Operator Theory and Harmonic Analysis

H.S. Shapiro

Department of Mathematics
Royal Institute of Technology
S-100 44 Stockholm
Sweden
`shapiro@math.kth.se`

ABSTRACT. There are very close ties between the two subjects of the title. In the early days of operator theory, this was mainly manifested in applications of harmonic (and complex) analysis to operator theory. For example, von Neumann proved that the norm of $p(T)$, where T is any contraction on Hilbert space and p a polynomial, cannot exceed the maximum of $|p(z)|$ for z on the unit circle, as an application of complex analysis. In like manner, Stone's theorems on groups of unitary operators were proved with the aid of Bochner's theorem characterizing those functions on the circle (or on the reals) which are the Fourier transforms of positive measures. The spectral theorem itself was obtained by various routes based on Bochner's theorem, the trigonometric and/or algebraic moment theorem, etc.

What is of more recent date (and the main focus of this exposition) are deep and interesting applications of operator theory to classical analysis. A breakthrough here was Bela Sz.-Nagy's discovery that every linear operator on a Hilbert space has a unitary dilation; without here entering into details this implies that many properties of general contractions can be deduced from corresponding properties of operators in the much nicer class of *unitary* ones, to be sure on a vastly bigger space but nonetheless allowing many nontrivial deductions. For example, von Neumann's theorem mentioned earlier is thus reduced to the corresponding, and much easier, problem where T is unitary. We can here truly say that operator theory, involving genuinely infinite-dimensional constructions, starts to have applications to concrete classical problems. A beautiful case in point is D. Sarason's recognition of the fundamental role played by commutativity in understanding (and getting radically new proofs of) classical interpolation and moment theorems, like those usually associated with the names of Pick-Nevanlinna, Caratheodory, and Schur. The abstract kernel of this approach was further clarified and generalized as the so-called "commutant lifting theorem" of Sz.-Nagy and Foias. This in turn led to new and far-reaching extensions of the classical results *e.g.* to matrix-valued functions, which also have important applications *e.g.* in control theory.

1. Introduction

Since the dawning of functional analysis in the early years of the twentieth century, it has had much interaction with, and inspiration from, harmonic analysis. *Formal trigonometric series* were the original impetus to various notions of generalized functions (distributions, hyperfunctions), absolutely convergent trigonometric series were the original model for what was to become the theory of normed rings (Banach algebras), and so on. Indeed, it is well known that the Lebesgue integral itself was first conceived in connection with the study of trigonometric series.

In this talk I want to focus on a part of this interaction, that between harmonic analysis (understood here in a broad sense, so as to encompass e.g. Hardy spaces of analytic functions on the unit disk) and *linear operators in Hilbert space*. That these have close links is evident if one considers (we shall further develop this point shortly) that the solution to the trigonometric moment theorem is *grosso modo* equivalent to the spectral theorem for unitary operators. The "common denominator" in this case is *positive definiteness*: characteristic for a Hilbert space H is its inner product, a positive definite Hermitian form on $H \times H$; whereas positive definite functions on *groups* (and semigroups) are a central notion in harmonic analysis.

Generally speaking, the interplay between classical analysis and functional analysis begins by someone taking a "second look" at some classical theorem, and finding that it contains the seeds of something more general. One postulates some abstract object having only part of the features of the original classical one, and tries to deduce analogous results. This procedure does not always lead to fruitful generalizations, but in rare cases "a miracle happens" and the general theory returns more than was put into it (thus, for example, the theory of commutative Banach algebras, originally modelled on absolutely convergent trigonometric series, turns out to be the tool *par excellence* for studying spectral decomposition of normal operators on Hilbert space).

In most cases there occurs no such miracle, but none the less the perspective opened up by abstract thinking may focus researchers upon new kinds of questions which in turn stimulate the development of classical analysis, and channel thought into new pathways. Thus, for example, study of the invariant subspace problem for general linear operators on Hilbert space has not thus far "paid off" by e.g. clarifying the structure of non-normal operators in a way that could be compared with what the spectral theorem accomplishes, and may never do so. But, there has been much valuable "fallout" from the massive efforts that have been expended on this problem. For example, finding the invariant subspaces for just one concrete integral operator (the so-called *Volterra operator*) has led to a new and beautiful proof of a deep "classical" theorem of Titchmarsh on the support of a convolution of two functions. Quest of invariant subspaces for *subnormal operators* has stimulated profound researches into approximation by rational functions in the complex plane, which already has produced results of independent importance.

The present talk is intended to present, for non-specialists, a small but hopefully interesting body of results illustrating the aforementioned interplay. I assume familiarity with the notion of a Hilbert space, and shall adopt the following notations and conventions.

All Hilbert spaces are *complex* and *separable*. In a Hilbert space H, $\langle f, g\rangle$ denotes the inner product of elements f and g. When several Hilbert spaces are involved, a subscript as in $\langle f, g\rangle_H$ may be used to specify in which Hilbert space the inner product is intended. By $\|f\|$ (or $\|f\|_H$) the norm of f is denoted. An *operator* T between Hilbert spaces H and K refers to a *continuous linear* map from H into K; when $K = H$, we say T is an operator on H. The set of all operators from H to K is denoted $L(H, K)$, and $L(H, H)$ is usually abbreviated $L(H)$.

The *span* of a subset E of H denotes the set of finite linear combinations of elements of E with complex coefficients. The *closed span of* E is the closure of this set.

Other notions like subspace, adjoint, etc, which will be used are completely standardized, and the reader may refer to any of the textbooks such as [AkGl], [DuSc], [Ma] or [RiSz]. Subspaces will be tacitly assumed to be *closed*. More specialized notions and notations will be defined as needed.

The following *basic geometric result* will be needed several times.

LEMMA 1.1. *Let H, H' be Hilbert spaces, and E any subset of H. Suppose φ is any injective map from E into H', such that*

$$\langle \varphi f, \varphi g\rangle_{H'} = \langle f, g\rangle_H \tag{1.1}$$

for all f, g in E. Then, there is a continuous linear map V from the closed span of E onto the closed span of $\varphi(E)$, which moreover is an isometry, such that $V|_E = \varphi$.

For the proof, see e.g. [AkGl, p. 77].

Outline of the talk. In § 2 we shall deduce the spectral theorem for unitary operators, at least in a special form, from the Herglotz-F. Riesz-Toeplitz characterization of the Fourier coefficients of positive measures on the circle, and discuss “the basic à priori inequality” that results therefrom and that underlies the *functional calculus*.

In § 3 we discuss the theorem of Wold-von Neumann-Kolmogorov on the structure of isometries, and the application of this to Beurling’s invariant subspace theorem, and to problems of extrapolation and prediction for stationary random processes.

In § 4 we discuss the theorems of B. Sz.-Nagy on *isometric lifting* and *unitary dilation* of contractive operators, with applications (von Neumann inequality, mmean ergodic theorem).

In § 5 we present some further ramifications of dilation theory, notably the “commutant lifting theorem” and discuss its application to problems of interpolation (of the type of the Pick-Nevanlinna problem) and moment theorems.

For the most part, we shall not give detailed proofs, but try to convey some of the main ideas and techniques needed in the proofs. We cannot give a complete bibliography (the Pick-Nevanlinna theorem alone would involve us with hundreds, if not several thousands, of references) but believe that the references we give, together with references in those works, suffice to give the interested reader a good orientation in the literature.

2. The spectral theorem and harmonic analysis

The starting point of our story will be the following theorem, discovered independently by G. Herglotz, F. Riesz and O. Toeplitz ([He, Ri, To]).

THEOREM 2.1. *Given a sequence $\{c_n\}_{n=\infty}^{\infty}$ of complex numbers, a necessary and sufficient condition that there exists a bounded non-negative measure μ on the (Borel sets of the) unit circle $\mathbb{T}$ satisfying*

$$\int e^{-in\theta} d\mu(\theta) = c_n \qquad , \qquad n \in \mathbb{Z} \tag{2.1}$$

is: for every $N \geqslant 0$, and all choices of complex numbers $t_0, t_1, \ldots, t_N$ we have

$$\sum_{j,k=0}^{N} c_{j-k} t_j \bar{t}_k \geqslant 0. \tag{2.2}$$

REMARKS.

(i) Condition (2.2) can also be formulated so that the principal finite sections of the infinite matrix

$$\begin{bmatrix} c_0 & c_1 & c_2 & c_3 & \ldots \\ c_{-1} & c_0 & c_1 & c_2 & \ldots \\ c_{-2} & c_1 & c_0 & c_1 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{bmatrix} \tag{2.3}$$

are non-negative definite matrices. The matrix (2.3) is called a *Toeplitz matrix*; this designates a matrix whose entries are constant along any diagonal parallel to the main diagonal.

(ii) From (2.2) it follows readily that $c_{-n} = \bar{c}_n$ for all n (in particular, c_0 is real (and non-negative)). These relations also follow at once from (2.1).

(iii) This theorem is nowadays seen as a special case of a more general theorem of Bochner about positive functions on locally compact Abelian groups, see e.g. [Ru].

(iv) In some variants one replaces (2.1) by the equivalent statement that the harmonic function

$$u(r,\theta) := \sum_{n=-\infty}^{\infty} c_n r^{|n|} e^{in\theta} \tag{2.4}$$

in the open unit disk $\mathbb{D}$ is non-negative.

We refer for the proof of Theorem 2.1 to [Zy, p. 138].

One of the well-known deductions of the *spectral theorem for unitary operators* is based on Theorem 2.1. Let us trace the main ideas. Recall first that an operator U in $L(H)$ is *unitary* if it is a bijection of H that preserves inner products: $\langle Uf, Ug\rangle = \langle f, g\rangle$ for all f, g in H (or, what comes to the same, $U^*U = UU^* = I$, the identity operator).

DEFINITION 2.2. *The unitary operator* U has simple spectrum *if there exists a vector* $f \in H$ *such that the linear manifold spanned by* $\{U^n f : n \in \mathbb{Z}\}$ *is dense in* H.

THEOREM 2.3. (spectral theorem for unitary operators with simple spectrum). *If* $U \in \mathcal{L}(H)$ *is unitary with simple spectrum, there is a bounded positive measure* μ *on* $\mathbb{T}$ *such that the linear operator* M *on* $L^2(\mathbb{T}, d\mu)$ *defined by*

$$M : \varphi \mapsto z\varphi \qquad , \qquad \varphi \in L^2(\mathbb{T}, d\mu) \tag{2.5}$$

where $z = e^{i\theta}$ *denotes a generic point of* $\mathbb{T}$, *is unitarily equivalent to* U.

If H and K are Hilbert spaces, operators $A \in L(H)$ and $B \in L(K)$ are *unitarily equivalent* if and only if there is a unitary map $U \in L(H, K)$ such that the diagram

$$\begin{array}{ccc} H & \xrightarrow{A} & H \\ {\scriptstyle U}\big\downarrow & & \big\downarrow{\scriptstyle U} \\ K & \xrightarrow{B} & K \end{array} \tag{2.6}$$

commutes i.e. $BU = UA$. (This is also expressible as "U *intertwines* A and B"). The analogous weaker relation when U is merely assumed to have an inverse in $L(K, H)$ is called *similarity*.

Remark. Observe that M has a simple spectrum, since trigonometric polynomials are dense in $L^2(\mathbb{T}, \mu)$.

Proof of Theorem. Define

$$c_n = \langle U^n f, f\rangle \qquad , \qquad n \in \mathbb{Z} \tag{2.7}$$

where $f \in H$ is such that $\{U^n f\}_0^\infty$ span H. Then, $\{c_n\}$ satisfy (2.2), since

$$\begin{aligned} \sum_{j,k=0}^{N} c_{j-k} t_j \bar{t}_k &= \sum_{j,k=0}^{N} \langle U^{j-k} f, f\rangle t_j \bar{t}_k \\ &= \left\| \sum_{j=0}^{N} t_j U^j f \right\|^2 \geqslant 0. \end{aligned}$$

Hence, by Theorem 2.1, there is μ in $M^+(\mathbb{T})$ (the class of bounded positive measures on $\mathbb{T}$) such that (2.1) holds, and hence

$$\langle U^n f, f\rangle = \int e^{-in\theta} d\mu(\theta) \qquad , \qquad n \in \mathbb{Z}.$$

Writing $n = j - k$, this says that $\langle U^j f, U^k f\rangle$ equals the inner product of $e^{-ij\theta}$ with $e^{-ik\theta}$ in $L^2(\mathbb{T}, d\mu)$. By virtue of Lemma 1.1, there exists a unitary map V from H onto $L^2(\mathbb{T}, d\mu)$ such that

$$VU^j f = e^{-ij\theta} \qquad , \qquad j \in \mathbb{Z}.$$

Then, if M_- denotes multiplication by $e^{-i\theta}$ on $L^2(\mathbb{T}, d\mu)$, the diagram

$$\begin{array}{ccc} H & \xrightarrow{U} & H \\ {\scriptstyle V}\big\downarrow & & \big\downarrow{\scriptstyle V} \\ L^2(\mathbb{T}, d\mu) & \xrightarrow[M_-]{} & L^2(\mathbb{T}, d\mu) \end{array}$$

commutes (it is enough to check that $M_-Vg = VUg$ holds for a set of g whose linear combinations are dense in H, e.g. for $g = U^n f$ with $n \in \mathbb{Z}$, and that is immediate). This proves the theorem (with M_- instead of M, but clearly M_- and M are unitarily equivalent).

COROLLARY 2.4. *For U as in Theorem 2.3, if $p(z) = \sum_{-N}^{N} a_j z^j$, then*

$$\|p(U)\| \leqslant \max_{z\in\mathbb{T}} |p(z)|. \tag{2.8}$$

Proof. Because of the unitary equivalence, it is enough to show that $\|p(M)\| \leqslant \max|p(e^{i\theta})|$, i.e. that $\|p(M)g\| \leqslant \max\left|p\big(e^{i\theta}\big)\right| \cdot \|g\|$ for all $g \in L^2(\mathbb{T}, \mu)$. But $p(M)g = p(e^{i\theta})g$ so

$$\|p(M)g\|^2 = \int \left|p\big(e^{i\theta}\big)\right|^2 \left|g\big(e^{i\theta}\big)\right|^2 d\mu \leqslant \big(\max\left|p\big(e^{i\theta}\big)\right|^2\big)\|g\|^2 \tag{2.9}$$

which completes the proof.

Now, what about a unitary operator that does not have simple spectrum? Suppose f_1 is any unit vector in H, and let H_1 denote the closed span of $\{U^n f_1 : n \in \mathbb{Z}\}$. If $H_1 = H$ we are in the case already treated, so suppose the contrary. Then, pick a unit vector f_2 in $H \ominus H_1$ (the orthogonal complement of H_1) and let $H_2 :=$ closed span of $\{U^n f_2 : n \in \mathbb{Z}\}$. After at most a countable number of steps in this way, we obtain a decomposition of H as the sum of mutually orthogonal subspaces H_j such that:

(i) H_j is invariant for U and $U^* = U^{-1}$ (in other words, the H_j *reduce* U).
(ii) The restriction of U to H_j (also called the *part of U in H_j*) is unitarily equivalent to the operator "multiplication by z" on $L^2(\mathbb{T}, d\mu_j)$ for some measure $\mu_j \in M^+(\mathbb{T})$.

Therefore, introducing a new Hilbert space K, the *direct sum* of the $L^2(\mathbb{T}, d\mu_j)$ (whose elements are thus vectors $\varphi = (\varphi_1, \varphi_2, \dots)$ where $\varphi_j \in L^2(\mathbb{T}, d\mu_j)$), and with $\|\varphi\|^2 := \sum \int |\varphi_j|^2 d\mu_j$ it is easy to check that V is unitarily equivalent to "multiplication by z" on K. Then (once some boring questions about measurability etc. are disposed of) we can deduce (now, with no hypothesis of "simple spectrum"):

THEOREM 2.5. *If U is unitary, then*

$$\|p(U)\| \leqslant \max \left|p\left(e^{i\theta}\right)\right| =: \|p\|_\infty \tag{2.10}$$

for every Laurent polynomial $p(z) = \sum_{-N}^{N} c_k z^k$.

Remarks. We could call (2.10) the *fundamental à priori inequality* for unitary operators. It is the "high ground" from which one can (by many laborious steps, to be sure) build up the full spectral theorem for unitary operators (which we have no need to formulate here; for this whole development the reader is referred to [RiSz, § 109] or [AkGl, Chapter 6]).

The "model" we have constructed for U, as "multiplication by z on K" is not a fully satisfactory one (even though it suffices for the deduction of the important Theorem 2.5) because K is not a uniquely determined or "canonical" space, in general many different choices for the H_j are possible. This is at bottom the problem called "spectral multiplicity" and can be dealt with by methods in [DuSc, Chapter 10] or [Ha1]. One can also bypass it, and use Theorem 2.5 as the point of departure for the full spectral theorem (see remarks and references below).

A parallel development can be done for self-adjoint operators: if $T = T^*$, F.J. Murray showed by an elementary argument that *for every polynomial* $f(\lambda) = \sum_{k=0}^{N} a_k \lambda^k$ *we have* (λ being here restricted to real values)

$$\|f(T)\| \leqslant \max |f(\lambda)|, |\lambda| \leqslant \|T\|. \tag{2.11}$$

See [Ma, p. 100], where (2.11) is proven, and from it the spectral theorem deduced, following an argument of Eberlein. Alternatively, the spectral theorem for self-adjoint operators can be deduced from that for unitary operators (and vice versa) by using "Cayley transforms".

Let us also observe that the spectral theorem for several *commuting* unitary operators can be deduced from the *multivariable generalization* of the Herglotz-F. Riesz-Toeplitz theorem (itself a special case of Bochner's theorem) by adapting the argument we have given above in the case of one operator. Again, using Cayley transforms this can be carried over to several commuting self-adjoint operators. (The case of *two* commuting self-adjoint operators is equivalent to the case of one "normal" operator, i.e. an operator commuting with its adjoint).

We shall not dwell on the matter. There are a great many known proofs and variants of the spectral theorem. One of the most powerful methods known for proving them uses the theory of commutative Banach algebras to obtain the *Gelfand-Naimark theorem*, giving spectral resolutions for *commuting families of normal operators* ([DuSc, Chapter 9]. The "complete spectral theorem" of von Neumann gives a canonical representation of a normal operator as a "direct integral" independently of any multiplicity assumptions regarding the spectrum.

It is characteristic that, regardless of the particular features of these various approaches, they are all deeply influenced by ideas originating in harmonic analysis. In the opposite direction, the spectral theorem has applications to harmonic analysis: the reader is invited to deduce the Herglotz-F. Riesz-Toeplitz theorem from Theorem 2.3.

Before ending this section, a few words are in order about "functional calculus". For any operator T, with spectrum $\sigma(T)$, there is an obvious way to define $f(T)$, where f is any *rational function with no poles on* $\sigma(T)$ (for example, for $f(z) = (z-a)^{-1}$, $f(T)$ is defined to be $(T-aI)^{-1}$, where I denotes the identity operator, which is well defined for $a \in \mathbb{C} \setminus \sigma(T)$). We thus get a "functional calculus", that is a map $f \mapsto f(T)$ which is a *continuous algebra-homomorphism* from the set of rational functions with no poles on $\sigma(T)$, to $L(H)$. Moreover, a more general calculus (the "Riesz-Dunford calculus") can be built up, based on Cauchy integrals, the "domain" of which is *all functions* f *holomorphic on a neighborhood of* $\sigma(T)$. One of the great thematic problems of operator theory is to enlarge the domain of functional calculus beyond these holomorphic f. But, this cannot be done so as to apply to all operators. For special classes of operators T, however, one can widen the class of f such that $f(T)$ has meaning. This is closely related to estimates for $\|f(T)\|$, when f is rational; we shall return to this point in $\S$ 4 in connection with von Neumann's inequality. For example, the estimate (2.10) makes possible a useful definition of $f(U)$ for any unitary operator U, for any f continuous on the unit circle and even (ultimately) for any bounded Borel function on $\sigma(U)$; full details of this are in the already cited textbooks.

3. Isometries, and the Wold decomposition

Once we leave the category of normal operators, we cannot expect unitarily equivalent models so simple and transparent as "multiplication by some bounded function on an $L^2(\mu)$ space". Nevertheless there are important non-normal operators that impose themselves on us, and much of the operator theory of recent years is dedicated to finding structural properties, or models of some kind, for various classes of non-normal operators. In this section we discuss *isometries*. A fundamental theorem attributed to von Neumann, Wold, and Kolmogorov (independently) is

THEOREM 3.1. ("Wold decomposition"). *Let V be an isometry on a Hilbert space H. Then, there is a (uniquely determined) decomposition of H as the direct sum of two mutually orthogonal subspaces $H = H_1 \oplus H_2$ such that*

(i) *H_1 and H_2 reduce V.*
(ii) *The part of V in H_1 is unitary.*
(iii) *The part of V in H_2 is unitarily equivalent to a block shift.*

Before proceeding, let us explain the terminology. If K is any Hilbert space, we may construct from it a new one $\widetilde{K}$ whose elements are sequences from K, that is vectors

$$\tilde{k} := (k_0, k_1, k_2, \dots) \qquad , \qquad k_j \in K$$

such that $\|\tilde{k}\|^2_{\widetilde{K}} := \sum_0^\infty \|k_j\|^2 < \infty$. The map $S \in L(\widetilde{K})$ taking $\tilde{k}$ to $(0, k_0, k_1, \dots)$ is called the *shift operator* on $\widetilde{K}$. It is an isometry, but certainly (assuming K has dimension at least one) it is not surjective, and so not unitary. It is not hard to show that, for any two Hilbert spaces K_1 and K_2, the shifts on $\widetilde{K}_1$ and $\widetilde{K}_2$ are unitarily equivalent if and only if dim K_1 (the dimension of K_1) equals dim K_2. Thus, each shift has associated to it a unique unitary invariant, a positive integer or $+\infty$ which is the dimension of the space K from which $\widetilde{K}$ is formed. It is called the *multiplicity* of the shift. Since some authors use the term to denote exclusively the shift of multiplicity one, we use the term *block shift* to denote a general shift.

A shift (or block shift) not only is not unitary, it is as far from unitarity as an isometry can be. Indeed, clearly

$$\lim_{n\to\infty} S^n \tilde{k} = 0 \qquad , \qquad \text{for all } \tilde{k} \in \widetilde{K}.$$

We may remark, in passing, that this relation is the basis for an abstract (axiomatic) definition of a block shift (cf. [GoGoKa]). The block shift as defined above is also called *unilateral*; this is something of a misnomer—the same shift is called *bilateral* if the ambient Hilbert space of sequences is *doubly infinite* (like $(\dots k_{-1}, k_0, k_1, \dots)$) in which case it is unitary.

If we examine the proof of Theorem 3.1 (which we shall not give here, see [GoGoKa, p. 654]) we note the following features:

a) H_1 is the intersection of the ranges of all V^n, $n \geqslant 1$.
b) H_2 is a block shift of dimension dim W, where W is the kernel of V^*. Moreover, H_2 is the direct sum of the subspaces $W, VW, V^2W, \dots$ which are mutually orthogonal. (For this reason, W has been christened "wandering subspace" belonging to V, by Halmos. It is also called the defect space of V for the obvious reason that, being the orthogonal complement of the range of V, it indicates how far V falls short of being unitary.)

Of course, H_1 can be $\{0\}$ (and then V is a ("pure") shift, or H_2 can be $\{0\}$ (and then V is unitary, and $W = \{0\}$).

Theorem 3.1 has some remarkable consequences, and we shall discuss two of these, plus a remarkable generalization obtained recently by S. Shimorin.

First of all, following P. Halmos [Ha2] we shall outline a deduction from Theorem 3.1 of Beurling's famous invariant subspace theorem. (A subspace M of H is *invariant* for the operator T, if $TM \subset M$.) Before formulating it, let us recall some definitions and notations. By $L^p(\mathbb{T})$ for $p > 0$ we denote the usual Lebesgue space of measurable functions on the unit circle $\mathbb{T}$, endowed with the norm $\|\ \ \|_p$ where

$$\|f\|_p = \left(\frac{1}{2\pi}\int_{\mathbb{T}} \left|f\left(e^{i\theta}\right)\right|^p d\theta\right)^{1/p} \qquad , \qquad \|f\|_\infty = \operatorname{ess\,sup}\left|f\left(e^{i\theta}\right)\right|.$$

For $p \geqslant 1$, $L^p(\mathbb{T})$ is a Banach space. By $H^p(\mathbb{T})$, for $p \geqslant 1$ we designate the subspace of $L^p(\mathbb{T})$ consisting of those functions whose negatively indexed Fourier coefficients are all zero. As is well known ([Du], [Ga]), each function in $H^p(\mathbb{T})$ is the nontangential limiting value almost everywhere of a holomorphic function g in the open unit disk $\mathbb{D}$ such that

$$\varlimsup_{r\to 1}\left(\frac{1}{2\pi}\int_0^{2\pi}\left|g\left(re^{i\theta}\right)\right|^p d\theta\right)^{1/p} =: \|g\|_{H^p(\mathbb{D})}$$

is finite. The correspondence of g with its boundary values induces an isometric isomorphism between $H^p(\mathbb{D})$ and $H^p(\mathbb{T})$. For all of this see e.g. [Du], to which we also refer for certain special terminology like *inner* and *outer* functions, etc.

The shift on the space of square summable sequences

$$a = \{a_0, a_1, a_2, \dots\}$$

of complete numbers is, in an obvious way resulting from Parseval's identity, unitarily equivalent to the operator "multiplication by $e^{i\theta}$" on $H^2(\mathbb{T})$, and also to "multiplication by z" on $H^2(\mathbb{D})$. To avoid excessive pedantry we shall feel free to use the term "shift operator" to denote any of these, according to the context of the discussion. Now, we state

THEOREM 3.2. (A. Beurling, 1949). *The invariant subspaces of the shift operator on $H^2(\mathbb{T})$ are precisely the sets*

$$\varphi H^2 := \left\{\varphi f : f \in H^2(\mathbb{T})\right\} \tag{3.1}$$

where φ is a unimodular function in $H^\infty(\mathbb{T})$ ("inner function").

Proof. It is clear that a set of the form (3.1) is a vector subspace of $H^2(\mathbb{T})$, also it is closed (being the range of the isometric operator "multiplication by φ") and invariant with respect to the shift (here, multiplication by $e^{i\theta}$). The deeper part of the theorem is the converse direction, whose proof à la Halmos [Ha2] we now proceed to sketch.

Thus, let M be a (proper) invariant subspace. Then, restricted to M, the shift is an isometry which we denote by V, and to this we apply Theorem 3.1 with M in the role of the Hilbert space H there. It is easy to check that the intersection of the ranges of V^n for $n = 1, 2, \dots$ is $\{0\}$, so V is unitarily equivalent to a block shift. It is intuitively plausible that the multiplicity of this block shift is one, but that is no tautology, and indeed verification of this is the heart of the proof. We must show that the wandering space W is one-dimensional. To this end, let φ denote any vector in W, which in our context is identified as $M \ominus e^{i\theta}M$. Thus, φ is in M, and orthogonal to $e^{i\theta}f$ to every f in M.

We can in particular choose $f = e^{ik\theta}\varphi$ for $k = 0, 1, 2, \dots$ and this gives

$$\int_{\mathbb{T}} |\varphi|^2 e^{in\theta} d\theta = 0 \qquad , \qquad n = 1, 2, \dots .$$

By complex conjugation this holds also for all negative integers n, hence $|\varphi(e^{i\theta})|^2$ is constant almost everywhere on $\mathbb{T}$.

Thus, W is a closed vector subspace of $H^2(\mathbb{T})$ with the remarkable property that every function in it has (almost everywhere on $\mathbb{T}$) constant absolute value. The reader should have no trouble supplying the deduction, from this, that dim W is at most one (hence, exactly one, since otherwise V would be a unitary map of M onto itself, which we have ruled out).

We conclude that every element f of $H^2(\mathbb{T})$ is uniquely representable in the form

$$\sum_{n=0}^{\infty} w_n e^{in\theta}$$

where $w_n \in W$ and $\sum \|w_m\|^2 < \infty$. But, $w_n = c_n\varphi$ for some fixed inner function φ, where $\sum |c_n|^2 < \infty$. Hence

$$f\left(e^{i\theta}\right) = \varphi\left(e^{i\theta}\right) \sum_0^{\infty} c_n e^{in\theta}$$

which completes the proof of the theorem.

One advantage of this proof is that it can be adapted to prove generalizations of Theorem 3.2, pertaining to Hardy spaces of *vector-valued functions* (due to Lax and Halmos, see [GoGoKa] for details and references). These generalized versions are important for applications e.g. in systems and control theory, as well as multivariate stochastic processes and seem very difficult to prove by classical arguments.

Another, very important, application of Theorem 3.1 is to *wide-sense stationary random sequences*. Indeed, this was the source of the interest in Theorem 3.1 on the part of Wold, and of Kolmogorov. For lack of space, we cannot develop this here in much detail, but refer for the background to [IbRo] and the appendix by Peller and Khruschev to [Ni].

We consider a "random sequence" $\{X_n\}_{n\in\mathbb{Z}}$. The X_n are elements of an $L^2(\Omega, d\rho)$ space for some measure space Ω and probability measure ρ on it (we omit the σ-algebra of subsets of Ω from the notation). We assume the X_n all have mean zero, and variance 1. Moreover, the sequence $\{X_n\}$ is "stationary" in the sense that the *covariance* $E(X_m\overline{X}_n)$, with E denoting mean value, depends only on $n - m$.

For many of the purposes of statistics, we can ignore the underlying "sample space" $(\Omega, d\rho)$ and study the more abstract model, whereby: $\{X_n\}_{n\in\mathbb{Z}}$ *are unit vectors in some Hilbert space* H*, and* $\langle X_m, X_n\rangle$ *depends only on* $m - n$. We may also suppose that H is the closed span of $\{X_n\}$ (otherwise, simply re-define H).

Denoting $\langle X_m, X_n\rangle := c_{m-n}$, one easily checks that $\{c_k\}_{k\in\mathbb{Z}}$ satisfy the hypotheses of Theorem 2.1, hence there is a positive measure μ on $\mathbb{T}$ satisfying $\int e^{ik\theta} d\mu(\theta) = c_k$ for all $k \in \mathbb{Z}$, and hence

$$\langle X_m, X_n\rangle = \int e^{i(m-n)\theta} d\mu(\theta) = \langle e^{im\theta}, e^{in\theta}\rangle_{L^2(\mathbb{T}, d\mu)}.$$

Thus, by Lemma 1.1, *there is a unitary map* U *from* H *onto* $L^2(\mathbb{T}, d\mu)$ *carrying* X_n *onto* $e^{in\theta}$*, for all* $n \in \mathbb{Z}$.

This makes possible the transformation of certain "statistical" questions concerning $\{X_n\}$ into purely "analytical" ones in $L^2(\mathbb{T}, d\mu)$. Consider, for example, the problem of *prediction*. Let us denote, for any integer m, the closed span of $\{X_n\}_{n \leqslant m}$ by $\mathbb{P}_m$ ("the past up to time m"). Suppose we "know" X_n for all $n < 0$, and wish to "predict" the value of X_0. A very widely used method is to use a "least-squares estimate". That is, use as the "estimate" for X_0 the random variable $\widehat{X}_0$ which is defined as the orthogonal projection of X_0 on $\mathbb{P}_{-1}$. If one assumes (as we shall) that the above unitary map U is "known" (also called "the spectral representation of the random sequence"), then this problem can be transformed to the $L^2(\mathbb{T}, d\mu)$ context, where it is a standard type of approximation problem (weighted L^2 approximation by trigonometric polynomials) and can be tackled by traditional methods (Gram-Schmidt orthogonalization, etc.). For example, the variance of the prediction error when the estimator $\widehat{X}_0$ is employed, $\|X_0 - \widehat{X}_0\|^2$ can in principle be computed as the squared distance from the constant function 1 to the closed span of $\{e^{in\theta}\}_{n<0}$ in $L^2(\mathbb{T}, \mu)$. Hopefully, this conveys a little of the flavor of "prediction theory".

Now, there is an important unitary map V of H: the map defined for all n by $VX_n = X_{n-1}$ is easily seen to extend by linearity and continuity to a unitary map of H (which we continue to denote by V). (In the "spectral model" this becomes multiplication by $e^{-i\theta}$).

Now, $\mathbb{P}_0$ is invariant for V, and $V \mid \mathbb{P}_0$ is an isometry of $\mathbb{P}_0$. Hence, by Theorem 3.1, we get a Wold decomposition of $\mathbb{P}_0$ reducing $V \mid \mathbb{P}_0$. What are the two components into which it resolves?

Let us look first at the extreme cases:

(i) $V \mid \mathbb{P}_0$ is unitary; in this case, by $V\mathbb{P}_0 = \mathbb{P}_0$, so $\mathbb{P}_0 = \mathbb{P}_{-1}$, and by stationarity, all $\mathbb{P}_n$, $n \in \mathbb{Z}$ are equal. This is the hallmark of a *purely deterministic process*. By means of the spectral model, we know this happens when, and only when, in the Lebesgue decomposition of the spectral measure μ,

$$d\mu = d\mu_s + w d\theta$$

(where μ_s is singular with respect to $d\theta$, and $w \in L^1(\mathbb{T}, d\theta)$) we have $\int_{\mathbb{T}} \log w d\theta = -\infty$. This is a famous theorem due to G. Szegö, M. Krein and A. Kolmogorov.

(ii) $V|\mathbb{P}_0$ is a block shift; in this case, $\bigcap_{n \leqslant 0} \mathbb{P}_n$ (called the "remote past") reduces to $\{0\}$. The best least squares estimate of X_0 on the basis of "old" information $\mathbb{P}_{-m}$, where m is a large positive integer, tends in norm to 0 as $m \to \infty$. The variance of the prediction error tends to 1. In other words, the best least-squares estimate of X_0 on the basis of very old observations is simply its mean value (here assumed to be zero). Such a process is called *purely indeterminate*.

The general case is an amalgam of (i) and (ii), that is, the Wold decomposition in the stochastic model says:

There is a unique splitting $X_n = X_n' + X_n''$ where $\{X_n'\}$ is purely determinate and $\{X_n''\}$ is purely indeterminate. Their closed spans H', H'' are orthogonal complements in H.

In the spectral model, the splitting can be described explicitly: it mirrors the Lebesgue decomposition of $d\mu$, such that when $\int \log w d\theta = -\infty$, $w d\theta$ is grouped together with $d\mu_s$ and we have a purely determinate process.

The prediction problem is only one of a great many problems about stationary sequences, involving questions of mixing, regularity and so forth. These translate into interesting and sometimes very deep, in some cases still unsolved problems of harmonic analysis. For further information see [IbRo], and the Peller-Khruschev appendix to [Ni].

In closing this section, let us remark that there has been much study in recent years of the invariant subspaces of the operator "multiplication by z" on the *Bergmann space* of the disk $\mathbb{D}$, that is, the space $AL^2(\mathbb{D})$ of analytic functions on $\mathbb{D}$ square integrable with respect to area measure. Here the situation is much more complicated than in the corresponding H^2 scenario of Beurling's theorem. For one thing, "multiplication by z" is now a contractive mapping, but not isometric. It turns out that there are invariant subspaces M for which $M \ominus zM$ is not one-dimensional; the dimension of this subspace (analogous to the "wandering subspace" in our discussion of Beurling's theorem) may be any positive integer, or even infinity.

Nevertheless, it has been proved by Aleman, Richter and Sundberg [AlRiSu], that the space $W := M \ominus zM$ still has the property that, together, $W, zW, z^2W, \ldots$ span M. This remarkable and deep result inspired S. Shimorin [Sh] to discover a new general theorem implying a Wold-type decomposition for certain classes of operators (including, of course, all isometries, but also many others) which in particular yields the main result of [AlRiSu].

4. Dilation theory

Let us start by reviewing some definitions. If $A \in L(H)$ and $A' \in L(H')$, where H, H' are Hilbert spaces with $H \subset H'$, A' is said to *lift* A (or, be a *lifting* of A) if the diagram

(4.1)
$$\begin{array}{ccc} H' & \xrightarrow{A'} & H' \\ {\scriptstyle P}\downarrow & & \downarrow{\scriptstyle P} \\ H & \xrightarrow[A]{} & H \end{array}$$

commutes, P being the orthogonal projector of H' on H; that is,

(4.2)
$$AP = PA'.$$

Applying A to both sides of (4.2) from the left, gives $A^2P = APA' = PA'A' = P(A')^2$, and now a simple inductive argument shows

(4.3)
$$A^nP = P(A')^n \qquad , \qquad n = 1, 2, \ldots .$$

If $T \in L(H, H')$, the operator PT (in $L(H)$) is called the *compression of* T *to* H. It is easy to check that this is equivalent to, denoting PT by A,

(4.4)
$$\langle Th_1, h_2 \rangle = \langle Ah_1, h_2 \rangle \text{ for all } h_1 \text{ and } h_2 \text{ in } H.$$

If A' is a lifting of A to H', then from (4.3)

$$\langle A^n h_1, h_2\rangle = \langle A^n P h_1, h_2\rangle = \langle P(A')^n h_1, h_2\rangle = \langle (A')^n h_1, h_2\rangle \tag{4.5}$$

for all pairs h_1, h_2 in H. Taking $n = 1$ and comparing with (4.4) shows: *if A' is a lifting of A, then A is the compression to H of the restriction $A' \mid H$.* But, lifting implies more, namely that A^n is, for every positive integer n, the compression of $(A'|H)^n$ to H.

The converse is not true, and the last relation turns out to be important enough to be given a name:

DEFINITION 4.1. *If $A \in L(H)$ and $A' \in L(H')$, where H, H' are Hilbert spaces with $H \subset H'$, A' is a dilation of A if and only if A^n is, for every positive integer n, the compression of $(A' \mid H)^n$ to H, or what is equivalent*

$$\left\langle \left(A'\right)^n h_1, h_2\right\rangle = \left\langle A^n h_1, h_2\right\rangle \tag{4.6}$$

for every pair h_1, h_2 in H and every positive integer n.

Remarks. Formerly, the term "power dilation" was sometimes used to denote this relation.

If H is an invariant subspace of A', and A is the part of A' in H, then clearly A' is a dilation of A.

Thus, "A' is a dilation of A" is implied by, but in general strictly weaker than each of the assertions "A' is an extension of A" and "A' is a lifting of A".

A landmark discovery in operator theory, due to B. Sz.-Nagy, is

THEOREM 4.2. *Every contraction (i.e. operator of norm at most one) on a Hilbert space has a dilation that is unitary.*

One can formulate this theorem more precisely: if $A \in L(H)$, then there is a unitary dilation U of A to some, in general larger Hilbert space H', with the further property that the closed span of $\{U^n H\}$ for n in $\mathbb{Z}$ is H'. This unitary dilation is, in a natural sense *minimal* in that, roughly speaking, H' is no larger than it has to be and moreover, modulo a natural concept of *isomorphisms of unitary dilations*, this minimal one is unique. (For details, we refer to [SzFo].)

To get a feeling for this remarkable result, the reader is urged to try to construct a unitary dilation (u.d.) of the operator A that is identically zero on H. Any u.d. U must satisfy $\langle U^n h_1, h_2\rangle = 0$ for all h_1, h_2 in H, from which it follows that the subspaces $\{U^n H\}$ for $n \in \mathbb{Z}$ are mutually orthogonal, i.e. H is a "doubly wandering" subspace for U in the larger space H' where U operates, so H' must be quite large!

Thanks to Theorem 4.2 some results known for unitary operators can be carried over to contractions. Thus,

COROLLARY 4.3. (von Neumann's inequality) *Let A be a contraction on the Hilbert space H, and p any polynomial, $p(\lambda) = \sum_{n=0}^N a_n \lambda^n$ with complex coefficients. Then*

$$\|p(A)\| \leqslant \max |p(\lambda)| \qquad \lambda \in \mathbb{C}, |\lambda| \leqslant 1. \tag{4.7}$$

Proof. Let $U \in L(H')$ be any unitary dilation of A. Then, for every h_1, h_2 in H we have
$$\big|\langle p(U)h_1, h_2\rangle\big| = \big|\langle p(A)h_1, h_2\rangle\big|.$$

Taking the supremum of the right side over all unit vectors h_2 in H gives
$$\|p(A)h_1\| \leqslant \|p(U)h_1\| \leqslant \|p(U)\| \, \|h_1\| \tag{4.8}$$
(the second norm being in the space H').

But, $\|p(U)\|$ does not exceed the maximum of $|p(\lambda)|$ for λ on the unit circle, as shown in § 2. Using this in (4.8) and maximizing over unit vectors h_1 in H gives (4.7), completing the proof.

Remarks. On the basis of (4.7) one can enlarge the functional calculus for contractions A, to encompass all operators $f(A)$ with f continuous on $\overline{\mathbb{D}}$ and in $H^\infty(\mathbb{D})$. This is, in turn, close to von Neumann's notion of "spectral sets", see [RiSz], [SzFo].

One can give proofs of (4.7) using only complex analysis and elementary Hilbert space theory, see [RiSz]. Here we have shown its derivation from Theorem 4.2 to illustrate a typical application of that theorem. Another very nice application is the *mean ergodic theorem*, see [RiSz].

We shall not prove Theorem 4.2, referring to [SzFo] or [GoGoKa], but will give a few indications. First of all, Theorem 4.2 follows from

THEOREM 4.4. *Every contractive operator on a Hilbert space has an isometric lifting.*

To see why this implies Theorem 4.2 we require

LEMMA 4.5. *Every isometric operator on a Hilbert space has a unitary extension.*

Assuming this for the moment, suppose A is a contraction on H and V is some isometric lifting to H'. Let U be a unitary extension of V to H''. Then, in view of earlier remarks, U is a dilation of V, and V is a dilation of A. It follows readily that U is a dilation of A.

We'll give an informal proof of Lemma 4.5. Let V be an isometry on a Hilbert space H. By virtue of the Wold decomposition there is a splitting of $H = H_1 \oplus H_2$ reducing V, such that $V \mid H_1$ is unitary and $V \mid H_2$ is unitarily equivalent to a "unilateral block shift", that is to the operator
$$S : w := (w_0, w_1, w_2, \dots) \mapsto (0, w_0, w_1, \dots)$$
on the Hilbert space of *one-sided sequences* of elements of some Hilbert space W, with norm of w equal to $(\sum_0^\infty \|w_j\|^2)^{1/2}$. (In fact, W can be taken as the wandering space $H \ominus VH$.) Now, S has an obvious unitary extension $\widetilde{S}$, namely the map of the Hilbert space of *two-sided sequences* $\tilde{w} := (\dots w_{-1}, w_0, w_1, \dots)$ given by the formula
$$\big(\widetilde{S}\tilde{w}\big)_n = \tilde{w}_{n-1} \qquad , \qquad n \in \mathbb{Z},$$
the "bilateral shift". Indeed, $\widetilde{S}$ is unitary and, restricted to the subspace for which all w_i with $i < 0$ vanish (which in an obvious way is identified with the above space of one-sided

sequences), $\widetilde{S}$ coincides with S. Thus, finally, the operator equal to V on H_1 and $\widetilde{S}$ on the appropriate space is a unitary extension of, strictly speaking, an operator unitarily equivalent to V. And, it is easily checked that if an operator has a unitary extension, then so does every operator unitarily equivalent to it (this is because a unitary operator always has a unitary extension to every larger Hilbert space).

Theorem 4.4 also allows a strengthened version, whereby the isometric lifting is to a "minimal" space, and this "minimal isometric lifting" is essentially unique, for details we refer again to [SzFo] and [GoGoKa].

As to the proof of Theorem 4.4, we won't give it in full but illustrate one of the main ideas by sketching the proof of an earlier (weaker) version due to Halmos, [1] whose pioneering writings have been instrumental for modern developments in operator theory:

THEOREM 4.6. *Every contractive operator on a Hilbert space is the compression of a unitary operator.*

Proof. Let $A \in L(H)$, $\|A\| \leqslant 1$. Then, $I - A^*A$ and $I - AA^*$ are self-adjoint operators with spectrum contained in the set $\mathbb{R}^+$ of non-negative real numbers. By the basic functional calculus for self-adjoint operators, they have "positive square roots", that is, there exist self-adjoint operators D_A and D_{A^*} with spectrum in $\mathbb{R}^+$ such that

$$\begin{aligned} D_A^2 &= I - A^*A \\ (D_{A^*})^2 &= I - AA^*. \end{aligned} \tag{4.9}$$

One calls D_A the *defect operator* associated to A. Observe that it is O if and only if A is an isometry, and both D_A and D_{A^*} are O if and only if A is unitary. Now, we have

$$A\big(A^*A\big)^n = \big(AA^*\big)^n A$$

and so, for every polynomial $p(t) = c_0 + c_1 t + \cdots + t^m$ with real coefficients $Ap(A^*A) = p(AA^*)A$. Here we can let p run through a sequence $\{p_j\}$ of polynomials converging uniformly to $t^{1/2}$ on $[0,1]$. Then, by functional calculus, $\|p_j(A^*A) - D_A\|$ and $\|p_j(AA^*) - D_{A^*}\|$ tend to zero, and we obtain the *fundamental intertwining identity*

$$AD_A = D_{A^*}A. \tag{4.10}$$

It is now easy to prove Theorem 4.6. Let A be a contraction on H. Then, we can define an operator on $H \oplus H$ by means of the "block matrix"

$$U := \begin{bmatrix} A & D_{A^*} \\ -D_A & A^* \end{bmatrix} \tag{4.11}$$

[1] A weaker version of Theorem 4.6 with "univary" replaced by "isometric" was discovered earlier by G. Julia, see [SzFo], p. 51, for the references. I am indebted to N.K. Nikolski for pointing this out to me.

with the convention that U maps an element $\left[\begin{smallmatrix} h_1 \\ h_2 \end{smallmatrix}\right]$ of $H \oplus H$ to

$$\begin{bmatrix} Ah_1 & + & D_{A^*}h_2 \\ -D_A h_1 & + & A^* h_2 \end{bmatrix}.$$

Clearly $U \in L(H \oplus H)$ and let us now verify that it fulfills the requirements of the theorem.

First of all, its compression to H (here identified as the subset of $H \oplus H$ consisting of elements $\left[\begin{smallmatrix} h_1 \\ 0 \end{smallmatrix}\right]$ with $h_1 \in H$) is A. Indeed, the first component of $U \left[\begin{smallmatrix} h_1 \\ 0 \end{smallmatrix}\right]$ is Ah_1. As for unitarity, U^* is represented by the block matrix

$$U^* = \begin{bmatrix} A^* & -D_A \\ D_{A^*} & A \end{bmatrix}$$

so, at least formally, the rules of matrix multiplication give

$$U^*U = \begin{bmatrix} A^*A + D_A^2 & A^*D_{A^*} - D_A A^* \\ D_{A^*}A - AD_A & D_{A^*}^2 + AA^* \end{bmatrix}.$$

The diagonal elements are I by virtue of (4.9), (4.10) while the element in row 2, column 1 vanishes by (4.11). The remaining off-diagonal element is the adjoint of this one, hence O. This completes the proof, for people who believe in block matrices. One can of course paraphrase all these calculations by introducing the orthogonal projector from $H \oplus H$ to $H \oplus \{0\}$ and avoiding block matrices, and it is perhaps an instructive exercise for the reader to carry out the proof in this way too. But, block matrices are a very convenient notational device.

This completes the proof of Theorem 4.6. Actually, we have proved a bit more: the unitary operator lives on $H \oplus H$. Moreover, if A is a contractive operator from H to K where K is any Hilbert space (so that $A^* \in L(K, H)$) then $A^*A \subset L(H)$, $AA^* \in L(K)$ and we can define defect operators $D_A \subset L(H)$ and $D_{A^*} \in L(K)$ as before. It is then easy to verify that *mutatis mutandis* the construction of a unitary operator as above goes through. For details see [GoGoKa] or [FoFr]. This is a highly nontrivial result even for finite matrices. It implies given any $m \times n$ matrix which is contractive from $\mathbb{C}^m$ to $\mathbb{C}^n$, we can embed it as the upper left corner of a unitary matrix of size $(m + n) \times (m + n)$ (see (4.12)).

The reader who has followed thus far should now have no trouble following the standard proofs of Theorems 4.4 and 4.2.

5. The further development of dilation theory—especially applications to interpolation problems

Since the discovery of the results reported on in the preceding section, they have continued to play a major role in the development of operator theory, especially to so-called "functional models" for contractions. This is beyond the scope of the present talk, see [Pe] (Douglas' article) and [SzFo]. But we do wish, in closing, to say something about one remarkable development of dilation theory with spectacular applications.

We shall introduce this by going back to the landmark paper [Sa1] which triggered this development. Sarason cast a new light on classical moment and interpolation problems. We can illustrate his approach with the example of the classical *interpolation problem of Pick and Nevanlinna* (henceforth abbreviated "PN problem"). In its simplest and purest form, this is:

(PN) *Given are distinct points* $z_1, \dots, z_n$ *in the open disk* $\mathbb{D}$*, and complex numbers* $w_1, \dots, w_n$*. Find necessary and sufficient conditions that there exist a holomorphic function* f *bounded in modulus by* 1 *in* $\mathbb{D}$ *(we denote this class henceforth by* B*) satisfying*

$$f(z_j) = w_j \qquad j = 1, 2, \dots, n. \tag{5.1}$$

There are many further questions that naturally arise from this one: If such f exists, is it unique? What can be said about functions in $H^\infty(\mathbb{D})$ *of least norm* satisfying (5.1)? If there is more than one f in B satisfying (5.1), can one describe the *totality* of these functions? And, how about the analogous problem with infinitely many points $\{z_j\}$? What happens if some z_j are permitted to coincide, say $z_1 = z_2$ and one prescribes the functional $f'(z_1)$ in (5.1) to replace the redundant $f(z_2)$? And so on. All of these variants have been studied, as well as many others (especially, the case where $H^\infty(\mathbb{D})$ is replaced by an analogous class of *vector valued* functions, which is important in applications). We cannot here enter into all these, but will give references at the end of this section. The remarkable thing is that the path trodden by Sarason has turned out to be fruitful in the study of all these variants, and moreover is virtually the only approach known to yield results for some of the vector-valued generalizations of (PN).

The original question was answered by the following theorem of Pick:

THEOREM 5.1. *Under the hypotheses of (PN) a necessary and sufficient condition for the existence of* f *in* B *satisfying (5.1) is that the matrix*

$$\left[\frac{1 - \bar{w}_j w_k}{1 - \bar{z}_j z_k}\right]_{j,k=1}^{n} \tag{5.2}$$

be non-negative definite.

Remarks. To get some feeling for the problem let us look first at simple cases. For $n = 1$ the problem is trivial: the desired f exists if and only if $|w_1| \leqslant 1$. Moreover, if $|w_1| = 1$, the solution is unique (the constant function equal to w_1), whereas if $|w_1| < 1$, there are infinitely many solutions, the totality of which is easily described (we shall return to this point in a moment). (The reader is invited to examine carefully the case $n = 2$, and compare with the Schwarz Lemma.)

It is remarkable that, by a recursive algorithm first proposed by I. Schur the general problem (PN) can be reduced to the trivial case $n = 1$. Indeed, whenever $f \in B$, $a \in \mathbb{D}$ and $f(a) = b$, either

(i) $|b| = 1$ and $f \equiv b$
(ii) $|b| < 1$, and in this case

$$g(z) := \frac{f(z) - b}{1 - \bar{b}f(z)} \tag{5.3}$$

is in B, and vanishes at $z = a$, so that

$$h(z) := g(z)\left(\frac{1 - \bar{a}z}{z - a}\right) = \frac{f(z) - b}{1 - \bar{b}f(z)} \cdot \frac{1 - \bar{a}z}{z - a} \tag{5.4}$$

is again a function in B. Thus, if (PN) is solvable and $|w_n| < 1$, then, taking $a = z_n$, $b = w_n$ in (5.4) we see that

$$\frac{f(z) - w_n}{1 - \bar{w}_n f(z)} \cdot \frac{1 - \bar{z}_n z}{z - z_n} =: F(z) \tag{5.5}$$

is again in B. Moreover

$$F(z_j) = \frac{w_j - w_n}{1 - \bar{w}_n w_j} \cdot \frac{1 - \bar{z}_n z_j}{z_j - z_n} \qquad (j = 1, 2, \ldots, n-1). \tag{5.6}$$

We also see from (5.5) that

$$F(z_n) = f'(z_n) \cdot \frac{1 - |z_n|^2}{1 - |w_n|^2}. \tag{5.7}$$

Thus, if $w_1, \ldots, w_n$ are admissible data for $f(z_1), \ldots, f(z_n)$ with $f \in B$ so are the n numbers appearing on the right hand sides of (5.6) and (5.7). Now, suppose we are able to solve the $(n-1)$-point version of (PN) for $F \in B$ satisfying (5.6). Then, from (5.5) solved for f we obtain a function in B. Indeed, solving (5.3) for f gives

$$f(z) = \frac{g(z) + b}{1 + \bar{b}g(z)} \tag{5.8}$$

which is in B if g is, so the solution to (5.5) is

$$f(z) = \frac{G(z) + w_n}{1 + \bar{w}_n G(z)} \tag{5.9}$$

where

$$G(z) := \left(\frac{z - z_n}{1 - \bar{z}_n z}\right) F(z) \tag{5.10}$$

is in B.

Recapitulating: If (5.6) is solvable with $F \in B$, then f defined by (5.9) and (5.10) is in B, and it is easy to check from these formulae that f satisfies all the n conditions (5.1). *Therefore*: If $|w_n| < 1$, the (PN) problem reduces to one with $n - 1$ points; whereas if $|w_n| = 1$ we see at a glance that the problem is solvable if and only if all remaining w_j are equal to w_n; but of course if $|w_n| > 1$ there is no solution. Let us call the last two cases *trivial cases*. We have thus arrived at *Schur's algorithm*:

If $|w_n| \geqslant 1$ we are done, trivially. If $|w_n| < 1$, we construct the new numbers

$$W_j := \frac{w_j - w_n}{1 - \bar{w}_n w_j} \cdot \frac{1 - \bar{z}_n z_j}{z_j - z_n} \qquad (j = 1, 2, \ldots, n-1).$$

If $|W_{n-1}| \geqslant 1$ we are done, trivially; if $|W_{n-1}| < 1$ we reduce to a problem for $n - 2$ points by the analogous formulae applied to $W_1, \ldots, W_{n-1}$; and so on.

We therefore compute recursively a sequence of complex numbers W_n, $W_{n-1}, \ldots$ and (PN) is solvable if and only if either all of these "Schur parameters" remain in $\mathbb{D}$, or one of them is on the unit circle and the remaining data at that point are equal to this one.

Volumes have been written about the Schur algorithm and various generalizations, an excellent source is [FoFr] and we'll give other references. Nevertheless, although this adequately answers (PN) for small n, it does *not*, at least in an obvious way, yield the elegant theorem 5.1, to whose proof we now return.

Following in Sarason's footsteps: if we are to find an operator-theoretic proof of Theorem 5.1, which involves an unknown function in H^∞, we must start by "encoding" such functions as linear operators on a suitable Hilbert space. Let us try $H^2 = H^2(\mathbb{T})$ as our Hilbert space. Then, any $\varphi \in H^\infty(\mathbb{T})$ induces a linear operator on H^2 by associating with it the *multiplication operator* $M_\varphi : g \mapsto \varphi g$ on H^2. It is easy to show the operator norm $\|M_\varphi\|$ equals $\|\varphi\|_\infty$. Moreover, these multiplication operators are distinguished within $L(H^2)$ by a simple property: denoting by M_z the "shift" on L^2 (i.e. multiplication by the independent variable) we have

LEMMA 5.2. (Sarason).*) If $A \in L(H^2)$ and A commutes with M_z, then $A = M_\varphi$ for some $\varphi \in H^\infty$.*

The converse is obvious. The reader should be able to prove Lemma 5.2, starting with the simple observations that A commutes with M_p for every polynomial p: $A(pf) = p(Af)$ for all $f \in H^2$, hence $Ap = \varphi p$ for every polynomial p, where $\varphi := A1$. The main point is to deduce that this φ, so far only known to be in H^2, is in H^∞.

Now that we have a nice way to encode H^∞ as operators on H^2, we need a suitable way to encode the data (5.1). Sarason had the inspired idea to use for this purpose the *Toeplitz operator with symbol* $\bar{f}$. For any $\varphi \in L^\infty(\mathbb{T})$, "multiplication by φ" is a bounded operator on $L^2(\mathbb{T})$. Its compression to the subspace $H^2(\mathbb{T})$ is called the *Toeplitz operator with symbol* φ, and usually denoted T_φ. Thus

(5.11) $$T_\varphi g = PM_\varphi g = P(\varphi g) \qquad , \qquad \varphi \in L^\infty(\mathbb{T}), g \in H^2(\mathbb{T})$$

where P denotes orthogonal projection of L^2 on H^2. It is easy to check the properties

(5.12) $$\|T_\varphi\| \leqslant \|\varphi\|_\infty \qquad \text{(with equality if } \varphi \in H^\infty).$$

(5.13) $$T_\varphi^* = T_{\bar\varphi}.$$

See, e.g. [Do], [Ni] for extensive discussion of this important class of operators. Now, another important property of Toeplitz operators *with conjugate-analytic symbol* is this: let, for $\zeta \in \mathbb{D}$,

(5.14) $$k_\zeta(z) = \frac{1}{1-\bar\zeta z}.$$

This is the "representing element" for the functional $f \mapsto f(\zeta)$ on $H^2(\mathbb{D})$, i.e.

(5.15) $$\langle f, k_\zeta\rangle = f(\zeta) \qquad , \qquad f \in H^2(\mathbb{D}).$$

Crucial for us is the identity

(5.16) $$T_{\bar\varphi} k_\zeta = \overline{\varphi(\zeta)} k_\zeta \qquad , \qquad \varphi \in H^\infty.$$

The proof is easy, to verify (5.16) it suffices to show, for each $f \in H^2$, that both sides have the same inner product with f. Now,

$$\langle T_{\bar\varphi} k_\zeta, f\rangle = \langle k_\zeta, T_\varphi f\rangle = \langle k_\zeta, \varphi f\rangle = \overline{\varphi(\zeta)}\ \overline{f(\zeta)} = \overline{\varphi(\zeta)}, \langle k_\zeta, f\rangle,$$

proving (5.16). Hence, Sarason's reformulation of (PN) is:

Given are distinct points $z_1, \dots, z_n$ in $\mathbb{D}$ and complex numbers

$$w_1, \dots, w_n$$

Find necessary and sufficient conditions that there exist a linear operator T on H^2 such that

(5.17) $$\|T\| \leqslant 1$$

(5.18) T commutes with the shift M_z

and

(5.19) $$T^* k_{z_j} = \bar w_j k_{z_j} \qquad (j = 1, 2, \dots, n).$$

In view of our preceding discussion, it is clear that this problem is completely equivalent to (PN). Actually, it is slightly more convenient to replace T by its adjoint, and then (denoting T^* by S) the above conditions become

$$\|S\| \leqslant 1$$

S commutes with the *backward shift* $M_z^* = T_{\bar z}$ on H^2

$$S k_{z_j} = \bar w_j k_{z_j} \qquad (j = 1, 2, \dots, n).$$

Now, let N denote the span of the vectors k_{z_j} $(j = 1, 2, \dots, n)$. If S exists satisfying 5, 5 and 5 then $SN \subset N$ so the norm of $S \mid N$ is at most one, that is

(5.20) $$\left\| \sum_{j=1}^n \bar w_j t_j k_{z_j} \right\|^2 \leqslant \left\| \sum_{j=1}^n t_j k_{z_j} \right\|^2$$

holds for all complex n-tuples $\{t_j\}$. Expanding the squares and using the identity $\langle k_{z_i}, k_{z_j}\rangle = k_{z_i}(z_j) = 1/(1 - \bar{z}_i z_j)$, and a little manipulation, one sees that (5.20) is equivalent to the non-negative definite character of the matrix (5.2). This already shows the *necessity* of that condition for (PN), and now we turn to the more difficult issue of *sufficiency*. So, suppose (5.2) is non-negative definite, which is equivalent to (5.20), i.e. to the assertion

(5.21) $S \mid N$ has norm at most one.

Thus, $S \mid N$ *is a contractive operator on* N, *and it commutes with the restriction to* N *of the backward shift* $T_{\bar{z}}$ (indeed, each operator admits all the k_{z_i} as eigenvectors).

To finish the proof it suffices to show:

(E) $S \mid N$ has an extension to H^2 that is a contraction, and commutes with $T_{\bar{z}}$.

Indeed, if (E) is proven, and the extension is denoted $\widetilde{S}$, then $\widetilde{S}$ commutes with $T_{\bar{z}}$ so $(\widetilde{S})^*$ commutes with $(T_{\bar{z}})^* = M_z$ and hence is of the form T_f for some f in H^∞ by Lemma 5.2.

Since T_f (which is simply multiplication by f) is a contraction, $\|f\|_\infty \leqslant 1$. Thus, $\widetilde{S} = T_{\bar{f}}$ and so (using 5):

$$\bar{w}_j k_j = \widetilde{S} k_j = T_{\bar{f}} k_j = \overline{f(z_j)} k_j$$

so $f(z_j) = w_j$ $(j = 1, 2, \ldots, n)$, and this completes the proof.

Sarason was aware that there was more at stake here than just another solution to some classical interpolation problems. To grasp what this is, write H (a general Hilbert space) in place of H^2, and consider a subspace N that is invariant for an operator A on H (in PN the role of A is played by $T_{\bar{z}}$). Let now $S \in L(N)$ commute with $A \mid N$. And we ask:

(EG) *Under these conditions, can* S *be extended to all of* H *as a linear operator with the same norm, which moreover commutes with* A?

That is, at bottom the fundament of Pick's theorem is something very general, and geometric: an extension property for linear operators from a subspace of a Hilbert space to the whole space with preservation of (i) the norm and (ii) a commutation relation. Now examples show that this is not always possible. But, after the appearance of Sarason's paper, Foias and Sz.-Nagy, using dilation theory, succeeded to show that (EG) *has an affirmative answer whenever* A *is a co-isometry on* H, that is, A^* is an isometry. This covers the case discussed by Sarason, since there A^* is the shift. (Sarason was able to prove (E) by *ad hoc* methods; the Foias-Sz.-Nagy result is a new milestone, and opened up a fast-developing branch of operator theory.) The assertion that the extension asked for by (EG) exists when A is a co-isometry is sometimes called the "commutant lifting theorem" (CLT) because, in terms of the adjoint operators, it becomes a question of *lifting* (rather than extending) an operator commuting with an *isometry*. This in turn has a further generalization to the *intertwining lifting theorem*. Also, the CLT turns out to be intimately connected to the possibility of simultaneous unitary dilation of two commuting operators (Ando's theorem) and von-Neumann inequalities for polynomial functions of two commuting contractions. For all this, see [FoFr], [SzFo] and [Sz].

For all historical background concerning (PN) and related problems, an excellent source is [FrKi], which includes reprints of the fundamental papers of Herglotz, Schur, Pick and Nevanlinna, as well as a masterful historical review by B. Fritzsche and B. Kirstein and a thorough bibliography. The monograph [DuFrKi] is devoted to matricial generalizations. Other valuable references will be given in our Appendix.

A1. **Acknowledgements**. It is a pleasure to thank Bernd Kirstein, Karim Kellay, and Mihai Putinar for valuable advice concerning the selection of material, and the literature. I regret that lack of space has made it unfeasible to take up some very important topics that relate to the current material, especially Hankel and Toeplitz operators. (See [Sa2], [Po1], [Ni] and [BoeKa] for this material, and references to other relevant literature). I also express my gratitude to Siv Sandvik, who produced this fine LaTeXmanuscript from my hieroglyphics, under time pressure.

A.2. **Suggested reading**. The reader who wishes to learn more about the material we have discussed is referred to the following literature, which supplements the sources we already have named.

Stationary random sequences, prediction theory: [Lam,Ya,HeSz,HeSa,Sa3], [Ho, Chapter 4].

Operator-theoretic approaches to interpolation problems. Here we have barely scratched the surface. One alternative approach, building on the theory of *Krein spaces* (a generalization of Hilbert spaces, where the inner product is given by an *indefinite* Hermitian form) to (PN) and related problems is due to Ball and Helton. A very nice introduction to this theory is Sarason's lectures in the volume [Po2]. Also recommended from the same volume is N.K. Nikolskii's lectures on Hankel and Toeplitz operators, where the pace is more leisurely than in the comprehensive, standard monograph [Ni]. See also N.J. Young's article in [Po2]. The monograph [Po1] is also recommended, and [Sa2].

An important topic we have not mentioned is the class of Hilbert spaces of entire functions due to de Branges [Br], which are becoming increasingly popular as new applications and connections are discovered —for instance, at the time of this writing to "frames" and sampling theory by J. Ortega-Cerda and K. Seip.

The collection [Pe] is a gold mine. It contains an excellent introduction by Sarason to invariant subspaces, as well as A.L. Shields' survey of *weighted* shift operators, and R.G. Douglas' survey article on *canonical models* for operators, giving the "state of the art" of dilation theory and its ramifications as of 1970. The conference volume [Lan] contains an article of Sarason wherein he describes, with his usual clarity, the relation of *moment theorems* to operator theory in Hilbert space. On this score see also the monograph [RoRo] which gives a compact and unified framework for the application of Hilbert space methods (especially commutant lifting) to problems of interpolation both in $\mathbb{D}$ and $\overline{\mathbb{D}}$, moment theorems, Loewner's theorem on monotone matrix functions, and more. For interpolation of vector and matrix valued functions, see [DuFrKi], [FoFr] and [BaGoRo].

Fairly recently, theorems of Pick-Nevanlinna type in two or more variables have started to appear, a pioneering role having been played by Jim Agler. See [AgMc] for an account up to 1997. A spectacular application of operator theory to "hyperbolic geometry" was given in 1990 by Agler [Ag], who proved, using dilation techniques, the celebrated theorem of L. Lempert, that the Carathéodory and Kobayashi metrics agree on bounded convex domains. (To find out what those terms mean, the reader may consult the nice introductory book [Kr]).

A lively interest in generalizations of Beurling's theorem as well as of (PN), often in combination, and using tools from dilation theory, persists up to the present day. See, for example [Qu], [McTr] and [GrRiSu]. This work has been catalyzed in part also by new vistas and rich structure encountered in the recent work on Bergmann spaces, starting with the ground breaking paper of Hedenmalm [He]. We refer the reader to a forthcoming monograph by Hedenmalm *et al*. The current work [AlRiSu2], with a wealth of interesting results for classical analysts as well as a good bibliography, is also recommended.

Applications of Toeplitz matrices and operators in science and engineering are many. An interesting variant of the Carathéodory-Fejér problem important for engineering applications due to T. Georgiou is [Ge] and was followed up in later work, see [ByLi] for a recent account. Those are also beautiful applications to statistical physics, a good survey article is [Boe].

References

[Ag] Agler, Jim, *Operator theory and the Carathéodory metric*, Invent. Math. **101** (1990) 483–500.

[AgMc] Agler, Jim and John E. McCarthy, *Complete Nevanlinna-Pick kernels*, preprint, 1997.

[Ak] Akhieser, N.I., *The Classical Moment Theorem*, Hafner, New York, 1965.

[AkGl] Akhieser, N.I. and I.M. Glazman, *Theory of Linear Operators in Hilbert Space*, Ungar, New York, 1963.

[AlRiSu] Aleman, A., S. Richter and C. Sundberg, *Beurlings theorem for the Bergman space*, Acta Math. **177** (1996) 275–310.

[AlRiSu2] Aleman, A., S. Richter and C. Sundberg, *The majorization function, and the index of invariant subspaces in the Bergmann spaces*, preprint, June 2000.

[BaGoRo] Ball, J., F. Gohberg and L. Rodman, *Interpolation of Rational Matrix Functions*, Birkhäuser, 1990.

[Be] Beurling, A., *On two problems concerning linear transformations in Hilbert space*, Acta Math. **81** (1949) 239–255.

[BeChRe] Berg, Christian, J. Christenssen and P. Ressel, *Harmonic Analysis on Semigroups*, Springer-Verlag, 1984.

[ByLi] Byrnes, C. and A. Lindquist, *On the partial stochastic realization problem*, IEEE Trans. Automatic Control, Vol. 42, No. 8, Aug. 1997, pp. 1049–1070.

[Boe] Böttcher, A., *The Onsager formula, the Fisher-Hartwig conjecture, and their influence on research into Toeplitz operators*, J. Statistical Phys. **78** (1995) 575–584.

[BoeKa] Böttcher, A. and Yu. Karlovich, *Carleson Curves, Muckenhoupt Weights, and Toeplitz Operators*, Birkhäuser, 1997.

[Br] de Branges, L., *Hilbert Spaces of Entire Functions*, Prentice-Hall, 1968.

[Do] Douglas, R., *Banach Algebra Techniques in Operator Theory*, Academic Press, 1972.

[DuFrKi] Dubovoj, V., B. Fritzsche and B. Kirstein, *Matricial Version of the Classical Schur Problem*, B.G. Teubner Verlag, 1992.

[DuSc] Dunford, N. and J. Schwartz, *Linear Operators, Part 2*, Wiley-Interscience 1958.

[Du] Duren, P.L., *Theory of H^p Spaces*, Academic Press, 1970.

[FoFr] Foias, C. and A. Frazho, *The Commutant lifting Approach to Interpolation Problems*, Birkhäuser, 1990.

[FrKi] Fritzsche, B. and B. Kirstein, ed., *Ausgewählte Arbeiten zu den Ursprüngen der Schur-Analysis*, Teubner-Archiv zur Mathematik, Band 16, B.G. Teubner Verlag, 1991.

[Ga] Garnett, J., *Bounded Analytic Functions*, Academic Press, 1981.

[Ge] Georgiou, J.T., *Partial realization of covariance sequences*, Ph.D. dissertation, Univ. Florida, Gainesville, 1983.

[GoGoKa] Gohberg, I., S. Goldberg and M. Kaashoek, *Classes of Linear Operators*, 2 vol., Birkhäuser, 1990.

[GrRiSu] Greene, D., S. Richter and C. Sundberg, *The structure of inner multipliers on spaces with complete Nevanlinna-Pick kernels*, preprint, June 2000.

[GrSz] Grenander, U. and G. Szegö, *Toeplitz Forms and Their Applications*, Calif. Monographs in Math. Sci. **2**, 1958.

[Ha1] Halmos, P., *Introduction to Hilbert Space and the Theory of Spectral Multiplicity*, Chelsea, New York, 1951.

[Ha2] Halmos, P., *A Hilbert Space Problem Book*, 2 ed., Springer-Verlag, 1982.

[Ha3] Halmos, P., *Shifts on Hilbert Spaces*, J. Reine Angew. Math. **208** (1961) 102–112.

[He] Hedenmalm, H., *A factorization theorem for square area integrable analytic functions*, J. Reine Angew. Math. **422** (1991) 45–68.

[HeSa] Helson, H. and D. Sarason, *Past and future*, Math. Scand. **21** (1967) 5–16.

[HeSz] Helson, H. and G. Szegö, *A problem of prediction theory*, Ann. Mat. Pura Appl. **51** (1960) 107–138.

[Ho] Hoffman, K., *Banach Spaces of Analytic Functions*, Prentice-Hall, 1962.

[IbRo] Ibragimov, I. and Yu. Rozanov, *Gaussian Stochastic Processes*, Springer-Verlag 1978.

[Kr] Krantz, S., *Complex Analysis—The Geometric Viewpoint*, Carus Math. Monographs **23**, Math. Assoc. of America, 1990.

[Lam] Lamperti, *Stochastic Processes*, Springer-Verlag, 1977.

[Lan] Landau, H., ed., *Moments in Mathematics*, Proc. Symposia Appl. Math. **37**, AMS, 1987.

[Ma] Maurin, K., *Methods of Hilbert Spaces*, Monografie Mat., Tom 45, PWN Publishers, Warsaw, 1967.

[McTr] McCullough, S. and T. Trent, *invariant subspaces and Nevanlinna-Pick kernels*, preprint.

[Ne] Nevanlinna, R., *Über beschränkte analytische Funktionen*, Ann. Acad. Sci. Fennicae **32.7** (1929) 1–75.

[Ni] Nikolskii, *Treatise on the Shift Operator*, Springer-Verlag, 1986.

[Pe] Pearcy, C., ed., *Topics in Operator Theory*, Math. Surveys, No. 13, AMS, 1974.

[Pi] Pick, G., Über die Beschränkungen analytischer Funktionen, welche durch vorgegebene Funktionswerte bewirkt werden, Math. Annalen **77** (1916) 7–23.

[Po1] Power, S., *Hankel Operators on Hilbert Space*, Research Notes in Math. **64**, Pitman, 1982.

[Po2] Power, S., *Operators and Function Theory*, proc. of NATO Adv. Study Inst., D. Reidel publishers, 1985.

[Qu] Quiggin, P., *For which reproducing kernels is Pick's theorem true?*, Int. Equ. Op. Theory **16** (1993) 244–246.

[Ri] Riesz, F., *Über ein Problem von Carathéodory*, J. Für Math. **146** (1916) 83–87.

[RiSz] Riesz, F. and B. Sz-Nagy, *Leçons d'Analyse Fonctionelle*, 3 ed., Akad. Kiado, Szeged, 1955.

[Ro] Rozanov, Yu., *Stationary Stochastic Processes*, Moscow, Fizmatgiz 1963 (Russian).

[RoRo] Rosenblum, M. and J. Rovnyak, *Hardy Classes and Operator Theory*, Oxford, 1985.

[Ru] Rudin, W., *Fourier Analysis on Groups*, Second printing, Interscience, 1967.
[Sa1] Sarason, D., *Generalized interpolation in H^∞*, Trans. AMS **127** (1967) 179–203.
[Sa2] Sarason, D., *Function Theory on the Unit Circle*, Notes, Virginia Polytech. Inst., 1978.
[Sa3] Sarason, D., *Addendum to "Past and future"*, Math. Scand. **30** (1972) 62–64.
[Sc] Schur, I., *Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind*, J. für Math. **147** (1917) 205–232 and ibid. **148** (1918) 122–145.
[Sh] Shimorin, S., *Wold-type decompositions and wandering subspaces for operators close to isometries*, preprint, Lund University, No. 1999:8.
[Sz] Sz.-Nagy, B., *Unitary Dilations of Hilbert Space Operators and Related Topics*, CBMS Regional Conf. Series in Math. **19**, AMS, 1974.
[SzFo] Sz.-Nagy, B. and C. Foias, *Analyse Harmonique des Opérateurs de l'Espace de Hilbert*, Akad. Kiado, Budapest, 1967.
[To] Toeplitz, O., *Über die Fouriersche Entwickelungen positiver Funktionen*, Rend. Palermo **32** (1911) 191–192.
[Wo] Wold, H., *A Study in the Analysis of Stationary Time Series*, Stockholm, 1938.
[Ya] Yaglom, A., *An Introduction to the Theory of Stationary Random Functions*, Prentice-Hall, 1962.
[Zy] Zygmund, A., *Trigonometric Series*, Cambridge, 1968.