu_*^*(\overline{z})\alpha + \alpha\nu(z) &= P_{M_+}\alpha - \alpha(N-z)^{-1}P_M\alpha + \alpha P_{M+} + \alpha P_M(N-z)^{-1}\alpha \\ &= P_{M_+}\alpha + \alpha P_{M_+} = \alpha.\end{aligned}$$

Here the last equality can be seen by (4.22). This proves the Lemma. □

4.4. Orthogonal separating structures

Definition. We say that an affine separating structure ω is orthogonal if N is normal and the decomposition of K is orthogonal:

$$K = H_{0,-} \oplus M_- \oplus M_+ \oplus H_{0,+}. \tag{4.35}$$

Given an orthogonal decomposition (4.35), A_1, A_2 two commuting selfadjoint operators on K satisfying

$$A_j H_{0,-} \subset H_-, \quad A_j H_- \subset H_- \oplus M_+, \quad A_j H_+ \subset H_+ \oplus M_-, \quad A_j H_{0,+} \subset H_-, \tag{4.36}$$

for $j = 1, 2$, and a direction $\xi = (\xi_1, \xi_2) \in \mathbb{C}^2$, the operator $N_\xi = \xi_1 A_1 + \xi_2 A_2$ forms an orthogonal separating structure ω_ξ with respect to (4.35). Conversely, given an operator N forming an orthogonal separating structure, one can put $A_1 = \operatorname{Re} N$, $A_2 = \operatorname{Im} N$ and form the family of structures $\{\omega_\xi : \xi \in \mathbb{C}^2\}$ as above.

It will be convenient to think of orthogonal separating structures in this way, as a family of structures $\{\omega_\xi : \xi \in \mathbb{C}^2\}$ generated by two commuting selfadjoint operators A_1, A_2.

For each direction $\xi \in \mathbb{C}^2$, we obtain a separating structure ω_ξ with associated operator $N_\xi = \xi_1 A_1 + \xi_2 A_2$. Hence, we can apply the results of the preceding section to the structure ω_ξ. We will mark with the subscript $_\xi$ the objects of the preceding section constructed for the operator N_ξ. Therefore, we will write V_ξ, ν_ξ, $\Lambda_{-1\xi}$, $\Lambda_{0\xi}$, $R_{-1\xi}$, $T_{0\xi}$, etc.

In the next Theorem, we relate the orthogonal separating structure $\{\omega_\xi\}$ with the notion of pool given in Section 4.1.

Theorem 4.9. *If $\{\omega_\xi\}$ is an orthogonal separating structure, we can construct a pool*

$$Б = (A_1, A_2; K, \Phi, M; \sigma_1, \sigma_2, \gamma),$$

by defining $\Phi = P_M$ and

$$\begin{aligned}\sigma_j P_M &= -i(P_+ A_j - A_j P_+), \qquad j = 1, 2, \\ \gamma P_M &= i(A_1 P_+ A_2 - A_2 P_+ A_1).\end{aligned}$$

The operators α_ξ defined in (4.5) coincide with the operators α_ξ defined by using (4.21) for $N = N_\xi$. Moreover, the operator γ_ξ defined in (4.5) can be computed as

$$\gamma_\xi = -(\alpha_\xi^* s_\xi + \alpha_\xi s_\xi^*). \tag{4.37}$$

Proof. Conditions (4.36) imply that σ_j are well defined. If $T = i(A_1P_+A_2 - A_2P_+A_1)$, then (4.36) implies that $K \ominus M \subset \ker T$. Since T is selfadjoint, then $M \supset TK$, so γ is also well defined. The operators σ_j and γ are clearly self adjoint. We compute

$$\sigma_2P_MA_1 - \sigma_1P_MA_2 = -i(P_+A_2 - A_2P_+)A_1 + i(P_+A_1 - A_1P_+)A_2 = -\gamma P_M,$$

so Б is a pool.

A simple computation yields that the operator α_ξ defined by (4.5) coincides with the operator α_ξ appearing in (4.21). To check (4.37), just restrict (4.4) to M. □

Example 4.10. As we already commented in Example 4.3, the separating structure generated by a subnormal operator S is orthogonal. Now we will compute the discriminant curve of its associated pool, according to Theorem 4.9. We fix $\xi = (1, i)$ so that the operator N_ξ is just the minimal normal extension N. Thus, we will omit the subscript ξ in the rest of this example.

We have,

$$\alpha = \begin{bmatrix} 0 & 0 \\ T_0 & 0 \end{bmatrix}, \qquad s = \begin{bmatrix} \Lambda_{-1} & 0 \\ T_0 & \Lambda_0 \end{bmatrix}.$$

Using formula (4.37) for γ, we see that the equation for the discriminant curve (4.12) is

$$\begin{aligned}
\det(z\alpha^* + w\alpha + \gamma) &= \det\left(\begin{bmatrix} -T_0^*T_0 & zT_0^* - T_0^*\Lambda_0 \\ wT_0 - T_0\Lambda_{-1}^* & -T_0T_0^* \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} T_0^* & 0 \\ 0 & T_0 \end{bmatrix}\begin{bmatrix} -T_0 & z - \Lambda_0 \\ w - \Lambda_{-1}^* & -T_0^* \end{bmatrix}\right) \\
&= |\det T_0|^2 \det\left(\begin{bmatrix} -T_0 & z - \Lambda_0 \\ w - \Lambda_{-1}^* & -T_0^* \end{bmatrix}\right) \\
&= |\det T_0|^2 \det T_0 \det(-T_0^* + (w - \Lambda_{-1}^*)T_0^{-1}(z - \Lambda_0)) \\
&= |\det T_0|^2 \det(-T_0T_0^* + T_0(w - \Lambda_{-1}^*)T_0^{-1}(z - \Lambda_0)) \\
&= -|\det T_0|^2 \det(C - (w - \Lambda_0^*)(z - \Lambda_0)).
\end{aligned}$$

Here we have used $T_0T_0^* = C$, which comes from (4.27) and $T_0\Lambda_{-1}^* = \Lambda_0T_0$, which is obtained using the fact that N is normal (and $R_{-1} = 0$).

The equation for the discriminant curve associated to the subnormal operator S in [Yak98a, Yak98b] was precisely

$$\det(C - (w - \Lambda_0^*)(z - \Lambda_0)) = 0.$$

Therefore, this shows that the discriminant curve of the corresponding pool is the same curve. ♠

The next Proposition relates the concepts of purity of a separating structure and irreducibility of a pool.

Proposition 4.11. *Let $\{\omega_\xi\}$ be an orthogonal separating structure and* Б *its associated pool, according to Theorem 4.9. Assume that the discriminant curve X_{aff} of* Б *is not all of $\mathbb{C}^2$. Then the following statements are equivalent:*

(i) The separating structure ω_ξ is pure for some nondegenerate direction $\xi \in \Xi$.

(ii) The separating structure ω_ξ is pure for every nondegenerate direction $\xi \in \Xi$.

(iii) The pool Б *is irreducible.*

Proof. First we show (i) $\Rightarrow$ (iii), so we assume that ω_ξ is pure for a certain direction $\xi \in \Xi$. This means that the set

$$K_0 = \{(N_\xi^* - w)^{-1} m : m \in M, w \notin \sigma(N^*)\}$$

spans K. Since the function $(\overline{z} - w)^{-1}$, $w \notin \sigma(N^*)$ can be approximated uniformly in $\sigma(N)$ by polynomials in z and $\overline{z}$ by the Stone-Weierstrass theorem (see, for instance, [Rud91, Theorem 5.7]), it follows that every member of the set K_0 can approximated by some $p(N_\xi, N_\xi^*)m$, where p is a polynomial in two variables. By (4.6), we see that $K_0 \subset \widehat{K}$, where $\widehat{K}$ is the principal subspace of Б. This means that $\widehat{K}$ must also span K, so that $\widehat{K} = K$ and the pool Б is irreducible.

Now we show (iii) $\Rightarrow$ (ii), so we assume that Б is irreducible and take an arbitrary direction $\xi \in \Xi$. The set

$$K_1 = \{p(N_\xi, N_\xi^*)m : m \in M, p \in \mathbb{C}[z, \overline{z}]\}$$

spans K. Since the discriminant curve X is not all of $\mathbb{C}^2$ and the spectrum of N_ξ lies in $z_\xi(X_{\mathrm{aff},\mathbb{R}})$, we see that $\sigma(N_\xi)$ has area 0. By the Hartogs-Rosenthal theorem (see [Con91, Theorem V.3.6]), every continuous function on $\sigma(N_\xi)$ can be uniformly approximated by rational functions with poles outside $\sigma(N_\xi)$. This applies to any polynomial $p(z, \overline{z})$. Since every rational function with poles outside $\sigma(N_\xi)$ can be approximated uniformly in $\sigma(N_\xi)$ by a linear combination of functions of the form $(z - w_k)^{-1}$, $z_k \notin \sigma(N_\xi)$, we see that the vector $p(N_\xi, N_\xi^*)m$ can be approximated by a linear combination of vectors $(N_\xi - w_k)^{-1}m$. This shows that if

$$K_2 = \bigvee_{w \notin \sigma(N_\xi)} (N_\xi - w)^{-1} M,$$

then $K_1 \subset K_2$. Hence, we get $K_2 = K$. This implies that $\omega_{\overline{\xi}}$ is pure, because $N_{\overline{\xi}}^* = N_\xi$. Therefore, we get (ii), because $\xi \in \Xi$ was arbitrary.

The remaining implication (ii) $\Rightarrow$ (i) is trivial. □

Example 4.12. Consider a pure subnormal operator of finite type S, the separating structure it generates according to Example 4.3, and its associated pool Б. We will show that the purity of S implies the irreducibility of Б.

Put $G_+ = P_+(K \ominus \widehat{K})$. Since $M_+ \subset \widehat{K}$, we have $G_+ \subset H_{0,+}$. Using this, it is easy to check that G_+ is invariant for N and N^*. We have $N|G_+ = S|G_+$, because $N|H_+ = S$. Moreover, (4.26) and (4.29) imply that $N^*|H_{0,+} = S^*|H_{0,+}$. Hence, $N^*|G_+ = S^*|G_+$. It follows that G_+ reduces S to a normal operator. Since S is pure, it must be $G_+ = 0$.

Similarly, one can prove that $G_- = P_-(K \ominus \widehat{K}) = 0$. One has to follow the reasoning above interchanging N and N^* and using the pure subnormal operator $S' = N^*|H_-$ instead of S. Since we have $G_- = G_+ = 0$, we get $K \ominus \widehat{K} = 0$, so the pool Б is irreducible. ♠

Now we will relate the analytic model for the separating structure, constructed in terms of the almost diagonalizing transform V_ξ and the L^2 model for the pool, constructed in terms of the transform W_ξ, by means of the Cauchy operators. First, observe that there exists a positive scalar measure ρ_ξ and a matrix-valued ρ_ξ-measurable function $\mathcal{E}_\xi$ such that

$$de_\xi(z) = \mathcal{E}_\xi(z) d\rho_\xi(z),$$

The Cauchy operators are defined by

$$(\mathcal{K}_\xi f)(z) = \int_{\mathbb{C}} \frac{f(u)}{u - z} d\rho_\xi(u),$$
$$(\overline{\mathcal{K}}_\xi f)(z) = \int_{\mathbb{C}} \frac{f(u)}{\overline{u} - \overline{z}} d\rho_\xi(u).$$

We will also denote by $\mathcal{E}_\xi$ the operator of multiplication by $\mathcal{E}_\xi(z)$ in $L^2(de_\xi)$.

Proposition 4.13. *We have the following relation between the transforms of an orthogonal separating structure and its associated pool:*

$$V_\xi = \mathcal{K}_\xi \mathcal{E}_\xi W_\xi$$

Proof. For $m \in M$, g bounded Borel and $x = g(N_\xi)m$, we have

$$(V_\xi x)(z) = P_M(N_\xi - z)^{-1} g(N_\xi)m = P_M \int_{\mathbb{C}} \frac{g(u)}{u-z} dE_\xi(u)m = (\mathcal{K}_\xi \mathcal{E}_\xi W_\xi x)(z).$$

The Proposition follows by density. $\square$

Proposition 4.14. *The mosaic ν_ξ has the following integral representation:*

$$\nu_\xi(z) = P_{M_+} + \int_{\mathbb{C}} \frac{de_\xi(u)\alpha_\xi}{u-z}.$$

Proof. It suffices to observe that

$$P_M(N_\xi - z)^{-1}|M = \int_{\mathbb{C}} \frac{de_\xi(u)}{u-z}$$

and to use Theorem 4.4 (vi). $\square$

The next Lemma is a bit technical but it will be very useful later. Recall that Q_ξ was the projection-valued function defined on the discriminant curve (see Section 4.2).

Lemma 4.15. *If α_ξ is invertible, we have*

$$\alpha_\xi \nu_\xi(z) \alpha_\xi^{-1}(z\alpha_\xi^* + \gamma_\xi) = (z\alpha_\xi^* + \gamma_\xi)\nu_\xi(z), \tag{4.38}$$

Therefore $\nu_\xi(z)$ commutes both with $\alpha_\xi^{-1}(z\alpha_\xi^* + \gamma_\xi)$ and with $Q_\xi(p)$, for those $p \in X_{aff}$ such that $z_\xi(p) = z$.

Proof. To prove (4.38), using Proposition 4.14, we have to check that

$$\begin{aligned} \alpha_\xi P_{M_+} \alpha_\xi^{-1}(z\alpha_\xi^* + \gamma) + \int_{\mathbb{C}} \frac{\alpha_\xi de_\xi(u)(z\alpha_\xi^* + \gamma_\xi)}{u-z} \\ = (z\alpha_\xi^* + \gamma_\xi)P_{M_+} + \int_{\mathbb{C}} \frac{(z\alpha_\xi^* + \gamma_\xi)de_\xi(u)\alpha_\xi}{u-z}. \end{aligned}$$

Now we use the identities $\alpha_\xi P_{M_+} \alpha_\xi^{-1} = P_{M_-}$, and $P_{M_-}\alpha_\xi^* = \alpha_\xi^* P_{M_+}$, and rearrange terms to see that the equation above is equivalent to

$$P_{M_-}\gamma_\xi - \gamma_\xi P_{M_+} = \int_{\mathbb{C}} \frac{(z\alpha_\xi^* + \gamma_\xi)de_\xi(u)\alpha_\xi - \alpha_\xi de_\xi(u)(z\alpha_\xi^* + \gamma_\xi)}{u-z}. \tag{4.39}$$

Using (4.10) and the relation obtained from it by taking adjoints, we get

$$\gamma_\xi de_\xi(u) = -(u\alpha_\xi^* + \overline{u}\alpha_\xi)de_\xi(u), \qquad de_\xi(u)\gamma_\xi = -de_\xi(u)(u\alpha_\xi^* + \overline{u}\alpha_\xi).$$

Substituting these identities into the right hand side of (4.39), we see that it equals

$$\int_{\mathbb{C}} \alpha_\xi de_\xi(u)\alpha_\xi^* - \alpha_\xi^* de_\xi(u)\alpha_\xi = \alpha_\xi\alpha_\xi^* - \alpha_\xi^*\alpha_\xi.$$

Hence we just need to check that

$$P_{M_-}\gamma_\xi - \gamma_\xi P_{M_+} = \alpha_\xi\alpha_\xi^* - \alpha_\xi^*\alpha_\xi.$$

This is an easy computation with matrices, using (4.37), (4.19) and (4.22).

From (4.38), it is obvious that $\nu_\xi(z)$ and $\alpha_\xi^{-1}(\alpha_\xi^* z + \gamma_\xi)$ commute. The fact that $\nu_\xi(z)$ and $Q_\xi(p)$ commute is a consequence of the definition of $Q_\xi(p)$ and the properties of the Riesz-Dunford calculus (see Appendix A.1). □

Proposition 4.16. *If α_ξ is invertible, we have*

$$P_M(N_\xi - z)^{-1}(N_\xi^* - \overline{w})^{-1}P_M = (\gamma_\xi + z\alpha_\xi^* + \overline{w}\alpha_\xi)^{-1}(P_M - \alpha_\xi\nu(z)\alpha_\xi^{-1}P_M - \nu_\xi^*(w)P_M),$$

for every pair (z,w) such that $z,w \notin \sigma(N)$ and $(z,\overline{w}) \notin X_{aff}$.

Proof. First rewrite (4.4) as

$$\alpha_\xi^* P_M(N_\xi - z) + \alpha P_M(N_\xi^* - \overline{w}) + (\gamma_\xi + z\alpha_\xi^* + \overline{w}\alpha_\xi)P_M = 0.$$

Multiplying by $(N_\xi - z)^{-1}(N_\xi^* - \overline{w})^{-1}P_M$ on the right and rearranging terms, we get

$$\begin{aligned} &P_M(N_\xi - z)^{-1}(N_\xi^* - \overline{w})^{-1}P_M = \\ &\quad -(\gamma_\xi + z\alpha_\xi^* + \overline{w}\alpha_\xi)^{-1}[\alpha_\xi^* P_M(N_\xi^* - \overline{w})^{-1}P_M + \alpha_\xi P_M(N_\xi - z)^{-1}P_M]. \end{aligned} \tag{4.40}$$

By Theorem 4.4 (vi),

$$\alpha_\xi^* P_M(N_\xi^* - \overline{w})^{-1}P_M = \nu_\xi^*(w)P_M - P_{M_+}.$$

Also,

$$\alpha_\xi P_M(N_\xi - z)^1 P_M = \alpha_\xi(\nu_\xi(z) - P_{M_+})\alpha_\xi^{-1}P_M = \alpha_\xi\nu_\xi(z)\alpha_\xi^{-1}P_M - P_{M_-},$$

because $\alpha P_{M_+} = P_{M_-}\alpha$. The Proposition follows by substituting these two equalities into (4.40). □

This Proposition implies that the data $\alpha_\xi, \gamma_\xi, \nu_\xi$ completely determines the separating structure $\{\omega_\xi\}$ whenever the structure is pure. Indeed, we see from the Proposition that the inner product

$$\langle (N_\xi^* - \overline{w})^{-1}m, (N_\xi^* - \overline{z})^{-1}m' \rangle \qquad m, m' \in M,\ z, w \notin \sigma(N_\xi),$$

depends only on $\alpha_\xi, \gamma_\xi, \nu_\xi$ and m, m', z, w.

If we have two pure structures $\{\omega_\xi\}$ and $\{\widetilde{\omega}_\xi\}$ with the same data $\alpha_\xi, \gamma_\xi, \nu_\xi$, then the operator Z defined by

$$Z(N_\xi^* - \overline{w})^{-1}m = (\widetilde{N}_\xi^* - \overline{w})^{-1}m$$

continues to a unitary operator. By the Riesz-Dunford functional calculus (see Appendix A.1), we see that

$$ZN_\xi^* x = \widetilde{N}_\xi^* Zx,$$

for vector of the form $x = (N_\xi^* - \overline{w})^{-1}m$, and hence for every $x \in K$ by density. This shows that N_ξ and $\widetilde{N}_\xi$ are unitarily equivalent.

Below we will see that in many cases, the mosaic function ν_ξ can be computed from the matrices α_ξ and γ_ξ alone by using the so called restoration formula and the discriminant curve (which, of course, is defined solely in terms of α_ξ and γ_ξ).

Our goal now is to define the two *halves* of $\widehat{X}$. This is a partition of $\widehat{X} \setminus \widehat{X}_\mathbb{R}$ in two sets $\widehat{X}_-$, $\widehat{X}_+$ such that $*(\widehat{X}_-) = \widehat{X}_+$, and such that we can recover the mosaic ν_ξ by the restoration formula

$$\nu_\xi(z) = \sum_{\substack{p \in \widehat{X}_+ \\ z_\xi(p) = z}} Q_\xi(p),$$

where Q_ξ is the projection-valued meromorphic function on $\widehat{X}$ that was defined in Section 4.2.

First recall the definition of Σ_ξ from (4.15) and note that

$$\Sigma_\xi = \begin{bmatrix} \Sigma_\xi^- & 0 \\ 0 & \Sigma_\xi^+ \end{bmatrix}, \qquad \text{where } \Sigma_\xi^- = T_{0\xi}^{-1} R_{-1\xi}^*, \text{ and } \Sigma_\xi^+ = R_{-1\xi}^{-1} T_{0\xi}^*.$$

Hence, $\sigma(\Sigma_\xi) = \sigma(\Sigma_\xi^-) \cup \sigma(\Sigma_\xi^+)$ and the map $\lambda \mapsto \overline{\lambda}^{-1}$ interchanges $\sigma(\Sigma_\xi^-)$ and $\sigma(\Sigma_\xi^+)$.

We will assume from now on that

$$\sigma(\Sigma_\xi^-) \cap \sigma(\Sigma_\xi^+) = \emptyset. \tag{S}$$

This happens, for instance, if X_∞ is a set of $\dim M$ different points, because then all the eigenvalues of Σ_ξ are distinct.

We define the meromorphic function λ_ξ on $\widehat{X}$ by

$$\lambda_\xi(p) = \left(\frac{w_\xi}{z_\xi}\right)(p), \qquad p \in \widehat{X}.$$

If (S) holds, we can partition $\widehat{X}_\infty = \widehat{X}_\infty^- \cup \widehat{X}_\infty^+$, where we put

$$\widehat{X}_\infty^\pm = \{p \in \widehat{X}_\infty : \lambda_\xi(p) \in \sigma(\Sigma_\xi^\pm)\}$$

(recall that if $p \in \widehat{X}_\infty$, then $\lambda_\xi(p) \in \sigma(\Sigma_\xi)$).

Now we give a Lemma which relates the behaviour of $\nu_\xi(z_\xi(p))$ and $Q_\xi(p)$ for p near $\widehat{X}_\infty$.

Lemma 4.17. *Define*

$$\begin{aligned} \varphi_+(p) &= [1 - \nu_\xi(z_\xi(p))]Q_\xi(p), \qquad p \in \widehat{X} \setminus z_\xi^{-1}(\sigma(N_\xi)) \\ \varphi_-(p) &= \nu_\xi(z_\xi(p))Q_\xi(p), \qquad p \in \widehat{X} \setminus z_\xi^{-1}(\sigma(N_\xi)). \end{aligned}$$

If $p_0 \in \widehat{X}_\infty^+$, then φ_+ vanishes in a neighbourhood of p_0. If $p_0 \in \widehat{X}_\infty^-$, then φ_- vanishes in a neighbourhood of p_0.

Proof. Recall that if $p \in \widehat{X}$ is such that $z_\xi(p) \neq 0$, then $\lambda_\xi(p) \in \sigma(\Sigma_\xi + z_\xi(p)^{-1}D_\xi)$. Let $V \subset \mathbb{C}$ be an open disk with centre $\lambda_\xi(p_0)$ and such that $\overline{V}$ does not contain any other eigenvalue of Σ_ξ.

If q is a point in $\widehat{X}$ and $\lambda_\xi(q) \in V$, let Δ_q be a simple closed curve in $\mathbb{C}$ around $\lambda_\xi(q)$ which is positively oriented, is contained entirely inside V, and is such that no other eigenvalue of $\Sigma_\xi + z_\xi(q)^{-1}D_\xi$ lies inside Δ_q. Recall that if such point q satisfies $q \in \widehat{X}_0$ and $z_\xi(q) \neq 0$, then

$$Q_\xi(q) = \Pi_{\lambda_\xi(q)}(\Sigma_\xi + z_\xi(q)^{-1}D_\xi) = \frac{1}{2\pi i}\int_{\Delta_q} \left(\lambda - \Sigma_\xi - z_\xi(q)^{-1}D_\xi\right)^{-1} d\lambda.$$

Put

$$L(z) = \{q \in \widehat{X} \ : \ z_\xi(q) = z, \ \lambda_\xi(q) \in V\}.$$

For $z \in \widehat{\mathbb{C}}$ near ∞, no eigenvalue of the matrix $\Sigma_\xi + z^{-1}D_\xi$ lies on ∂V, by continuity of the spectrum. Hence, for a general $z \in \widehat{\mathbb{C}}$ near ∞,

$$\sum_{q \in L(z)} Q_\xi(q) = \frac{1}{2\pi i}\int_{\partial V} \left(\lambda - \Sigma_\xi - z^{-1}D_\xi\right)^{-1} d\lambda.$$

Since

$$\Pi_{\lambda_\xi(p_0)}(\Sigma_\xi) = \frac{1}{2\pi i}\int_{\partial V} (\lambda - \Sigma_\xi)^{-1} d\lambda,$$

it follows that

$$\sum_{q \in L(z)} Q_\xi(q) \xrightarrow[z \to \infty]{} \Pi_{\lambda_\xi(p_0)}(\Sigma_\xi). \tag{4.41}$$

Assume that $p_0 \in \widehat{X}_\infty^+$. This means that $\lambda_\xi(p_0) \in \sigma(\Sigma_\xi^+)$. Put

$$\psi(z) = (1 - \nu_\xi(z))\left(\sum_{q \in L(z)} Q_\xi(q)\right).$$

Since $Q_\xi(q)$ and $\nu_\xi(z_\xi(q))$ commute for $q \in \widehat{X}_0 \setminus z_\xi^{-1}(\sigma(N_\xi))$ (see Lemma 4.15), $\psi(z)$ is projection-valued and meromorphic in a punctured neighbourhood of ∞. Also, by (4.41), we have $\psi(\infty) = 0$, because $\nu_\xi(\infty) = P_{M_+}$ and $P_{M_-}\Pi_{\lambda_\xi(p_0)}(\Sigma_\xi) = 0$. Since $\psi(z)$ is continuous at $z = \infty$, it follows that $\psi(z)$ vanishes in a neighbourhood of ∞.

The proof of the Lemma for the case when $p_0 \in \widehat{X}_\infty^+$ concludes by observing that

$$\varphi_+(p) = \psi(z_\xi(p))Q_\xi(p)$$

for a general p near p_0. The case where $p_0 \in \widehat{X}_\infty^-$ is treated in a similar way. □

We claim that, if (S) holds, then $\widehat{X}_\mathbb{R} \cap \widehat{X}_\infty = \emptyset$. To see this, observe that Σ_ξ cannot have eigenvalues λ with $|\lambda| = 1$, because such an eigenvalue will be fixed by the map $\lambda \mapsto \overline{\lambda}^{-1}$, but this map interchanges the disjoint sets $\sigma(\Sigma_\xi^-)$ and $\sigma(\Sigma_\xi^+)$. If $p \in \widehat{X}_\mathbb{R}$, then $|\lambda_\xi(p)| = 1$, because $z_\xi(p) = \overline{w_\xi}(p)$. Hence, $p \notin \widehat{X}_\infty$, because Σ_ξ has no eigenvalues of modulus 1.

Put

$$\Gamma = z_\xi(\widehat{X}_\mathbb{R}).$$

Since $\widehat{X}_\mathbb{R} \cap \widehat{X}_\infty = \emptyset$, we see that Γ is a compact curve in $\mathbb{C}$. In particular, $\widehat{X}_\infty \cap z_\xi^{-1}(\Gamma) = \emptyset$. It is worthy to note that the mosaic ν_ξ is holomorphic on $\widehat{\mathbb{C}} \setminus \Gamma$ except for a finite number of points. Indeed, by Lemma 4.14, it is holomorphic outside the support of e_ξ. Using (4.13), it suffices to note that $X_{\mathrm{aff},\mathbb{R}}$ and $\widehat{X}_\mathbb{R}$ differ by a finite number of points.

For the proof of the next Lemma, we will need to use the following version of the Privalov-Plemelj jump formula. Let $\Gamma : [0,1] \to \mathbb{C}$ be a parametrized piecewise smooth curve. We also denote by Γ its image $\Gamma([0,1])$. This curve has a well defined tangent except at a finite number of points. Let $\psi(s)$ be the angle that the tangent line at $\Gamma(s)$ makes with the real axis.

Let F be a function defined on $\mathbb{C} \setminus \Gamma$ and fix a point $z_0 = \Gamma(s_0) \in \Gamma$. Put $\psi_0 = \psi(s_0)$. Let

$$z_{\pm,\varepsilon} = z_0 \pm \varepsilon i e^{i\psi_0}$$

be the two points $z_{-,\varepsilon}$ and $z_{+,\varepsilon}$ that are on the line normal to Γ at z_0 and are at a distance ε of z_0. If the limit

$$\lim_{\varepsilon \to 0} F(z_{+,\varepsilon}) - F(z_{-,\varepsilon})$$

exists, we call that number the jump of F at z_0 and we denote it by $\operatorname{Jump} F(z_0)$.

Now we consider μ a finite complex Borel measure on Γ and its Cauchy integral

$$F(z) = \frac{1}{2\pi i} \int_\Gamma \frac{d\mu(u)}{u - z}, \qquad z \in \mathbb{C} \setminus \Gamma.$$

The function $F(z)$ has nontangential boundary values from each side at almost every point of Γ. Indeed, F belongs to the Smirnov class $E^p(\mathbb{C} \setminus \Gamma)$ for every $p < 1$. The jump of F at $z_0 \in \Gamma$ is precisely the difference of these two boundary values.

Denote by $\frac{d\mu}{|dz|}$ the Radon-Nikodym derivative of the absolutely continuous part of μ with respect to the arc-length measure on Γ, and put

$$\frac{d\mu}{dz}(z_0) = e^{-i\psi_0} \frac{d\mu}{|dz|}(z_0).$$

Then the Privalov-Plemelj jump formula states that

$$\operatorname{Jump} F(z_0) = \frac{d\mu}{dz}(z_0).$$

for almost every $z_0 \in \Gamma$ (with respect to arc-length measure on Γ).

See [CMR06] for a monograph devoted to the properties of the Cauchy integral in the case when $\Gamma = \mathbb{T}$. Many of the results mentioned above can be obtained from this case by conformal mapping. Privalov's results about the boundary behaviour of analytic functions were originally published in [Pri50] in Russian. A German translation of this work can be found in [Pri56]. Unfortunately, the author is not aware of any English translation of these works. An introduction to the Smirnov class E^p can be found in [Dur70, Chapter 10].

Another fact that we will need is Privalov's uniqueness theorem. This states that if f is holomorphic on a connected open set Ω bounded by a rectifiable curve and has zero nontangential boundary values at a subset of positive measure of $\partial\Omega$, then f is identically zero. It is easy to see that this is also true for holomorphic functions on a Riemann surface. A proof of this theorem for the case when $\Omega = \mathbb{D}$ can be found in [Koo98, Section III.D].

Lemma 4.18. *Let U be an open connected set in $\widehat{X} \setminus \widehat{X}_\mathbb{R}$. If $U \cap \widehat{X}_\infty^+ \neq \emptyset$, then*

$$\nu_\xi(z_\xi(p)) Q_\xi(p) = Q_\xi(p), \qquad p \in U \setminus z_\xi^{-1}(\Gamma).$$

If $U \cap \widehat{X}_\infty^- \neq \emptyset$, then

$$\nu_\xi(z_\xi(p)) Q_\xi(p) = 0, \qquad p \in U \setminus z_\xi^{-1}(\Gamma).$$

Proof. We will give the proof for the case $U \cap \widehat{X}^+_\infty \neq \emptyset$. The other case is symmetric. Take $p_0 \in U \cap \widehat{X}^+_\infty$, and define $\varphi_+(p)$ as in the previous Lemma. We know that $\varphi_+ \equiv 0$ near p_0. We will use a continuation argument to show that $\varphi_+ \equiv 0$ on all of $U \setminus z_\xi^{-1}(\Gamma)$.

Fix an arbitrary point $p \in U \setminus z_\xi^{-1}(\Gamma)$. We can make a finite list $\Omega_0, \dots, \Omega_k$, where Ω_j are connected components of $U \setminus z_\xi^{-1}(\Gamma)$, the boundaries $\partial\Omega_j$ and $\partial\Omega_{j+1}$ have a common arc Γ_j contained in U, the point p_0 lies in Ω_0, and the point p lies in Ω_k.

Since φ_+ is identically zero on a neighbourhood of p_0, it is identically zero on all Ω_0. We are going to prove that this implies that φ_+ is also identically zero on Ω_1. Iterating this argument, we will finally show that φ_+ is identically zero on Ω_k. In particular, $\varphi_+(p) = 0$. If we can show that φ_+ has zero nontangential limit from Ω_1 at almost every $q \in \Gamma_0$, the common arc of $\partial\Omega_0$ and $\partial\Omega_1$ inside U, then we will have $\varphi_+ \equiv 0$ in Ω_1 by the Privalov's uniqueness theorem.

Since we are interested only on what happens a.e. on Γ_0, we can take a $q_0 \in \Gamma_0$ such that $q_0 \in \widehat{X}_0$ and $dz_\xi(q_0) \neq 0$ (because $\widehat{X} \setminus \widehat{X}_0$ is finite, and $dz_\xi(q) = 0$ only for a finite number of points $q \in \widehat{X}$). This second condition implies that $z = z_\xi(q)$ gives a local coordinate near q_0. We put $z_0 = z_\xi(q_0)$, and write everything using this coordinate z. We must study

$$\varphi_+(z) = [1 - \nu_\xi(z)]Q_\xi(z).$$

Since $\varphi_+(z)$ is identically zero for z on one side of Γ (the side corresponding to $z_\xi(\Omega_0)$), to see that the nontangential limit from the other side at z_0 is zero, it is enough to show that $\operatorname{Jump}\varphi_+(z_0) = 0$.

The function $Q_\xi(z)$ is continuous at z_0. Also, the nontangential boundary value of $\nu_\xi(z)$ exists a.e. on Γ, because by Proposition 4.14, $\nu_\xi(z) - P_{M_+}$ is the Cauchy integral of a finite Borel measure. We assume that $\nu_\xi(z)$ has nontangential boundary values at z_0. Hence,

$$\operatorname{Jump}\varphi_+(z_0) = \operatorname{Jump}F(z_0),$$

where

$$F(z) = -(\nu_\xi(z) - P_{M_+})Q_\xi(z_0).$$

The function $F(z)$ is the Cauchy integral of the measure

$$-2\pi i\, de_\xi\, \alpha_\xi Q_\xi(z_0).$$

By the Privalov-Plemelj jump formula,

$$\operatorname{Jump}F(z_0) = -2\pi i \frac{de_\xi}{dz}(z_0)\alpha_\xi Q_\xi(z_0).$$

We have $w_\xi(q_0) \neq \overline{z_\xi}(q_0)$, because $U \cap \widehat{X}_\mathbb{R} = \emptyset$. Therefore, we can define a function ψ analytic in a neighbourhood of $\sigma(z_\xi(q_0)\Sigma_\xi + D_\xi)$ such that $\psi(u) = (u - \overline{z_\xi}(q_0))^{-1}$ in a small neighbourhood of $w_\xi(q_0) \in \sigma(z_\xi(q_0)\Sigma_\xi + D_\xi)$ and $\psi(u) = 0$ outside of this neighbourhood. Put

$$\Psi = \psi(z_\xi(q_0)\Sigma_\xi + D_\xi).$$

Then we get by the Riesz-Dunford calculus (see Appendix A.1) that

$$Q_\xi(q_0) = (z_\xi(q_0)\Sigma_\xi + D_\xi - \overline{z_\xi}(q_0))\Psi.$$

This implies that

$$\operatorname{Jump}F(z_0) = -2\pi i \frac{de_\xi}{dz}(z_0)\alpha_\xi(z_0\Sigma_\xi + D_\xi - \overline{z_0})\Psi = 2\pi i \frac{de_\xi}{dz}(z_0)(z_0\alpha_\xi^* + \overline{z_0}\alpha_\xi + \gamma_\xi)\Psi.$$

By the relation obtained by taking adjoints in (4.10), we see that

$$\frac{de_\xi}{dz}(z_0)(z_0\alpha_\xi^* + \overline{z_0}\alpha_\xi + \gamma_\xi) = 0,$$

which implies that $\operatorname{Jump} F(z_0) = 0$. This finishes the proof, because it shows that φ_+ has zero jump at almost every point of Γ_0, and therefore zero boundary value from Ω_1 at almost every point of Γ_0. Hence, φ_+ must be identically zero on Ω_1 and we can iterate the argument. ☐

Using this Lemma, we can define $\widehat{X}_-$ and $\widehat{X}_+$, and prove the restoration formula for the mosaic ν_ξ.

Theorem 4.19. *Assume that $\xi \in \Xi$ is a nondegenerate direction such that α_ξ is invertible. Suppose that (S) holds. Then there exists a partition of $\widehat{X} \setminus \widehat{X}_\mathbb{R}$ into two* halves *$\widehat{X}_-$ and $\widehat{X}_+$ with the following properties:*

(a) Each half is the union of some of the connected components of $\widehat{X} \setminus \widehat{X}_\mathbb{R}$.

(b) The two halves are conjugate: $*(\widehat{X}_-) = \widehat{X}_+$, and $*(\widehat{X}_+) = \widehat{X}_-$.

(c) If a component $\widehat{X}_j$ intersects $\widehat{X}_\mathbb{R}$, then it intersects both halves $\widehat{X}_-$ and $\widehat{X}_+$.

(d) If a component $\widehat{X}_j$ does not intersect $\widehat{X}_\mathbb{R}$, then it is contained either in $\widehat{X}_-$ or in $\widehat{X}_+$. Moreover, $*(\widehat{X}_j)$ is a different component of $\widehat{X}$.

(e) The restoration formula holds:

$$\nu_\xi(z) = \sum_{\substack{p\in\widehat{X}_+ \\ z_\xi(p)=z}} Q_\xi(p), \qquad z \in \widehat{\mathbb{C}} \setminus \Gamma, \tag{4.42}$$

where $\Gamma = z_\xi(\widehat{X}_\mathbb{R})$.

Proof. We define $\widehat{X}_-$ as the union of the connected components of $\widehat{X} \setminus \widehat{X}_\mathbb{R}$ which intersect $\widehat{X}_\infty^-$, and similarly for $\widehat{X}_+$. By Lemma 4.18, $\widehat{X}_-$ and $\widehat{X}_+$ are disjoint (note that $Q_\xi(p)$ cannot vanish).

Now we observe that every connected component of $\widehat{X} \setminus \widehat{X}_\mathbb{R}$ must intersect $\widehat{X}_\infty$, and hence $\widehat{X} = \widehat{X}_- \cup \widehat{X}_\mathbb{R} \cup \widehat{X}_+$. Indeed, assume that U is a connected component of $\widehat{X} \setminus \widehat{X}_\mathbb{R}$ which doesn't intersect $\widehat{X}_\infty$. Let $\widehat{X}_j$ be the connected component of $\widehat{X}$ which contains U.

The theory of real algebraic curves shows that there are two possibilities for the (complex) Riemann surface of an irreducible real algebraic curve: either the set of points not fixed by the involution induced by complex conjugation is connected, or either it consists of precisely two connected components, which are interchanged by the involution. In the second case, we say that the surface is separated. See [GH81, Section 3] for an exposition of the topological properties of real algebraic curves.

In our case, there are two possible cases: either $\widehat{X}_j$ contains no real point, or $\widehat{X}_j$ contains real points, and hence, it is fixed by the involution $*$ (because $*$ permutes the components of $\widehat{X}$ and some points of $\widehat{X}_j$ are fixed by $*$). In the first case, we have $U = \widehat{X}_j$. This is a contradiction, because $\widehat{X}_j$ must contain points of $\widehat{X}_\infty$.

In the second case, $\widehat{X}_j$ is the Riemann surface of an irreducible real algebraic curve. Since $\widehat{X}_j$ contains points of $\widehat{X}_\infty$, and U and $\widehat{X}_\mathbb{R}$ do not intersect $\widehat{X}_\infty$, the surface $\widehat{X}_j$ must be separated, and U must be one of the two connected components of $\widehat{X}_j \setminus \widehat{X}_\mathbb{R}$. Therefore, $\widehat{X}_j \setminus \widehat{X}_\mathbb{R} = U \cup *(U)$.

Since $*$ maps $\widehat{X}_\infty$ onto $\widehat{X}_\infty$, we see that $*(U)$ does not intersect $\widehat{X}_\infty$. Hence, $\widehat{X}_j$ does not intersect $\widehat{X}_\infty$. This is a contradiction again.

Therefore, we see that $\widehat{X}_-$ and $\widehat{X}_+$ are a partition of $\widehat{X}\backslash\widehat{X}_\mathbb{R}$. Property (a) holds by construction, and it is also clear that (b) is true. Properties (c) and (d) are obtained using the fact that the involution $*$ interchanges $\widehat{X}_\infty^-$ and $\widehat{X}_\infty^+$.

To obtain the restoration formula, we use (4.16) to get the equation

$$\sum_{\substack{p\in\widehat{X}\\ z_\xi(p)=z}} Q_\xi(p) = I_M.$$

Then we multiply this equation on the left by $\nu_\xi(z)$ and use Lemma 4.18. □

Example 4.20. In the case of the separating structure generated by a subnormal operator S, the operator α_ξ has the form

$$\alpha_\xi = \frac{1}{2}\begin{bmatrix} 0 & -(\xi_1+i\xi_2)T_0^* \\ (\xi_1-i\xi_2)T_0 & 0 \end{bmatrix},$$

where $T_0 = X|M_-$ (see Example 4.3). Hence, the operator Σ_ξ is

$$\Sigma_\xi = -\alpha_\xi^{-1}\alpha_\xi^* = \begin{bmatrix} \dfrac{\overline{\xi_1}-i\overline{\xi_2}}{\xi_1-i\xi_2}I_{M_-} & 0 \\ 0 & \dfrac{\overline{\xi_1}+i\overline{\xi_2}}{\xi_1+i\xi_2}I_{M_+}. \end{bmatrix}$$

Therefore, the sets $\sigma(\Sigma_\xi^-)$ and $\sigma(\Sigma_\xi^+)$ consist of only one point each. Moreover, we see that these points are different if and only if the direction ξ is nondegenerate, i.e., if $\operatorname{Im}\xi_1\overline{\xi_2}\neq 0$. This means that we can carry out the construction given above to define the halves $\widehat{X}_-$ and $\widehat{X}_+$, as long as we choose a general direction ξ. ♠

A necessary remark is that Theorem 4.19 does not prove that the discriminant curve $\widehat{X}$ is separated in the sense that we gave in the Preface, which was that each component $\widehat{X}$ is divided in two connected components when we remove its real points. Here, it may happen that some components have no real points and so, belong either to $\widehat{X}_-$ or to $\widehat{X}_+$, and there is a conjugate component in the other half of the curve $\widehat{X}$. These components are in some sense degenerate. However, this partition into halves $\widehat{X}_-$ and $\widehat{X}_+$ should be good enough to allow the development of the theory. For instance, in [Yak98a], it happened that the only degenerate components that appeared were those of degree one.

Another remark is that the restoration formula (4.42) imposes a strong condition on the spectrum of N_ξ. Using the restoration formula, we see that $\nu_\xi(z)$ is discontinuous at Γ except for a finite number of points. Since we know that $\nu_\xi(z)$ is holomorphic outside $\operatorname{supp} e_\xi$, we get that $\Gamma\subset\operatorname{supp} e_\xi\subset\sigma(N_\xi)$. We also know that e_ξ is supported in Γ and perhaps a finite number of additional points. This implies that if the pool associated with the separating structure is irreducible, then $\sigma(N_\xi)$ is precisely the curve Γ and perhaps a finite number of isolated points.

4.5. Generalized compression

This section deals with the notion of the generalized compression and its application to separating structures to obtain vessels. First we will give the abstract definition of the generalized compression of a linear operator, and then we will pass to study it in the context of separating structures.

Definition. Let $K \supset H \supset G$ be vector spaces and $A : K \to K$ a linear map which satisfies the two following conditions:

$$AG \cap H \subset G, \tag{C1}$$

and

$$AH \subset AG + H. \tag{C2}$$

Then we define the compression $\widetilde{A} : H/G \to H/G$ by the following procedure. Given a vector $h \in H$, using (C2), we can find a $g \in G$ such that $h' = A(h - g) \in H$. Then we define

$$\widetilde{A}(h + G) = h' + G.$$

To check that this is well defined, we must see that if $h \in G$, then $h' \in G$, but this is a consequence of (C1). Hence, the compression $\widetilde{A}$ is a linear map on the quotient space H/G.

In the context of Hilbert spaces, i.e., when K, H, G are Hilbert spaces and $A \in \mathcal{B}(K)$, one should replace (C1) by

$$\overline{AG} \cap H \subset G, \tag{C1*}$$

and also require that if $L = \overline{AG} \cap (K \ominus G)$, then

$$L \dot{+} H \text{ is a direct sum} \tag{C3}$$

(note that $L \cap H = 0$ by (C1*)). Let

$$P : L \dot{+} H \to H \tag{4.43}$$

be the parallel projection onto H according to this direct sum decomposition. Now we see that (C2) implies that $AH \subset L + H$, and that the compression $\widetilde{A}$ is

$$\widetilde{A}(h + G) = PAh + G, \qquad h \in H.$$

Hence, $\widetilde{A}$ is bounded and $\|\widetilde{A}\| \le \|P\|\|A\|$. If we identify the quotient space H/G with the space $R = H \ominus G$, then we see that

$$\widetilde{A} = P_R PA|R. \tag{4.44}$$

To see that this generalizes the classical notion of a compression, assume that $K = H_1 \oplus H_2 \oplus H_3$ and that A has the structure

$$A = \begin{bmatrix} * & 0 & 0 \\ * & A_0 & 0 \\ * & * & * \end{bmatrix}$$

according to this decomposition, so that A is a dilation of A_0 and A_0 is its classical compression (see Section 1.1). Then we put $G = H_3$, $H = H_2 \oplus H_3$. We have $AG \subset G$, so that (C1*) holds. Also, $AH \subset H$, and (C2) holds. Moreover, $L = 0$, which implies that (C3) holds and $P = I_H$ (see (4.43)). We identify the quotient H/G with the space $R = H \ominus G = H_2$. Now, (4.44) shows that $\widetilde{A} = P_{H_2} A|H_2 = A_0$. Hence, in this setting, the generalized compression coincides with the classical compression.

Since we are interested in compressing the operators A_1, A_2 in a separating structure to obtain operators $\widetilde{A}_1$, $\widetilde{A}_2$ forming a (commutative) pool, we should know when the compressions of two commuting operators also commute.

Lemma 4.21. *Assume that A_1 and A_2 are commuting linear maps on K and let $\widetilde{A}_1$ and $\widetilde{A}_2$ be their respective compressions. If*

$$A_1 A_2 G \cap H \subset G, \tag{4.45}$$

then the compressions $\widetilde{A}_1$ and $\widetilde{A}_2$ commute.

Proof. Take an $x \in H$. Then there are vectors $g, g', l, l' \in G$ such that

$$\begin{aligned} y &= A_1(x-g) \in H, & y' &= A_2(x-h') \in H, \\ z &= A_2(y-l) \in H, & z' &= A_1(y'-l') \in H. \end{aligned}$$

By definition of the compression, $\widetilde{A}_2\widetilde{A}_1(x+G) = z+G$ and $\widetilde{A}_1\widetilde{A}_2(x+G) = z'+G$. We must check that $z - z' \in G$. We compute

$$z - z' = A_2(A_1(x-g)-l) - A_2(A_1(x-g')-l') = A_1A_2(g'-g+l'-l).$$

The vector on the right hand side of this equation is in A_1A_2G and the vector on the left hand side is in H. It suffices to use (4.45). □

Now we pass to the compression of separating structures. Recall that a tuple

$$\omega = (K, A_1, A_2, H_{0,-}, M_-, M_+, H_{0,+})$$

is called an orthogonal separating structure if A_1 and A_2 are selfadjoint operators on K,

$$K = H_{0,-} \oplus M_- \oplus M_+ \oplus H_{0,+},$$

and

$$\begin{aligned} A_j H_{0,-} &\subset H_-, & A_j H_- &\subset H_- + M_+, \\ A_j H_+ &\subset M_- + H_+, & A_j H_{0,+} &\subset H_+, \end{aligned}$$

for $j = 1, 2$, where

$$H_- = H_{0,-} + M_-, \qquad H_+ = M_+ + H_{0,+}.$$

Suppose that A_1 and A_2 are two selfadjoint operators on K which are included in two orthogonal separating structures ω and $\widehat{\omega}$, so that:

$$\omega = (K, A_1, A_2, H_{0,-}, M_-, M_+, H_{0,+}), \qquad \widehat{\omega} = (K, A_1, A_2, \widehat{H}_{0,-}, \widehat{M}_-, \widehat{M}_+, \widehat{H}_{0,+}).$$

We write $\widehat{\omega} \prec \omega$ if

$$H_- \subset \widehat{H}_-, \qquad H_+ \supset \widehat{H}_+. \tag{4.46}$$

Observe that conditions (4.46) and those involved in the definition of the separating structures ω and $\widehat{\omega}$ remain invariant if we exchange the subscripts $_+$ and $_-$, remove the hat $\widehat{\ }$ from those spaces which had it, and add it to those spaces which did not have it. This kind of symmetry will be called *hat-symmetry* and will be useful later.

We will now construct the compression of two structures such that $\widehat{\omega} \prec \omega$. We start by defining the operators $\beta_j : M_+ \to M_-$ and $\widehat{\beta}_j : \widehat{M}_+ \to \widehat{M}_-$ by

$$\beta_j = P_{M_-} A_j | M_+, \qquad \widehat{\beta}_j = P_{\widehat{M}_-} A_j | \widehat{M}_+, \qquad j = 1, 2. \tag{4.47}$$

Note that their adjoints are

$$\beta_j^* = P_{M_+} A_j | M_-, \qquad \widehat{\beta}_j^* = P_{\widehat{M}_+} A_j | \widehat{M}_-,$$

Using the formula given in Theorem 4.9 for the rates σ_j of the pool Б generated by ω and the rates $\widehat{\sigma}_j$ of the pool $\widehat{Б}$ generated by $\widehat{\omega}$, we see that

$$\sigma_j = \begin{bmatrix} 0 & i\beta_j \\ -i\beta_j^* & 0 \end{bmatrix}, \qquad \widehat{\sigma}_j = \begin{bmatrix} 0 & i\widehat{\beta}_j \\ -i\widehat{\beta}_j^* & 0 \end{bmatrix}, \tag{4.48}$$

according to the decompositions $M = M_- \oplus M_+$ and $\widehat{M} = \widehat{M}_- \oplus \widehat{M}_+$.

Hence, if we assume the non-degeneracy condition (4.14) for both σ_1, σ_2 and $\widehat{\sigma}_1, \widehat{\sigma}_2$, replacing A_1 and A_2 by $t_1 A_1 + t_2 A_2$ and $t_3 A_1 + t_4 A_2$, where $t_k \in \mathbb{R}$ and $t_1 t_4 - t_2 t_3 \neq 0$, we may assume that

$$\beta_j \text{ and } \widehat{\beta}_j \text{ are invertible, for } j = 1, 2. \tag{Iβ}$$

We will do so hereafter.

We define the space

$$R = \widehat{H}_- \ominus H_- = H_+ \ominus \widehat{H}_+ = H_+ \cap \widehat{H}_-. \tag{4.49}$$

This space will be the compression space when we compress A_1 and A_2 to either $\widehat{H}_-/H_-$ or $H_+/\widehat{H}_+$, because it can be identified with both quotients.

We will need the following technical lemma.

Lemma 4.22. *The following relation holds:*

$$(P_{M_-} - I)\widehat{M}_- \subset R.$$

Proof. First, $(P_{M_-} - I)\widehat{M}_- \subset \widehat{H}_-$, because $M_- \subset H_- \subset \widehat{H}_-$. Second,

$$P_{H_-}(P_{M_-} - I)|\widehat{M}_- = (P_{M_-} - P_{H_-})|\widehat{M}_- = -P_{H_{0,-}}|\widehat{M}_-.$$

Since $\widehat{\beta}_1$ is invertible,

$$\widehat{M}_- = \widehat{\beta}_1 \widehat{M}_+ = P_{\widehat{M}_-} A_1 \widehat{M}_+.$$

Using $P_{H_{0,-}} P_{\widehat{H}_+} = 0$, we get

$$\begin{aligned} P_{H_{0,-}} \widehat{M}_- &= P_{H_{0,-}} P_{\widehat{M}_-} A_1 \widehat{M}_+ = P_{H_{0,-}} (P_{\widehat{M}_-} + P_{\widehat{H}_+}) A_1 \widehat{M}_+ \\ &= P_{H_{0,-}} A_1 \widehat{M}_+ \subset P_{H_{0,-}} A_1 H_+ = 0. \end{aligned}$$

This implies the relation

$$P_{H_{0,-}} P_{\widehat{M}_-} = 0, \tag{4.50}$$

which will be useful later.

We see that $P_{H_-}(P_{M_-} - I)|\widehat{M}_- = 0$. This finishes the proof, because $R = \widehat{H}_- \ominus H_-$. □

Now we define the operators $\tau_- : \widehat{M}_- \to M_-$ and $\tau_+ : \widehat{M}_+ \to M_+$ by

$$\tau_- = P_{M_-}|\widehat{M}_-, \qquad \tau_+ = P_{M_+}|\widehat{M}_+. \tag{4.51}$$

Note that the adjoints of these operators are

$$\tau_-^* = P_{\widehat{M}_-}|M_-, \qquad \tau_+^* = P_{\widehat{M}_+}|M_+.$$

The following Lemma relates these two operators with the possibility of compressing the operators A_1 and A_2 using the generalized compression as defined above.

Lemma 4.23. *Put $H = H_+$ and $G = \widehat{H}_+$. The operators A_1 and A_2 satisfy the conditions (C1*), (C2), and (C3) needed for the construction of their compressions to $H_+/\widehat{H}_+$ if and only if τ_- is invertible.*

Similarly, the operators can be compressed to $\widehat{H}_-/H_-$ if and only if τ_+ is invertible.

Proof. First assume that τ_- is invertible. We have

$$\overline{A_j\widehat{H}_+} \cap H_+ \subset (\widehat{M}_- + \widehat{H}_+) \cap H_+ = (\widehat{M}_- \cap H_+) + \widehat{H}_+,$$

because $\widehat{H}_+ \subset H_+$. Now we check that

$$\widehat{M}_- \cap H_+ = 0, \tag{4.52}$$

so that we get condition (C1*). Assume that $x \in \widehat{M}_- \cap H_+$. Then $\tau_- x = P_{M_-}x = 0$, so that $x = 0$, because τ_- is invertible.

To prove condition (C2), we first check that

$$M_- \subset H_+ + \widehat{M}_-. \tag{4.53}$$

Take $m_- \in M_-$. Since τ_- is invertible, $m_- = \tau_- \widehat{m}_-$ for some $\widehat{m}_- \in \widehat{M}_-$. Then, by Lemma 4.22,

$$m_- - \widehat{m}_- = (P_{M_-} - I)\widehat{m}_- \in R \subset H_+.$$

Hence, $m_- \in H_+ + \widehat{M}_-$.

Now we see that

$$A_jH_+ \subset H_+ + M_- \subset H_+ + \widehat{M}_- = A_j\widehat{H}_+ + H_+.$$

Here the last inequality comes from the fact that

$$A_j\widehat{H}_+ + \widehat{H}_+ = \widehat{M}_- + \widehat{H}_+, \tag{4.54}$$

which is true because $\widehat{\beta}_j$ is onto.

Condition (C3) holds because

$$L = \overline{A_j\widehat{H}_+} \cap \widehat{H}_- \subset \widehat{M}_-,$$

so that L is finite-dimensional and therefore the sum $L \dotplus H_+$ is always direct. Hence, A_j can be compressed to $H_+/\widehat{H}_+$.

Let us now assume that A_j can be compressed to $H_+/\widehat{H}_+$ and prove that τ_- is invertible. By (C1*) for A_j instead of A,

$$A_j\widehat{H}_+ \cap H_+ \subset \widehat{H}_+.$$

Since (4.54) holds, we have $\widehat{M}_- \cap H_+ \subset \widehat{H}_+$. This implies (4.52), and from this it follows that τ_- is injective. We also have

$$M_- \subset A_jH_+ + H_+ \subset A_j\widehat{H}_+ + H_+ = \widehat{M}_- + H_+.$$

Here the first inclusion comes from the relation obtained by removing all the hats $\widehat{\ }$ in (4.54) (the relation obtained is true because β_j is onto), the second inclusion comes from (C2) for A_j instead of A, and the last equality uses again (4.54). Hence, we have (4.53), and from this it follows that τ_- is onto. This proves the first statement of the Lemma.

To prove the second statement, we apply hat-symmetry to see that A_1, A_2 can be compressed to $\widehat{H}_-/H_-$ if and only if τ_+^* (which is the hat-symmetric of τ_-) is invertible. $\square$

Lemma 4.24. *The following relations hold:*

$$\beta_j \tau_+ = \tau_- \widehat{\beta}_j, \qquad j = 1, 2.$$

Proof. Since $\widehat{M}_+ \subset \widehat{H}_+ \subset H_+$,

$$A_j|\widehat{M}_+ = A_j P_{H_{0,+}}|\widehat{M}_+ + A_j P_{M_+}|\widehat{M}_+.$$

Hence,

$$P_{M_-} A_j|\widehat{M}_+ = P_{M_-} A_j P_{H_{0,+}}|\widehat{M}_+ + P_{M_-} A_j P_{M_+}|\widehat{M}_+ = \beta_j \tau_+,$$

because $P_{M_-} A_j P_{H_{0,+}} = 0$.

Also, since $A_j \widehat{M}_+ \subset \widehat{H}_+ + \widehat{M}_-$,

$$P_{M_-} A_j|\widehat{M}_+ = P_{M_-} P_{\widehat{H}_+} A_j|\widehat{M}_+ + P_{M_-} P_{\widehat{M}_-} A_j|\widehat{M}_+ = \tau_- \widehat{\beta}_j,$$

because $\widehat{H}_+ \subset H_+$ implies $P_{M_-} P_{\widehat{H}_+} = 0$. We have obtained the desired equality. □

Since we assume that β_j and $\widehat{\beta}_j$ are invertible (see (Iβ) on page 75), Lemma 4.24 implies that τ_+ is invertible if and only if τ_- is invertible. By Lemma 4.23, we see that A_1 and A_2 can be compressed to $H_+/\widehat{H}_+$ if and only if they can be compressed to $\widehat{H}_-/H_-$, and that this happens whenever both τ_- and τ_+ are invertible. From now on, we will assume (Iβ) and

$$\tau_- \text{ and } \tau_+ \text{ are invertible.} \tag{Iτ}$$

Now we are ready to give a formula for the compression of A_1 and A_2 to $H_+/\widehat{H}_+$.

Lemma 4.25. *If we identify R with the quotient space $H_+/\widehat{H}_+$, then the compression of A_j to $H_+/\widehat{H}_+$ has the form*

$$\widetilde{A}_j = P_R(A_j - \tau_-^{-1}\beta_j P_{M_+})|R.$$

Proof. We will write the proof for $\widetilde{A}_1$. The same argument applies to $\widetilde{A}_2$. Let $h \in R$. Then, by definition of the compression, there is a $g \in \widehat{H}_+$ such that $A_1(h-g) \in H_+$, and $\widetilde{A}_1 h = P_R A_1(h-g)$. Now, since $A_1 g \in \widehat{H}_+ + \widehat{M}_-$, and $R = H_+ \cap \widehat{H}_-$, we have $P_R A_1 g = P_R P_{\widehat{M}_-} A_1 g$. Since $g \in \widehat{H}_+$, we have $P_{\widehat{M}_-} A_1 g = P_{\widehat{M}_-} A_1 P_{\widehat{M}_+} g = \widehat{\beta}_1 P_{\widehat{M}_+} g$. Hence,

$$P_R A_1 g = P_R \widehat{\beta}_1 P_{\widehat{M}_+} g = P_R \tau_-^{-1} \tau_- \widehat{\beta}_1 P_{\widehat{M}_+} g = P_R \tau_-^{-1} \beta_1 \tau_+ P_{\widehat{M}_+} g.$$

By hat-symmetry in (4.50), we get $P_{\widehat{H}_{0,+}} P_{M_+} = 0$, which taking adjoints becomes $P_{M_+} P_{\widehat{H}_{0,+}} = 0$. This shows that

$$\tau_+ P_{\widehat{M}_+}|\widehat{H}_+ = P_{M_+}(P_{\widehat{M}_+} + P_{\widehat{H}_{0,+}})|\widehat{H}_+ = P_{M_+}|\widehat{H}_+.$$

Since $g \in \widehat{H}_+$, we see that

$$P_R A_1 g = P_R \tau_-^{-1} \beta_1 P_{M_+} g = P_R \tau_-^{-1} P_{M_-} A_1 P_{M_+} g = P_R \tau_-^{-1} P_{M_-} A_1 g.$$

Here the last equality holds because $g \in H_+$ and $P_{M_-} A_1 P_{H_{0,+}} = 0$. The condition $A_1(h-g) \in H_+$ implies $P_{M_-} A_1 g = P_{M_-} A_1 h$. Hence,

$$P_R A_1 g = P_R \tau_-^{-1} P_{M_-} A_1 h = P_R \tau_-^{-1} \beta_1 P_{M_+} h,$$

where the last equality is true because $h \in H_+$. This proves the Lemma, because $\widetilde{A}_1 h = P_R A_1(h-g)$. □

We have two different options to construct the compression of A_j to R. We can either do the compression to the quotient space $H_+/\widehat{H}_+$ or to the quotient space $\widehat{H}_-/H_-$. Both of these spaces are identified with R, but the compression produces different operators. We decide to denote by $\widetilde{A}_j$ the compression of A_j to $H_+/\widehat{H}_+$. The surprising fact is that the compression of A_j to $\widehat{H}_-/H_-$ is just the adjoint $\widetilde{A}_j^*$.

Proposition 4.26. *For $j = 1, 2$, let $\widetilde{A}_j$ denote the compression of A_j to the quotient space $H_+/\widehat{H}_+$, which we identify with R. Then the compression of A_j to the quotient space $\widehat{H}_-/H_-$, also identified with R, is $\widetilde{A}_j^*$.*

We also have the following formula for $\widetilde{A}_j^*$:

$$\widetilde{A}_j^* = P_R(A_j - \beta_j^* \tau_-^{-*} P_{\widehat{M}_-})|R$$

(Here $\tau_+^{-*} = (\tau_+^{-1})^*$).

Proof. By applying hat-symmetry to Lemma 4.25, we see that the formula for the compression of A_j to $\widehat{H}_-/H_-$ is

$$P_R(A_j - \tau_+^{-*} \widehat{\beta}_j^* P_{\widehat{M}_-})|R.$$

Note that R is hat-symmetric to itself, τ_- is hat-symmetric to τ_+^*, and β_j is hat-symmetric to $\widehat{\beta}_j^*$.

Now we compute the adjoint of the second part in the formula for $\widetilde{A}_j$ given in Lemma 4.25, using Lemma 4.24:

$$[P_R \tau_-^{-1} \beta_j P_{M_+}|R]^* = P_R \beta_j^* \tau_-^{-*} P_{\widehat{M}_-}|R = P_R \tau_+^{-*} \widehat{\beta}_j^* P_{\widehat{M}_-}|R.$$

The first part now follows, because $(P_R A_j|R)^* = P_R A_j|R$. The formula for $\widetilde{A}_j^*$ has been obtained throughout the proof. $\square$

We will now compute the rates and the gyrations of a vessel in which the compressions $\widetilde{A}_1^*$ and $\widetilde{A}_2^*$ can be included. The next Lemma motivates the definition of the window operator $\widetilde{\Phi}$ and shows that the rates of the vessel will coincide with the rates σ_j.

Lemma 4.27. *Define the operator $\widetilde{\Phi} : R \to M = M_- \oplus M_+$ by*

$$\widetilde{\Phi} = \begin{bmatrix} -\tau_-^{-*} P_{\widehat{M}_-}|R \\ P_{M_+}|R \end{bmatrix}. \tag{4.55}$$

Then

$$\frac{1}{i}(\widetilde{A}_j^* - \widetilde{A}_j) = \widetilde{\Phi}^* \sigma_j \widetilde{\Phi}, \qquad j = 1, 2,$$

where σ_j are given by (4.48).

Proof. Using the formulas for $\widetilde{A}_j$ and $\widetilde{A}_j^*$ given in Lemma 4.25 and Proposition 4.26, we see that

$$\widetilde{A}_j^* - \widetilde{A}_j = P_R(-\beta_j^* \tau_-^{-*} P_{\widehat{M}_-} + \tau_-^{-1} \beta_j P_{M_+})|R. \tag{4.56}$$

Now we compute

$$\begin{aligned} \widetilde{\Phi}^* \sigma_j \widetilde{\Phi} &= \begin{bmatrix} -P_R \tau_-^{-1} & P_R \end{bmatrix} \begin{bmatrix} 0 & i\beta_j \\ -i\beta_j^* & 0 \end{bmatrix} \begin{bmatrix} -\tau_-^{-*} P_{\widehat{M}_-}|R \\ P_{M_+}|R. \end{bmatrix} \\ &= P_R(i\beta_j^* \tau_-^{-*} P_{\widehat{M}_-} - i\tau_-^{-1} \beta_j P_{M_+})|R = \frac{1}{i}(\widetilde{A}_j^* - \widetilde{A}_j). \end{aligned}$$

$\square$

The next two lemmas are calculations needed to compute the gyrations of the vessel.

Lemma 4.28. *The following equality holds for $j = 1, 2$:*

$$P_{\widehat{M}_-}(A_j \tau_-^* - \beta_j^*)|M_- = (\tau_-^* P_{M_-} A_j - P_{\widehat{M}_-} A_j P_{\widehat{H}_{0,-}})|M_-.$$

Proof. Since $M_- \subset \widehat{H}_-$, we have

$$P_{\widehat{M}_-} A_j |M_- = P_{\widehat{M}_-} A_j (P_{\widehat{M}_-} + P_{\widehat{H}_{0,-}})|M_- = (P_{\widehat{M}_-} A_j \tau_-^* + P_{\widehat{M}_-} A_j P_{\widehat{H}_{0,-}})|M_-.$$

Since $A_j M_- \subset H_- + M_+$, using the relation obtained by taking adjoints in (4.50), we get

$$P_{\widehat{M}_-} A_j |M_- = P_{\widehat{M}_-}(P_{M_-} + P_{M_+}) A_j |M_- = (\tau_-^* P_{M_-} A_j + P_{\widehat{M}_-} \beta_j^*)|M_-.$$

Hence, we see that

$$(P_{\widehat{M}_-} A_j \tau_-^* + P_{\widehat{M}_-} A_j P_{\widehat{H}_{0,-}})|M_- = (\tau_-^* P_{M_-} A_j + P_{\widehat{M}_-} \beta_j^*)|M_-.$$

The Lemma now follows by rearranging terms. $\square$

Lemma 4.29. *The following two relations hold for $j = 1, 2$:*

$$P_{M_+} \widetilde{A}_j^* = P_{M_+} A_j \widetilde{\Phi} + P_{M_+} A_j P_{H_{0,+}} |R. \tag{4.57}$$

$$-\tau_-^{-*} P_{\widehat{M}_-} \widetilde{A}_j^* = P_{M_-} A_j \widetilde{\Phi} - \tau_-^{-*} P_{\widehat{M}_-} A_j P_{\widehat{H}_{0,-}} (I - \tau_-^{-*} P_{\widehat{M}_-})|R. \tag{4.58}$$

Proof. First we prove (4.57). Take a fixed $r \in R$ and put $m_- = \tau_-^{-*} P_{\widehat{M}_-} r \in M_-$. We note that $r - m_- \in \widehat{H}_{0,-}$, because $r - m_- \in \widehat{H}_-$ (recall that $R \subset \widehat{H}_-$ and $M_- \subset H_- \subset \widehat{H}_-$), and

$$P_{\widehat{M}_-}(r - m_-) = P_{\widehat{M}_-} r - \tau_-^* m_- = 0.$$

It follows that $A(r - m_-) \in \widehat{H}_-$. By the definition of the compression, we see that

$$\widetilde{A}_j^* r + H_- = A(r - m_-) + H_-,$$

because $\widetilde{A}_j^*$ is the compression of A_j to the quotient $\widehat{H}_-/H_-$.

Since $P_{M_+}|H_- = 0$, the operator P_{M_+} is well defined in the quotient $\widehat{H}_-/H_-$. This implies that

$$P_{M_+} \widetilde{A}_j^* r = P_{M_+} A(r - m_-).$$

Therefore, we see that

$$P_{M_+} \widetilde{A}_j^* = P_{M_+} A_j |R - P_{M_+} A_j \tau_-^{-*} P_{\widehat{M}_-} |R,$$

because $r \in R$ was arbitrary. Writing

$$P_{M_+} A_j |R = P_{M_+} A_j P_{M_+} |R + P_{M_+} A_j P_{H_{0,+}} |R,$$

which is true because $R \subset H_+$, and using the definition of $\widetilde{\Phi}$ given in (4.55), we get (4.57).

To prove (4.58), we first apply hat-symmetry in (4.57) to obtain

$$P_{\widehat{M}_-} \widetilde{A}_j = (P_{\widehat{M}_-} A_j P_{\widehat{M}_-} - P_{\widehat{M}_-} A_j \tau_+^{-1} P_{M_+} + P_{\widehat{M}_-} A_j P_{\widehat{H}_{0,-}})|R.$$

(Note that hat-symmetry interchanges the operators $\widetilde{A}_j$ and $\widetilde{A}_j^*$). Since

$$P_{\widehat{M}_-} A_j \tau_+^{-1} = \widehat{\beta}_j \tau_+^{-1} = \tau_-^{-1} \beta_j$$

by Lemma 4.24, we get

$$P_{\widehat{M}_-} \widetilde{A}_j = (P_{\widehat{M}_-} A_j P_{\widehat{M}_-} - \tau_-^{-1}\beta_j P_{M_+} + P_{\widehat{M}_-} A_j P_{\widehat{H}_{0,-}})|R.. \tag{4.59}$$

Now we will use (4.56) to compute $P_{\widehat{M}_-}(\widetilde{A}_j^* - \widetilde{A})$. We have

$$P_{\widehat{M}_-} P_R | M_+ = P_{\widehat{M}_-} P_{\widehat{H}_-} | M_+ = P_{\widehat{M}_-} | M_+,$$

because $R = \widehat{H}_- \cap H_+$ and $M_+ \subset H_+$. Similarly,

$$P_{\widehat{M}_-} P_R | \widehat{M}_- = P_{\widehat{M}_-} P_{H_+} | \widehat{M}_- = P_{\widehat{M}_-}(I - P_{M_-}) | \widehat{M}_- = (I - \tau_-^* \tau_-) | \widehat{M}_-.$$

Here, the second equality is true by the relation obtained by taking adjoints in (4.50). Using these last two identities in (4.56), we see that

$$\begin{aligned} P_{\widehat{M}_-}(\widetilde{A}_j^* - \widetilde{A}_j) &= (-P_{\widehat{M}_-} \beta_j^* \tau_-^{-*} P_{\widehat{M}_-} + (I - \tau_-^*\tau_-)\tau_-^{-1}\beta_j P_{M_+})|R \\ &= (-P_{\widehat{M}_-} \beta_j^* \tau_-^{-*} P_{\widehat{M}_-} + \tau_-^{-1}\beta_j P_{M_+} - \tau_-^*\beta_j P_{M_+})|R. \end{aligned} \tag{4.60}$$

By (4.59) and (4.60),

$$\begin{aligned} P_{\widehat{M}_-} \widetilde{A}_j^* &= P_{\widehat{M}_-} \widetilde{A}_j + P_{\widehat{M}_-}(\widetilde{A}_j^* - \widetilde{A}_j) \\ &= (P_{\widehat{M}_-} A_j P_{\widehat{M}_-} + P_{\widehat{M}_-} A_j P_{\widehat{H}_{0,-}} - P_{\widehat{M}_-} \beta_j^* \tau_-^{-*} P_{\widehat{M}_-} - \tau_-^* \beta_j P_{M_+})|R \\ &= (P_{\widehat{M}_-}(A_j \tau_-^* - \beta^*) \tau_-^{-*} P_{\widehat{M}_-} - \tau_-^* \beta_j P_{M_+} + P_{\widehat{M}_-} A_j P_{\widehat{H}_{0,-}})|R. \end{aligned}$$

In this last equality, we have just rearranged terms. Using Lemma 4.28, we see that the last expression equals

$$(\tau_-^* P_{M_-} A_j \tau_-^{-*} P_{\widehat{M}_-} - P_{\widehat{M}_-} A P_{\widehat{H}_{0,-}} \tau_-^{-*} P_{\widehat{M}_-} - \tau_-^* \beta_j P_{M_+} + P_{\widehat{M}_-} A_j P_{\widehat{H}_{0,-}})|R.$$

Multiplying by $-\tau_-^{-*}$ on the left and using (4.55), we get (4.58). □

The next Lemma shows that (3.24) is satisfied for $\widetilde{A}_1^*$ and $\widetilde{A}_2^*$ in place of A_1 and A_2, and with $\gamma^{\text{out}} = \gamma_{12}^{\text{out}} = \gamma$.

Lemma 4.30. *The compressions $\widetilde{A}_1^*$ and $\widetilde{A}_2^*$ satisfy the three term relationship*

$$\sigma_2 \widetilde{\Phi} \widetilde{A}_1^* - \sigma_1 \widetilde{\Phi} \widetilde{A}_2^* + \gamma \widetilde{\Phi} = 0,$$

where $\sigma_1, \sigma_2, \gamma$ are the matrices that appear in the three term relationship (4.1) for the pool Б associated with the separating structure ω according to Theorem 4.9.

Proof. Multiplying the three term relationship (4.1) for Б by P_{M_-} on the left and $\widetilde{\Phi}$ on the right, using (4.48) and $\Phi\widetilde{\Phi} = \widetilde{\Phi}$ (which is true because $\Phi = P_M$), we get

$$i\beta_2 P_{M_+} A_1 \widetilde{\Phi} - i\beta_1 P_{M_+} A_2 \widetilde{\Phi} + P_{M_-} \gamma \widetilde{\Phi} = 0.$$

Using (4.57), we have

$$i\beta_2 P_{M_+}\widetilde{A}_1^* - i\beta_1 P_{M_+}\widetilde{A}_2^* + P_{M_-}\gamma\widetilde{\Phi} - i\beta_2 P_{M_+}A_1 P_{H_{0,+}}|R + i\beta_1 P_{M_+}A_2 P_{H_{0,+}}|R = 0.$$

Now we will show that since A_1 and A_2 commute, the last two terms in the left hand side of the preceding equality cancel. Since (4.36) holds, we can write A_1 and A_2 as tridiagonal matrices according to the decomposition $K = H_{0,-}\oplus M_-\oplus M_+\oplus H_{0,+}$, in a way similar to (4.19). Indeed,

$$A_j = \begin{bmatrix} * & * & 0 & 0 \\ * & * & \beta_j & 0 \\ 0 & * & * & P_{M_+}A_j P_{H_{0,+}} \\ 0 & 0 & * & * \end{bmatrix}, \qquad j = 1, 2.$$

Multiplying the second row of A_1 by the fourth column of A_2, we obtain the operator $\beta_1 P_{M_+}A_2 P_{H_{0,+}}$. Symetrically, multiplying the second row of A_2 by the fourth row of A_1, we obtain $\beta_2 P_{M_+}A_1 P_{H_{0,+}}$. Since $A_1A_2 = A_2A_1$, we must have

$$\beta_1 P_{M_+}A_2 P_{H_{0,+}} = \beta_2 P_{M_+}A_1 P_{H_{0,+}}. \tag{4.61}$$

As we have already remarked, this implies that

$$i\beta_2 P_{M_+}\widetilde{A}_1^* - i\beta_1 P_{M_+}\widetilde{A}_2^* + P_{M_-}\gamma\widetilde{\Phi} = 0.$$

Using (4.48) again and (4.55), we get

$$P_{M_-}(\sigma_2\widetilde{\Phi}\widetilde{A}_1^* - \sigma_1\widetilde{\Phi}\widetilde{A}_2^* + \gamma\widetilde{\Phi}) = 0. \tag{4.62}$$

Now we multiply (4.1) by P_{M_+} on the left and $\widetilde{\Phi}$ on the right. We get

$$-i\beta_2^* P_{M_-}A_1\widetilde{\Phi} + i\beta_1^* P_{M_-}A_2\widetilde{\Phi} + P_{M_+}\gamma\widetilde{\Phi} = 0.$$

Using (4.58), this rewrites as

$$\begin{aligned} & i\beta_2^* P_{M_-}\tau_-^{-*}P_{\widehat{M}_-}\widetilde{A}_1^* - i\beta_1^*\tau_-^{-*}P_{\widehat{M}_-}\widetilde{A}_2^* + P_{M_+}\gamma\widetilde{\Phi} \\ & - i\beta_2^*\tau_-^{-*}P_{\widehat{M}_-}A_1 P_{\widehat{H}_{0,-}}(I - \tau_-^{-*}P_{\widehat{M}_-})|R \\ & + i\beta_1^*\tau_-^{-*}P_{\widehat{M}_-}A_1 P_{\widehat{H}_{0,-}}(I - \tau_-^{-*}P_{\widehat{M}_-})|R = 0. \end{aligned}$$

Now we will see that the last two terms in the left hand side of this equality cancel. Using Lemma 4.24, we have

$$\begin{aligned} & \beta_2^*\tau_-^{-*}P_{\widehat{M}_-}A_1 P_{\widehat{H}_{0,-}} - \beta_1^*\tau_-^{-*}P_{\widehat{M}_-}A_1 P_{\widehat{H}_{0,-}} \\ & \qquad = \tau_+^{-*}\left(\widehat{\beta}_2^* P_{\widehat{M}_-}A_1 P_{\widehat{H}_{0,-}} - \widehat{\beta}_1^* P_{\widehat{M}_-}A_2 P_{\widehat{H}_{0,-}}\right). \end{aligned}$$

Using hat-symmetry in (4.61), we see that the expression in brackets is zero.

Therefore, we have

$$i\beta_2^* P_{M_-}\tau_-^{-*}P_{\widehat{M}_-}\widetilde{A}_1^* - i\beta_1^*\tau_-^{-*}P_{\widehat{M}_-}\widetilde{A}_2^* + P_{M_+}\gamma\widetilde{\Phi} = 0.$$

Using (4.48) and (4.55), this yields

$$P_{M_+}(\sigma_2\widetilde{\Phi}\widetilde{A}_1^* - \sigma_1\widetilde{\Phi}\widetilde{A}_2^* + \gamma\widetilde{\Phi}) = 0. \tag{4.63}$$

The Lemma now follows from (4.62) and (4.63). □

Theorem 4.31. *Suppose that A_1, A_2 are two selfadjoint operators on K which are included in two separating structures ω and $\widehat{\omega}$ such that $\widehat{\omega} \prec \omega$, as in (4.46). Assume that (Iβ) and (Iτ) hold (see pages 75 and 77). Define R from (4.49) and let $\widetilde{A}_j$ be the compression of A_j to R, considered as the quotient $H_+/\widehat{H}_+$, for $j = 1, 2$. Assume that the compression operators $\widetilde{A}_j$ commute. Let $\sigma_1, \sigma_2, \gamma$ be the matrices that appear in the three term relationship (4.1) for the pool Б associated with ω according to Theorem 4.9, and let $\widehat{\sigma}_1, \widehat{\sigma}_2, \widehat{\gamma}$ be the corresponding matrices for $\widehat{\omega}$. Define $\widetilde{\Phi}$ from (4.55) and let $\widehat{\widetilde{\Phi}}$ be the operator obtained from $\widetilde{\Phi}$ by applying hat-symmetry, i.e.,*

$$\widehat{\widetilde{\Phi}} = \begin{bmatrix} P_{\widehat{M}_-}|R \\ -\tau_+^{-1} P_{M_+}|R \end{bmatrix} : R \to \widehat{M} = \widehat{M}_- \oplus \widehat{M}_+.$$

Then, the following tuples are commutative vessels:

(a) $(\widetilde{A}_1^*, \widetilde{A}_2^*; R, \widetilde{\Phi}, M; \sigma_j, \gamma^{in} = \gamma - i(\sigma_1 \widetilde{\Phi}^* \widetilde{\Phi} \sigma_2 - \sigma_2 \widetilde{\Phi}^* \widetilde{\Phi} \sigma_1), \gamma^{out} = \gamma)$.

(b) $(\widetilde{A}_1, \widetilde{A}_2; R, -\widetilde{\Phi}, M; -\sigma_j, \gamma^{in} = -\gamma, \gamma^{out} = -\gamma + i(\sigma_1 \widetilde{\Phi}^* \widetilde{\Phi} \sigma_2 - \sigma_2 \widetilde{\Phi}^* \widetilde{\Phi} \sigma_1))$.

(c) $(\widetilde{A}_1, \widetilde{A}_2; R, \widehat{\widetilde{\Phi}}, \widehat{M}; \widehat{\sigma}_j, \gamma^{in} = \widehat{\gamma} - i(\widehat{\sigma}_1 \widehat{\widetilde{\Phi}}^* \widehat{\widetilde{\Phi}} \sigma_2 - \sigma_2 \widehat{\widetilde{\Phi}}^* \widehat{\widetilde{\Phi}} \sigma_1), \gamma^{out} = \widehat{\gamma})$.

(d) $(\widetilde{A}_1^*, \widetilde{A}_2^*; R, -\widehat{\widetilde{\Phi}}, \widehat{M}; -\widehat{\sigma}_j, \gamma^{in} = -\widehat{\gamma}, \gamma^{out} = -\widehat{\gamma} + i(\widehat{\sigma}_1 \widehat{\widetilde{\Phi}}^* \widehat{\widetilde{\Phi}} \sigma_2 - \sigma_2 \widehat{\widetilde{\Phi}}^* \widehat{\widetilde{\Phi}} \sigma_1),)$.

Proof. The fact that (a) is a vessel is just a consequence of the preceding Lemma 4.27 and Lemma 4.30. Then (b) is just the adjoint vessel of (a). The vessel (c) is obtained from (a) by hat-symmetry, and (d) is the adjoint vessel of (c). $\square$

It is worthy to mention that Lemma 4.21 gives a sufficient condition for the compressions $\widetilde{A}_1$ and $\widetilde{A}_2$ to commute, which is required in this Theorem.

A. Functional calculus

A functional calculus for Hilbert space operators is a homomorphism $\Psi : \mathcal{A} \to \mathcal{B}(H)$, where $\mathcal{A}$ is a certain algebra of complex functions defined on some set and $\mathcal{B}(H)$ denotes the set of bounded operators in a Hilbert space. If the algebra $\mathcal{A}$ has a unit, we require that Ψ maps the unit to the identity operator I, which is the unit of $\mathcal{B}(H)$. The easiest functional calculus we can think of is obtained by fixing an operator $T \in \mathcal{B}(H)$, putting $\mathcal{A} = \mathbb{C}[z]$, the algebra of polynomials in z and defining

$$\Psi(a_n z^n + \cdots + a_1 z + a_0) = a_n T^n + \cdots + a_1 T + a_0 I.$$

We can call this calculus the polynomial calculus for the operator T.

Usually, one wishes to extend the polynomial calculus to a larger algebra of functions. Therefore, $\mathcal{A}$ will be an algebra containing the polynomials. In many cases, $\mathcal{A}$ will have some topology defined on it, and we will require the map Ψ to have some continuity properties. For instance, it is not difficult to check that the polynomial calculus defined above is continuous when one puts the topology of uniform convergence on compact subsets of $\mathbb{C}$ on $\mathbb{C}[z]$ and the operator norm topology on $\mathcal{B}(H)$.

Whenever $\mathbb{C}[z] \subset \mathcal{A}$, we can put $T = \Psi(z)$ and say that the calculus is a calculus for the operator T. Then, it is usual to write $f(T)$ instead of $\Psi(f)$ when we know the functional calculus that we are using.

Another simple functional calculus for an operator $T \in \mathcal{B}(H)$ is the $\mathrm{Rat}(\sigma(T))$ functional calculus. Here we put $\mathcal{A} = \mathrm{Rat}(\sigma(T))$, where $\mathrm{Rat}(K)$ denotes the algebra of rational functions with poles off a compact set K. Recall that $\sigma(T)$ is the spectrum of T, the set of all $\lambda \in \mathbb{C}$ such that $T - \lambda I$ does not have a bounded inverse in $\mathcal{B}(H)$. The $\mathrm{Rat}(\sigma(T))$ functional calculus is defined by

$$\Psi(p(z)/q(z)) = p(T)q(T)^{-1},$$

where $p, q \in \mathbb{C}[z]$, q does not vanish on $\sigma(T)$, and $p(T), q(T)$ are defined using the polynomial calculus for T.

Here, one should check that the operator $q(T)$ is invertible. This bring us to another interesting property that some functional calculi have: the spectral mapping property. If the algebra $\mathcal{A}$ of a functional calculus for T is an algebra of functions defined on $\sigma(T)$, we say that the calculus has the spectral mapping property if

$$\sigma(f(T)) = f(\sigma(T)).$$

(Here the set on the right hand side is just notation for $\{f(\lambda) : \lambda \in \sigma(T)\}$. It is easy to see that to check that the operator $q(T)$ above is invertible, it suffices to check that the polynomial calculus has the spectral mapping property.

Below we will present two different and very useful types of functional calculi.

A.1. Riesz-Dunford functional calculus

The Riesz-Dunford funcional calculus is an extension of the two simple calculi defined above for functions holomorphic on $\sigma(T)$. It can be constructed more in general for Banach algebras (see, for instance, [Rud91, Section 10.21]), but here we will stick to the setting described above.

Fix an operator $T \in \mathcal{B}(H)$. We say that a function is holomorphic on $\sigma(T)$ if it is defined and holomorphic on some (open) neighbourhood of $\sigma(T)$. The algebra $\mathcal{A}$ for the Riesz-Dunford calculus is the algebra of all the functions holomorphic on $\sigma(T)$. Since by Runge's theorem any function holomorphic on $\sigma(T)$ can be approximated uniformly on compact sets by functions in $\mathrm{Rat}(\sigma(T))$, the Riesz-Dunford calculus could be defined by using the continuity properties of the $\mathrm{Rat}(\sigma(T))$ calculus for T and taking limits. However, it is much more powerful to use a generalization of Cauchy's integral formula.

Take a function f be holomorphic on $\sigma(T)$, and let U_f be some neighbourhood of $\sigma(T)$ on which f is holomorphic and Γ_f a contour in U_f which "surrounds" $\sigma(T)$. By "surrounds" we mean that

$$\frac{1}{2\pi i}\int_{\Gamma_f}\frac{d\lambda}{\lambda - z} = \begin{cases} 1, & \text{if } z \in \sigma(T), \\ 0, & \text{if } z \notin U_f, \end{cases}$$

so that Cauchy's intregral formula holds in $\sigma(T)$ for functions holomorphic in U_f, using the contour Γ_f (see [Rud87, Section 10.34] for a statement of the "global" Cauchy integral formula). Then we define

$$f(T) = \Psi(f) = \frac{1}{2\pi i}\int_{\Gamma_f} f(\lambda)(\lambda I - T)^{-1} d\lambda.$$

The map Ψ defined above is indeed a homomorphism which extends the $\mathrm{Rat}(\sigma(T))$ calculus. It has the following continuity property: if $\{f_n\}_{n=1}^{\infty}$ and f are holomorphic on some neighbourhood U of $\sigma(K)$ and f_n converges to f uniformly on compact subsets of U, then $f_n(T)$ converges to $f(T)$ in the operator norm topology. Moreover, this functional calculus satisfies the spectral mapping property. Also, it is clear from the definition that for any $f \in \mathcal{A}$, the operator $f(T)$ commutes with every operator S that commutes with T.

A construction which can be done using the Riesz-Dunford calculus is that of the Riesz projections. Let K be a subset of $\sigma(T)$ which is closed and open in the relative topology. We define f_K to be constantly 1 on a small neighbourhood of K and constantly 0 on another disjoint small neighbourhood of $\sigma(T) \setminus K$. Then f_K is holomorphic on $\sigma(T)$. The Riesz projection associated to K is $\Pi_K T = f_K(T)$. It is a parallel projection, i.e., $(\Pi_K T)^2 = \Pi_K T$, because $f_K^2 = f_K$.

Using the definition for the functional calculus, we see that $\Pi_K T$ has the formula

$$\Pi_K T = \frac{1}{2\pi i}\int_{\Gamma_K} (\lambda I - T)^{-1} d\lambda,$$

where Γ_K is a contour which "surrounds" K and "leaves $\sigma(T) \setminus K$ outside".

The subspace $(\Pi_K T)H$ is invariant for T. If $K_1, \ldots, K_n$ is a partition of $\sigma(T)$ into subsets which are closed and open in the relative topology, then we have

$$I = \Pi_{K_1} T + \ldots + \Pi_{K_n} T,$$

and

$$(\Pi_{K_j} T)(\Pi_{K_k} T) = 0, \qquad j \neq k.$$

This implies that the space H decomposes in direct sum of subspaces invariant for T:

$$H = (\Pi_{K_1} T)H \dotplus \ldots \dotplus (\Pi_{K_n} T)H.$$

In the case when H is finite-dimensional, the spectrum of T is a finite collection of points, which are the eigenvalues of T, $\lambda_1, \ldots, \lambda_n$. Then, we can consider the partition of $\sigma(T)$ given by the subsets $K_j = \{\lambda_j\}$, to obtain Riesz projections $\Pi_{\lambda_j} T$ associated with each of the eigenvalues of T. The subspaces $(\Pi_{\lambda_j} T)H$ are then the generalized eigenspaces (also called root spaces) which appear in the Jordan normal form of T.

A.2. Spectral theorem and functional calculus for normal operators

The spectral theorem for normal operators can be seen as a consequence of the theory of commutative C^*-algebras and the Gelfand transform (see [Rud91, Section 12.22]). One of the most common forms of the spectral theorem consists in representing a normal operator T as an integral against a projection valued measure.

Let $\mathfrak{M}$ be a σ-algebra on a set X. A function $E : \mathfrak{M} \to \mathcal{B}(H)$ is called an spectral measure if it satisfies the following properties:

(a) $E(\emptyset) = 0$ and $E(X) = I$.

(b) $E(A)$ is an orthogonal projection for every $A \in \mathfrak{M}$.

(c) $E(A \cap B) = E(A)E(B)$ for every $A, B \in \mathfrak{M}$.

(d) If $A, B \in \mathfrak{M}$ and $A \cap B = \emptyset$, then $E(A \cup B) = E(A) + E(B)$.

(e) For every pair of vectors $h, h' \in H$, the function defined by $A \mapsto \langle E(A)h, h'\rangle$ is a complex measure.

If E is a spectral measure on X and $f : X \to \mathbb{C}$ is bounded and measurable, the integral

$$T = \int_X f(\lambda)\, dE(\lambda)$$

defines a bounded operator on H by

$$\langle Th, h'\rangle = \int_X f(\lambda)\, \langle dE(\lambda)h, h'\rangle, \qquad \forall h, h' \in H.$$

The spectral theorem states that if $T \in \mathcal{B}(H)$ is a normal operator then there is a spectral measure E on the Borel subsets of $\sigma(T)$ such that

$$T = \int_{\sigma(T)} \lambda\, dE(\lambda).$$

The spectral projections $E(A)$ for A a Borel subset of $\sigma(T)$ commute with every operator S which commutes with T. Also, the support of E is $\sigma(T)$, in the sense that $E(U) \neq 0$ for every nonempty (relatively) open subset U of $\sigma(T)$.

This allows us to define a functional calculus for T. If $f : \sigma(T) \to \mathbb{C}$ is bounded Borel, we can define

$$f(T) = \int_{\sigma(T)} f(\lambda)\, dE(\lambda). \tag{A.1}$$

This defines a functional calculus $\Psi : \mathcal{A} \to \mathcal{B}(H)$, where $\mathcal{A}$ is the algebra of bounded Borel functions on $\sigma(T)$. This calculus extends the Riesz-Dunford functional calculus. Moreover, it satisfies

$$\|f(T)\| \leq \sup_{z \in \sigma(T)} |f(z)|,$$

and $\overline{f}(T) = f(T)^*$. The operator $f(T)$ commutes with every S which commutes with T.

Let us recall that selfadjoint operators and unitary operators are particular kinds of normal operators. Moreover, $\sigma(T) \subset \mathbb{R}$ for a selfadjoint operator and $\sigma(T) \subset \mathbb{T}$ for a unitary operator.

A functional calculus for matrix-valued functions can also be defined (see Section 2.6 for the notation). If $f : \sigma(T) \to M_s$ is bounded Borel, then we can also define $f(T)$ using (A.1), where the integral is defined componentwise: if $f(\lambda) = [f_{jk}(\lambda)]_{jk}$, then

$$f(T) = \int_{\sigma(T)} f(\lambda)\, dE(\lambda) = \left[\int_{\sigma(T)} f_{jk}(\lambda)\, dE(\lambda)\right]_{jk}$$

This calculus extends the $M_s[z]$ calculus defined on Section 2.6. The usual properties of the scalar-valued functional calculus also hold for the matrix-valued calculus. In particular, we have the inequality $\|f(T)\| \leq \sup_{z\in\sigma(T)} \|f(z)\|$.

A version of the spectral theorem for commuting normal operators also exists (see, for instance [AM02, Appendix D]). Let $T_1, \ldots, T_n \in \mathcal{B}(H)$ be commuting normal operators. Note that in this case, all the operators $T_1, \ldots, T_n, T_1^*, \ldots, T_n^*$ commute by the Fuglede-Putnam-Rosenblum theorem (see [Rud91, Theorem 12.16]). Then there is a compact set $X \subset \mathbb{C}^n$ and a spectral measure on the Borel subsets of X such that

$$T_j = \int_X \lambda_j\, dE(\lambda_1, \ldots, \lambda_n), \qquad j = 1, \ldots, n.$$

A functional calculus can be defined for bounded Borel functions on X by

$$f(T_1, \ldots, T_n) = \int_X f(\lambda_1, \ldots, \lambda_n)\, dE(\lambda_1, \ldots, \lambda_n).$$

This calculus extends the usual $\mathbb{C}[z_1, \ldots, z_n]$ polynomial calculus and has analogous properties to the calculus for a single operator. A calculus for matrix-valued functions can also be defined. If $T_1, \ldots, T_n$ are selfadjoint, then $X \subset \mathbb{R}^n$, and if $T_1, \ldots, T_n$ are unitary, then $X \subset \mathbb{T}^n$.

Bibliography

[ABHN11] W. Arendt, C. J. K. Batty, M. Hieber, and F. Neubrander, *Vector-valued Laplace transforms and Cauchy problems*, 2nd ed., Monographs in Mathematics, vol. 96, Birkhäuser/Springer Basel AG, Basel, 2011. MR2798103 (2012b:47109)

[Agl90] J. Agler, *On the representation of certain holomorphic functions defined on a polydisc*, Topics in operator theory: Ernst D. Hellinger memorial volume, Oper. Theory Adv. Appl., vol. 48, Birkhäuser, Basel, 1990, pp. 47–66. MR1207393 (93m:47013)

[AKM12] J. Agler, G. Knese, and J. E. McCarthy, *Algebraic pairs of isometries*, J. Operator Theory **67** (2012), no. 1, 215–236. MR2881540 (2012m:47034)

[AM02] J. Agler and J. E. McCarthy, *Pick interpolation and Hilbert function spaces*, Graduate Studies in Mathematics, vol. 44, American Mathematical Society, Providence, RI, 2002. MR1882259 (2003b:47001)

[BV03] J. A. Ball and V. Vinnikov, *Overdetermined multidimensional systems: state space and frequency domain methods*, Mathematical systems theory in biology, communications, computation, and finance (Notre Dame, IN, 2002), IMA Vol. Math. Appl., vol. 134, Springer, New York, 2003, pp. 63–119, DOI 10.1007/978-0-387-21696-6_3, (to appear in print). MR2043236 (2005b:47170)

[CD13] M.-D. Choi and K. R. Davidson, *A 3×3 dilation counterexample*, Bulletin of the London Mathematical Society **45** (2013), no. 3, 511-519, DOI 10.1112/blms/bds109.

[CMR06] J. A. Cima, A. L. Matheson, and W. T. Ross, *The Cauchy transform*, Mathematical Surveys and Monographs, vol. 125, American Mathematical Society, Providence, RI, 2006. MR2215991 (2006m:30003)

[CD75] M. J. Crabb and A. M. Davie, *von Neumann's inequality for Hilbert space operators*, Bull. London Math. Soc. **7** (1975), 49–50. MR0365179 (51 #1432)

[Con91] J. B. Conway, *The theory of subnormal operators*, Mathematical Surveys and Monographs, vol. 36, American Mathematical Society, Providence, RI, 1991. MR1112128 (92h:47026)

[Dru83] S. W. Drury, *Remarks on von Neumann's inequality*, Banach spaces, harmonic analysis, and probability theory (Storrs, Conn., 1980/1981), Lecture Notes in Math., vol. 995, Springer, Berlin, 1983, pp. 14–32, DOI 10.1007/BFb0061886, (to appear in print). MR717226 (85b:47006)

[Dur70] P. L. Duren, *Theory of H^p spaces*, Pure and Applied Mathematics, Vol. 38, Academic Press, New York, 1970. MR0268655 (42 #3552)

[FF90] C. Foias and A. E. Frazho, *The commutant lifting approach to interpolation problems*, Operator Theory: Advances and Applications, vol. 44, Birkhäuser Verlag, Basel, 1990. MR1120546 (92k:47033)

[GH81] B. H. Gross and J. Harris, *Real algebraic curves*, Ann. Sci. École Norm. Sup. (4) **14** (1981), no. 2, 157–182. MR631748 (83a:14028)

[GKVVW09] A. Grinshpan, D. S. Kaliuzhnyi-Verbovetskyi, V. Vinnikov, and H. J. Woerdeman, *Classes of tuples of commuting contractions satisfying the multivariable von Neumann inequality*, J. Funct. Anal. **256** (2009), no. 9, 3035–3054, DOI 10.1016/j.jfa.2008.09.012. MR2502431 (2010f:47015)

[JKM12] M. T. Jury, G. Knese, and S. McCullough, *Nevanlinna-Pick interpolation on distinguished varieties in the bidisk*, J. Funct. Anal. **262** (2012), no. 9, 3812–3838, DOI 10.1016/j.jfa.2012.01.028. MR2899979

[Koo98] P. Koosis, *Introduction to H_p spaces*, 2nd ed., Cambridge Tracts in Mathematics, vol. 115, Cambridge University Press, Cambridge, 1998. With two appendices by V. P. Havin [Viktor Petrovich Khavin]. MR1669574 (2000b:30052)

[LA89] M. S. Livšic and Y. Avishai, *A study of solitonic combinations based on the theory of commuting nonselfadjoint operators*, Linear Algebra Appl. **122/123/124** (1989), 357–414, DOI 10.1016/0024-3795(89)90659-9. MR1019994 (90m:58089)

[LKMV95] M. S. Livšic, N. Kravitsky, A. S. Markus, and V. Vinnikov, *Theory of commuting nonselfadjoint operators*, Mathematics and its Applications, vol. 332, Kluwer Academic Publishers Group, Dordrecht, 1995. MR1347918 (96m:47003)

[Lot94] B. A. Lotto, *von Neumann's inequality for commuting, diagonalizable contractions. I*, Proc. Amer. Math. Soc. **120** (1994), no. 3, 889–895, DOI 10.2307/2160483. MR1169881 (94e:47012)

[LS94] B. A. Lotto and T. Steger, *von Neumann's inequality for commuting, diagonalizable contractions. II*, Proc. Amer. Math. Soc. **120** (1994), no. 3, 897–901, DOI 10.2307/2160484. MR1169882 (94e:47013)

[MAR07] R. A. Martínez-Avendaño and P. Rosenthal, *An introduction to operators on the Hardy-Hilbert space*, Graduate Texts in Mathematics, vol. 237, Springer, New York, 2007. MR2270722 (2007k:47051)

[Mir95] R. Miranda, *Algebraic curves and Riemann surfaces*, Graduate Studies in Mathematics, vol. 5, American Mathematical Society, Providence, RI, 1995. MR1326604 (96f:14029)

[Nik86] N. K. Nikol′skiĭ, *Treatise on the shift operator*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 273, Springer-Verlag, Berlin, 1986. Spectral function theory; With an appendix by S. V. Hruščev [S. V. Khrushchëv] and V. V. Peller; Translated from the Russian by Jaak Peetre. MR827223 (87i:47042)

[Nik02a] N. K. Nikolski, *Operators, functions, and systems: an easy reading. Vol. 1*, Mathematical Surveys and Monographs, vol. 92, American Mathematical Society, Providence, RI, 2002. Hardy, Hankel, and Toeplitz; Translated from the French by Andreas Hartmann. MR1864396 (2003i:47001a)

[Nik02b] ———, *Operators, functions, and systems: an easy reading. Vol. 2*, Mathematical Surveys and Monographs, vol. 93, American Mathematical Society, Providence, RI, 2002. Model operators and systems; Translated from the French by Andreas Hartmann and revised by the author. MR1892647 (2003i:47001b)

[Par70] S. Parrott, *Unitary dilations for commuting contractions*, Pacific J. Math. **34** (1970), 481–490. MR0268710 (42 #3607)

[Pri50] I. I. Privalov, *Graničnye svoĭstva analitičeskih funkciĭ*, Gosudarstv. Izdat. Tehn.-Teor. Lit., Moscow-Leningrad, 1950 (Russian). 2d ed.] MR0047765 (13,926h)

[Pri56] I. I. Priwalow, *Randeigenschaften analytischer Funktionen*, Zweite, unter Redaktion von A. I. Markuschewitsch überarbeitete und ergänzte Auflage. Hochschulbücher für Mathematik, Bd. 25, VEB Deutscher Verlag der Wissenschaften, Berlin, 1956 (German). MR0083565 (18,727f)

[Rud87] W. Rudin, *Real and complex analysis*, 3rd ed., McGraw-Hill Book Co., New York, 1987. MR924157 (88k:00002)

[Rud91] ———, *Functional analysis*, 2nd ed., International Series in Pure and Applied Mathematics, McGraw-Hill Inc., New York, 1991. MR1157815 (92k:46001)

[SNFBK10] B. Sz.-Nagy, C. Foias, H. Bercovici, and L. Kérchy, *Harmonic analysis of operators on Hilbert space*, Revised and enlarged edition, Universitext, Springer, New York, 2010. MR2760647 (2012b:47001)

[SV05] A. Shapiro and V. Vinnikov, *Rational transformations of systems of commuting nonselfadjoint operators* (2005), available at http://arxiv.org/abs/math/0511075v1.

[Var74] N. Th. Varopoulos, *On an inequality of von Neumann and an application of the metric theory of tensor products to operators theory*, J. Functional Analysis **16** (1974), 83–100. MR0355642 (50 #8116)

[Vin98] V. Vinnikov, *Commuting operators and function theory on a Riemann surface*, Holomorphic spaces (Berkeley, CA, 1995), Math. Sci. Res. Inst. Publ., vol. 33, Cambridge Univ. Press, Cambridge, 1998, pp. 445–476. MR1634421 (99e:47018)

[Yak98a] D. V. Yakubovich, *Subnormal operators of finite type. I. Xia's model and real algebraic curves in* $\mathbf{C}^2$, Rev. Mat. Iberoamericana **14** (1998), no. 1, 95–115, DOI 10.4171/RMI/236. MR1639287 (2000b:47060)

[Yak98b] ———, *Subnormal operators of finite type. II. Structure theorems*, Rev. Mat. Iberoamericana **14** (1998), no. 3, 623–681, DOI 10.4171/RMI/247. MR1681587 (2001d:47032)

Index

C^*-algebra, 28
D_ξ, 54
E_ξ, 53
$H^2(\mathscr{U})$, 3
$H^\infty(\mathscr{U}, \mathscr{Y})$, 4
$L^2(\mathscr{U})$, 3
$M(L)$, 7
$M_+(L)$, 7
M_z, 4
$M_{e^{it}}$, 3
N_ξ, 52
Q_ξ, 55
R, 75
$S(\xi, z)$, 43
$S(z)$, 40
V, V_*, 57, 58
W_ξ, 53
$H_{0,-}, H_{0,+}, H_-, H_+, M_-, M_+$, 56
Φ, 38, 52
Б, 52
ω, 58
Σ_ξ, 54
Θ_T, 14
$X_{\text{aff}}, X, \widehat{X}, \ldots$, 54
$\widehat{X}_-, \widehat{X}_+$, 67
Ξ, 52
α_ξ, 52
$\beta_j, \widehat{\beta}_j$, 74
γ, 52
γ_ξ, 52
$\gamma^{\text{in}}, \gamma^{\text{out}}$, 49
$\gamma^{\text{in}}_{kj}, \gamma^{\text{out}}_{kj}$, 44
ν, ν_*, 57, 58
σ, σ_k, 38, 41, 52
τ_-, τ_+, 75
$\widehat{S}(z)$, 50
ξ, 52
e_ξ, 53
z_ξ, w_ξ, 54

algebraic curve, 49, 54
  projective, 55
  real, 54, 71
algebraic variety, 48
analytic function
  bounded, 4
  constant, 6
  contractive, 5
  inner, 5
Arveson theorem, 29

Banach algebra, 28
Beurling-Lax-Halmos theorem, 15
bilateral shift, 7
blow-up, 56

Cayley-Hamilton theorem, 48, 49
channel operator, 57
channel space, 56
characteristic function, 14, 40
  complete, 43
  joint, 50
colligation, 38, 41
  commutative, 41
  coupling, 39, 42
  decomposition, 39, 41
  irreducible, 41
  strict, 38, 41
  unitary equivalence, 41
compatibility conditions, 42, 43, 45
compression, 1, 73
  generalized, 72
contraction
  $C_{0\cdot}$, 15
  $C_{\cdot 0}$, 11
  completely non-unitary, 15

desingularization, 56
dilation, 1, 23
  isometric, 8, 23

regular, 30
unitary, 10, 23
weak, 1
discriminant
curve, 49, 54, 63
halves, 67
ideal, 48
polynomial, 49
variety, 48

energy (conservation), 39, 43
extension, 1, 23

functional calculus
Borel, 85
polynomial, 83
rational, 83
Riesz-Dunford, 83

gyration, 45, 80

Hardy space, 3
hat-symmetry, 74

inner function, 5
inner space, 38

lifting, 1, 23
isometric, 8, 23

matching condition, 46
matrix-valued
functional calculus, 86
polynomial, 27
mosaic, 57

nondegenerate direction, 52

outer space, 38

Poisson formula, 4
pool, 52, 62
irreducible, 52, 63
principal subspace, 41, 44, 52
Privalov uniqueness theorem, 69
Privalov-Plemelj jump formula, 69

rate, 38, 41, 52, 78
representation, 29
restoration formula, 51, 67, 71
restriction, 1
Riemann surface, 56
Riesz projection, 55, 84

separating structure
affine, 56
orthogonal, 62
pure, 57, 63
shift
bilateral, 7
multiplicity, 7
unilateral, 7
space
inner, 38
outer, 38
spectral measure, 53, 85
subnormal operator, 58, 63
system
cascade connection, 40, 46
control, 39, 42
transfer function, 40, 51
Sz.-Nagy–Foiaş model, 11, 15

transform
almost diagonalizing, 57
Cauchy, 64, 69

unilateral shift, 7

vector bundle, 50
vessel, 45, 82
adjoint, 48
coupling, 46
decomposition, 47

wandering space, 7
window, 38