Hankel and Toeplitz operators

Lecture notes by Roman Bessonov

version 03.03.2026

Contents

1 Hilbert inequality **3**

1.1 The inequality and its elementary proof 3

1.2 The Hilbert matrix 4

1.3 Hankel matrices 5

1.4 Nehari theorem. Another proof of Hilbert inequality 7

2 Schur interpolation problem **9**

2.1 Schur interpolation problem 9

2.2 Toeplitz matrices 9

2.3 Brown-Halmos and Sarason theorems 10

2.4 Solution of Schur problem 11

2.5 Some examples 13

3 Hardy spaces, I **15**

3.1 Definitions and basic theory 15

3.2 Riesz projectors. Hankel and Toeplitz operators 17

4 Proofs of Nehari, Brown-Halmos and Sarason theorems **21**

4.1 Warm-up: the operator norm of a multiplication operator 21

4.2 Proof of Nehari theorem 22

4.3 Proof of Brown-Halmos theorem 24

4.4 Proof of Sarason theorem 26

5 Shift operator **30**

5.1 Wiener theorem 30

5.2 Beurling theorem 31

5.3 Some examples 33

6 Hardy spaces, II **35**
6.1 Convergence test for Blaschke products 35
6.2 Boundary values of Blaschke products 36
6.3 Blaschke factors of functions in Hardy spaces 37
6.4 Logarithmic integrability for functions in the Hardy spaces 38
6.5 Inner-outer factorization 40

7 Hankel operators of finite rank **44**
7.1 Coinvariant subspaces 44
7.2 Kronecker theorem 47

8 Compact Hankel operators **49**
8.1 Sarason algebra 49
8.2 Proof of Hartman theorem 50

9 Asymptotic behavior of Toeplitz determinants **53**
9.1 Szegő theorem on Toeplitz determinants 53
9.2 Orthogonal polynomials 54
9.3 Schur algorithm 58
9.4 Proof of Szegő theorem 61

10 Homeworks **64**
10.1 Homework I 64
10.2 Homework II 66

Credits **68**
Credits 68
Licence 68
Last version & source 68

1 Hilbert inequality

Our tour to the land of Hankel and Toeplitz operators starts with Hilbert inequality. Its proof by means of Hankel operators dates back to the very beginning of functional analysis. Another historical example, solution of Schur interpolation problem by means of Toeplitz operators, will be discussed in the next section.

1.1 The inequality and its elementary proof

Theorem 1.1 (Hilbert inequality). *For any sequences $\{a_n\}$, $\{b_n\}$ of complex numbers, we have*

$$\sum_{m,\,n\geq 1} \frac{|a_m|\cdot|b_n|}{m+n} \leq \pi \sqrt{\sum_{n\geq 1}|a_n|^2} \cdot \sqrt{\sum_{n\geq 1}|b_n|^2}.$$

Proof. We may assume that both sequences $\{a_n\}$, $\{b_n\}$ are nonnegative and square summable. For a positive real x, denote by $\lceil x \rceil$ the smallest integer that is larger than x. Define functions f, g on $(0, +\infty)$ by $f(x) = a_{\lceil x \rceil}$, $g(x) = b_{\lceil x \rceil}$, $x > 0$. We have

$$\begin{aligned}\sum_{m,n\geq 1} \frac{a_m b_n}{m+n} &\leq \int_0^\infty \int_0^\infty \frac{f(x)g(y)}{x+y}\, dx\, dy, \\ &= \int_0^{\frac{\pi}{2}} \int_0^\infty \frac{f(r\cos^2\varphi)g(r\sin^2\varphi)}{r\cos^2\varphi + r\sin^2\varphi} |J_{x,y}(r,\varphi)|\, dr\, d\varphi,\end{aligned}$$

by the change of variables

$$\begin{cases} x = r\cos^2\varphi, \\ y = r\sin^2\varphi. \end{cases}$$

Here, the Jacobian $J_{x,y}(r,\varphi)$ is given by

$$J_{x,y}(r,\varphi) = \det \begin{bmatrix} x'_r & x'_\varphi \\ y'_r & y'_\varphi \end{bmatrix} = \det \begin{bmatrix} \cos^2\varphi & -2r\cos\varphi\sin\varphi \\ \sin^2\varphi & 2r\sin\varphi\cos\varphi \end{bmatrix} = r\sin 2\varphi.$$

It follows that

$$\sum_{m,n\geq 1} \frac{a_m b_n}{m+n} \leq \int_0^{\frac{\pi}{2}} \int_0^\infty \frac{f(r\cos^2\varphi)g(r\sin^2\varphi)}{r} \cdot r\sin 2\varphi\, dr\, d\varphi.$$

By Cauchy-Schwarz inequality,

$$\int_0^\infty f(r\cos^2\varphi)g(r\sin^2\varphi)\,dr \le \left(\int_0^\infty f^2(r\cos^2\varphi)\,dr \cdot \int_0^\infty g^2(r\sin^2\varphi)\,dr\right)^{\frac{1}{2}}$$
$$= \frac{\|f\|_{L^2(0,\infty)}\|g\|_{L^2(0,\infty)}}{|\cos\varphi\sin\varphi|}.$$

Therefore,

$$\sum_{m,n\ge1}\frac{a_m b_n}{m+n} \le \int_0^{\frac{\pi}{2}} 2\|f\|_{L^2(0,\infty)}\|g\|_{L^2(0,\infty)}\,d\varphi = \pi\|f\|_{L^2(0,\infty)}\|g\|_{L^2(0,\infty)}.$$

It remains to note that

$$\|f\|_{L^2(0,\infty)} = \sqrt{\sum_{n\ge1}|a_n|^2}, \qquad \|g\|_{L^2(0,\infty)} = \sqrt{\sum_{n\ge1}|b_n|^2},$$

by the definition of functions f, g. □

1.2 The Hilbert matrix

Since the right hand side of Hilbert inequality contains a product of norms in the Hilbert space $\ell^2(\mathbb{N},\mathbb{C})$ of complex square-summable sequences indexed by natural numbers $\mathbb{N}$, one may wonder whether one can reformulate Hilbert inequality in a way that makes it a statement from functional analysis. This is indeed possible.

Definition 1.2. The **Hilbert matrix** is defined by

$$H = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

We also will deal with the following submatrix of the Hilbert matrix,

$$H_0 = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \cdots \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Both H and H_0 can be considered as linear operators on the space $\ell^2(\mathbb{N},\mathbb{C})$.

Lemma 1.3. *Hilbert inequality is equivalent to the bound $\|H_0\| \le \pi$ for the operator norm of H_0 as an operator on $\ell^2(\mathbb{N}, \mathbb{C})$.*

Proof. Let us first note that Hilbert inequality is equivalent to the inequality

$$\left| \sum_{m,n\ge 1} \frac{a_m \overline{b_n}}{m+n} \right| \le \pi$$

for all sequences $a = \{a_n\}$, $b = \{b_n\}$ with finitely many nonzero terms such that $\|a\|_{\ell^2(\mathbb{N},\mathbb{C})} \le 1$, $\|b\|_{\ell^2(\mathbb{N},\mathbb{C})} \le 1$. Denote the set of such sequences by $B_{\ell^2_0(\mathbb{N},\mathbb{C})}[0,1]$. For every $a, b \in B_{\ell^2_0(\mathbb{N},\mathbb{C})}[0,1]$, we have

$$|\langle H_0 a, b\rangle| = \left| \sum_{m,n\ge 1} \frac{a_m \overline{b_n}}{m+n} \right|,$$

with $\langle \cdot, \cdot \rangle$ denoting the inner product in $\ell^2(\mathbb{N}, \mathbb{C})$. Let us recall that the norm of a bounded linear operator on a Hilbert space $\mathcal{H}$ can be computed by the formula

$$\|T\|_{\mathcal{H}\to\mathcal{H}} = \sup_{x,y \in B_{\mathcal{H}_0}[0,1]} |\langle Tx, y\rangle_{\mathcal{H}}|,$$

where the supremum is taken over $B_{\mathcal{H}_0}[0,1]$ – the intersection of the unit ball in $\mathcal{H}$ with a dense linear subset $\mathcal{H}_0$ of $\mathcal{H}$. In particular, for $\mathcal{H} = \ell^2(\mathbb{N}, \mathbb{C})$, $B_{\mathcal{H}_0}[0,1] = B_{\ell^2_0(\mathbb{N},\mathbb{C})}[0,1]$, and $T = H_0$, we have

$$\|H_0\| = \sup_{a,b \in B_{\ell^2_0(\mathbb{N},\mathbb{C})}[0,1]} \left| \sum_{m,n\ge 1} \frac{a_m \overline{b_n}}{m+n} \right|.$$

From here we see that Hilbert inequality is equialent to the estimate $\|H_0\| \le \pi$. □

1.3 Hankel matrices

Hilbert matrix has a very specific structure – its elements are constant on anti-diagonals. In other words, the Hilbert matrix has Hankel structure. Let us denote $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$.

Definition 1.4. The Hankel matrix generated by a sequence $\{c_k\}_{k\in\mathbb{Z}_+}$ of complex numbers is the infinite matrix of the form

$$H_{\{c_k\}} = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots \\ c_1 & c_2 & c_3 & \cdots \\ c_2 & c_3 & c_4 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

In particular, the Hilbert matrix H corresponds to the choice $c_k = \frac{1}{k+1}$, $k \in \mathbb{Z}_+$, and its submatrix H_0 – to the choice $c_k = \frac{1}{k+2}$, $k \in \mathbb{Z}_+$. We can consider any Hankel matrix as a linear operator on the Hilbert space $\ell^2(\mathbb{Z}_+, \mathbb{C})$ densely defined by its bilinear form

$$\langle H_{\{c_k\}} a, b \rangle = \sum_{m,n \geq 0} c_{m+n} a_m \overline{b_n}$$

on the linear subset of sequences $a, b \in \ell_0^2(\mathbb{Z}_+, \mathbb{C})$ with finitely many non-zero elements. This operator might be bounded or unbounded depending on the choice of $\{c_k\}_{k \in \mathbb{Z}_+}$. Similarly to the Hilbert matrix, the boundedness is equivalent to the inequality

$$\|H_{\{c_k\}}\| = \sup_{a,b \in B_{\ell_0^2(\mathbb{Z}_+,\mathbb{C})}[0,1]} \left| \langle H_{\{c_k\}} a, b \rangle \right| < \infty.$$

We have the following simple necessary condition for the boundedness.

Lemma 1.5. *If $H_{\{c_k\}}$ extends to a bounded operator on $\ell^2(\mathbb{Z}_+, \mathbb{C})$, then $\{c_k\} \in \ell^2(\mathbb{Z}_+, \mathbb{C})$.*

Proof. Idea of the proof is very simple: if $H_{\{c_k\}}$ is bounded, then $\{c_k\} = H_{\{c_k\}} e_0$ must belong to the space $\ell^2(\mathbb{Z}_+, \mathbb{C})$. Since $H_{\{c_k\}}$ is in general defined only by its bilinear form, we should argue more accurately. Take an integer $N \in \mathbb{Z}_+$ and let $\{c_{N,k}\}$ be the sequence $\{c_0, c_1, \ldots, c_N, 0, \ldots\}$. Then

$$\sup_{b \in B_{\ell_0^2(\mathbb{Z}_+,\mathbb{C})}[0,1]} \left| \langle H_{\{c_k\}} e_0, b \rangle \right| \leq \|H_{\{c_k\}}\|$$

implies $\|\{c_{N,k}\}\|_{\ell^2(\mathbb{Z}_+,\mathbb{C})} \leq \|H_{\{c_k\}}\|$ for each N. It follows that $\{c_k\} \in \ell^2(\mathbb{Z}_+, \mathbb{C})$ and, moreover, $\|\{c_k\}\|_{\ell^2(\mathbb{Z}_+,\mathbb{C})} \leq \|H_{\{c_k\}}\|$. $\square$

Definition 1.6. A function $\varphi \in L^2[0, 2\pi]$ is called a **symbol** of the Hankel matrix $H_{\{c_k\}}$ if its Fourier decomposition with respect to the orthonormal basis $\{e^{ikt}\}_{k \in \mathbb{Z}}$ has the form

$$\varphi(t) = \sum_{k \in \mathbb{Z}} \hat{\varphi}(k) e^{ikt}, \qquad \hat{\varphi}(k) = c_{-k-1}, \quad k \geq 0.$$

Note that

$$\hat{\varphi}(k) = \frac{1}{2\pi} \int_0^{2\pi} \varphi(t) e^{-ikt} \, dt, \qquad k \in \mathbb{Z},$$

for every function $\varphi \in L^2[0, 2\pi]$. If exists, the symbol φ of $H_{\{c_k\}}$ is highly non-unique. Indeed, we limit only negative Fourier coefficients of φ. One can rephrase the definition of the symbol in the following way: a function $\varphi \in L^2[0, 2\pi]$ is a symbol of $H_{\{c_k\}}$ if and only if

$$\varphi(t) = \ldots + c_1 e^{-2it} + c_0 e^{-it} + d_0 + d_1 e^{it} + d_2 e^{2it} + \ldots$$

where the series converges in $L^2[0, 2\pi]$ (equivalently, the coefficients are square-summable). The left part of the series, corresponding to negative Fourier coefficients, is determined by the given sequence $\{c_k\} \in \ell^2(\mathbb{Z}_+, \mathbb{C})$, while the right part, corresponding to nonnegative Fourier coefficients, is arbitrary. Lemma 1.5 and Parseval identity imply that every bounded Hankel matrix $H_{\{c_k\}}$ has symbols in the space $L^2[0, 2\pi]$. Indeed, one can take, say,

$$\varphi_{st}(t) = \cdots + c_1 e^{-2it} + c_0 e^{-it}.$$

This symbol is often called the standard symbol of the Hilbert matrix $H_{\{c_k\}}$.

1.4 Nehari theorem. Another proof of Hilbert inequality

The most basic and commonly used fact about Hankel matrices is Nehari theorem.

Theorem 1.7 (Nehari theorem). *A Hankel matrix $H_{\{c_k\}}$ admits a bounded extension to the Hilbert space $\ell^2(\mathbb{Z}_+, \mathbb{C})$ if and only if it has a symbol $\varphi \in L^\infty[0, 2\pi]$. Moreover, we have*

$$\|H_{\{c_k\}}\| = \inf\{\|\varphi\|_{L^\infty[0,2\pi]} \mid \varphi \ \text{ is a symbol of } \ H_{\{c_k\}}\}.$$

Let us derive Hilbert inequality from Nehari theorem.

Proof of Hilbert inequality by means of Nehari theorem. By Lemma 1.3, we need to check that $\|H_0\| \leq \pi$ for the submatrix matrix H_0 of the Hilbert matrix H. Clearly, it suffices to prove that $\|H\| \leq \pi$. Since the matrix H has Hankel form, we can use Nehari theorem. So, now we turn to construct a symbol φ of H with $\|\varphi\|_{L^\infty[0,\pi]} \leq \pi$. We first guess a formula for φ using informal arguments, and then check that this formula indeed determines the required symbol. By definition, any symbol of the Hilbert matrix should be of the form

$$\ldots + c_k e^{-(k+1)it} + \ldots + c_1 e^{-2it} + c_0 e^{-it} + d_0 + d_1 e^{it} + \ldots$$

where a square-summable sequence $\{d_k\}$ is arbitrary. Let us consider first the standard symbol,

$$\varphi_{st} = \sum_{k=1}^{\infty} \frac{e^{-ikt}}{k} = -\log(1 - e^{-it}).$$

The last formula here needs a justification, but we skip it and instead ask themselves how to make a bounded function from $-\log(1+e^{-it})$ by changing Fourier coefficients with positive indexes. Since the absolute value of the imaginary part of $\log(1+e^{-it})$ is bounded by $\pi/2$ (it corresponds to the argument of $1+e^{-it}$, a point in the half-plane $\operatorname{Re} z > 0$) the natural choice will be

$$\varphi_{st} - \overline{\varphi_{st}} = \sum_{k=1}^{\infty} \frac{e^{-ikt}}{k} - \sum_{k=1}^{\infty} \frac{e^{ikt}}{k} = \log \frac{1-e^{it}}{1-e^{-it}} = \log(-e^{it}) = \log e^{i(t-\pi)} = i(t-\pi).$$

All computations here are informal, but the answer, $i(t-\pi)$, makes sense. Let us now check directly that $\varphi = i(t-\pi)$ is a symbol of Hilbert matrix. To verify this, we compute the Fourier coefficients:

$$\begin{aligned}
\hat{\varphi}(-k-1) =& \frac{1}{2\pi} \int_0^{2\pi} i(t-\pi) e^{-i(-k-1)t} dt \\
=& \frac{1}{2\pi} \int_0^{2\pi} ite^{i(k+1)t} dt - \frac{1}{2\pi} \int_0^{2\pi} i\pi e^{i(k+1)t} dt \\
=& \frac{1}{2\pi} \cdot \frac{t}{k+1} e^{i(k+1)t} \Big|_{t=0}^{2\pi} - \frac{1}{2\pi} \int_0^{2\pi} \frac{e^{i(k+1)t}}{k+1} dt - \frac{1}{2\pi} \int_0^{2\pi} i\pi e^{i(k+1)t} dt \\
=& \frac{1}{k+1} = c_k,
\end{aligned}$$

as desired. Finally, we have $\|H\| \leq \|\varphi\|_{L^\infty[0,2\pi]} = \pi$, and Hilbert inequality follows. $\square$

2 Schur interpolation problem

In this section we consider the classical Schur problem for analytic functions and solve it using some results for Toeplitz matrices.

2.1 Schur interpolation problem

Denote by H^∞ the Banach space of all bounded analytic functions f in the open unit disk $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ equipped with the norm

$$\|f\|_\infty = \sup\{|f(z)| : z \in \mathbb{D}\}.$$

Problem 2.1 (Schur problem)**.** Given $a_0, \dots, a_n \in \mathbb{C}$, decide if there is $f \in H^\infty$ such that

- $f = \sum_{k=0}^{\infty} b_k z^k$ with $b_k = a_k$ for $k = 0, \dots, n$,
- $\|f\|_\infty \leq 1$.

Surprisingly, the problem has "if and only if" solution in terms of coefficients $a_0, \dots, a_n$. The solution was found by I.Schur in 1917, his proof was elementary. We will give a proof of Schur theorem by means of Toeplitz operators.

2.2 Toeplitz matrices

Definition 2.2. The **Toeplitz matrix** generated by a sequence of complex numbers $\{c_n\}_{n\in\mathbb{Z}}$ is the matrix

$$T_{\{c_n\}} = \begin{bmatrix} c_0 & c_1 & c_2 & c_3 & \ddots \\ c_{-1} & c_0 & c_1 & c_2 & \ddots \\ c_{-2} & c_{-1} & c_0 & c_1 & \ddots \\ c_{-3} & c_{-2} & c_{-1} & c_0 & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

Define also the corresponding Toeplitz operator T by its bilinear form

$$\langle Tx, y \rangle = \sum_{k,n\in\mathbb{Z}_+} c_{k-n} x_k \overline{y_n}$$

on sequences $x, y \in \ell_0^2(\mathbb{Z}_+)$ with finitely many nonzero elements.

Lemma 2.3. *Let T be the Toeplitz operator generated by a sequence $\{c_n\}_{n\in\mathbb{Z}}$. If T extends to a bounded operator on the whole space $\ell^2(\mathbb{Z}_+)$, then $\{c_n\}_{n\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$.*

Proof. The argument is similar to the Hankel case. Intuitively, since T is bounded and $e_0 \in \ell^2(\mathbb{Z}_+)$, then the sequences $\{c_{-n}\}_{n\in\mathbb{Z}_+} = Te_0$ and $\{\overline{c_n}\}_{n\in\mathbb{Z}_+} = T^* e_0$ belong to $\ell^2(\mathbb{Z}_+)$. Taken together, $\{c_n\}_{n\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$. Details related to rigorously rewriting this argument in terms of bilinear forms are left for the reader. □

Definition 2.4. A function $\varphi \in L^2[0, 2\pi]$ is called a **symbol** of the Toeplitz matrix $T_{\{c_n\}}$ if its Fourier decomposition with respect to the orthonormal basis $\{e^{ikt}\}_{k\in\mathbb{Z}}$ has the form

$$\varphi(t) = \sum_{k\in\mathbb{Z}} \hat{\varphi}(k)e^{ikt}, \qquad \hat{\varphi}(k) = c_{-k}, \quad k \in \mathbb{Z}.$$

Recall that

$$\hat{\varphi}(k) = \frac{1}{2\pi} \int_0^{2\pi} \varphi(t)e^{-ikt}\, dt, \qquad k \in \mathbb{Z},$$

for every function $\varphi \in L^2[0, 2\pi]$. One can rephrase the definition of the symbol in the following way: a function $\varphi \in L^2[0, 2\pi]$ is a symbol of $T_{\{c_k\}}$ if and only if

$$\varphi(t) = \ldots + c_2e^{-2it} + c_1e^{-it} + c_0 + c_{-1}e^{it} + c_{-2}e^{2it} + \ldots$$

where the series converges in $L^2[0, 2\pi]$ (equivalently, the coefficients are square-summable). Lemma 1.5 and Parseval identity imply that every bounded Toeplitz matrix $T_{\{c_k\}}$ has the unique symbol $\varphi \in L^2[0, 2\pi]$. In other words, the correspondence between Toeplitz matrices and their symbols is bijective. Indeed, all Fourier coefficients of the symbol are uniquely determined by the entries of Toeplitz matrix and vice versa.

2.3 Brown-Halmos and Sarason theorems

The norm of Toeplitz matrix can be computed in terms of its symbol.

Theorem 2.5 (Brown-Halmos)**.** *Let $T_{\{c_n\}}$ be a Toeplitz matrix and let $\varphi = \sum_{k\in\mathbb{Z}} c_{-k}e^{ikt}$ be its symbol. Then*

$$\|T_{\{c_n\}}\|_{\mathcal{B}(\ell^2(\mathbb{Z}_+))} = \|\varphi\|_{L^\infty[0,2\pi]}.$$

Let us illustrate the usage of Brown-Halmos theorem in some simple situations.

Example 2.6. Identity matrix is the Toeplitz matrix $T_{\{c_n\}}$ generated by the sequence

$$\{c_n\} = \ldots, 0, 0, 1, 0, 0, \ldots, \qquad c_0 = 1.$$

The symbol of this matrix is the constant function with value 1. Clearly, the norm of the identity operator is 1. This agrees with the Brown-Halmos theorem:

$$\|T_{\{c_n\}}\| = \|1\|_{L^\infty[0,2\pi]} = 1.$$

Example 2.7. The shift operator $S : \{x_n\}_{n\in\mathbb{Z}_+} \mapsto \{x_{n-1}\}_{n\in\mathbb{Z}_+}$ on $\ell^2(\mathbb{Z}_+)$ (we set $x_{-1} = 0$) has Toeplitz structure of its matrix in the standard basis of $\ell^2(\mathbb{Z}_+)$. More precisely its matrix is the Toeplitz matrix $T_{\{c_n\}}$ generated by the sequence

$$\{c_n\} = \ldots, 0, 1, 0, 0, 0, \ldots, \qquad c_{-1} = 1.$$

The symbol of this Toeplitz matrix $T_{\{c_n\}}$ is $\varphi = e^{it}$. Note that S is an isometry on $\ell^2(\mathbb{Z}_+)$, i.e., $\|Sx\| = \|x\|$ for every $x \in \ell^2(\mathbb{Z}_+)$. In particular, we have $\|S\| = 1$. This also agrees with the Brown-Halmos theorem: $\|e^{it}\|_{L^\infty[0,2\pi]} = 1$.

Theorem 2.8 (Sarason)**.** *Let T_n be an upper triangular $n\times n$ matrix with constant diagonals. Then there exists an upper triangular Toeplitz matrix T such that*

1. *the left-upper submatrix of T of size $n \times n$ is T_n,*

2. $\|T\| = \|T_n\|$.

The same holds for lower triangular Toeplitz matrices.

Note that the first condition in Sarason theorem itself is not very restricting. For example, one can construct a Toeplitz matrix T satisfying (1.) as follows: extend the diagonals of T_n with the same values and fill the remaining entries with zeros. Moreover, it is clear that $\|T\| \geq \|T_n\|$ for any extension of T_n to an infinite Toeplitz matrix T. Hence, the existence of T satisfying the exact equality $\|T\| = \|T_n\|$ is the nontrivial part.

2.4 Solution of Schur problem

With a sequence $a_0, \ldots, a_n \in \mathbb{C}$ we associate the following Toeplitz matrix T_{n+1},

$$T_{n+1} = \begin{bmatrix} a_0 & 0 & 0 & \ldots & 0 \\ a_1 & a_0 & 0 & \ldots & 0 \\ a_2 & a_1 & a_0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_n & a_{n-1} & a_{n-2} & \ldots & a_0 \end{bmatrix}$$

In terms of this matrix (we consider it as an operator on the Hilbert space $\mathbb{C}^{n+1}$) the solution of Schur problem can be formulated as follows.

Theorem 2.9 (Schur theorem)**.** *Schur problem with initial data $a_0, \dots, a_n$ admits a solution if and only if $I - T_{n+1}^* T_{n+1} \geq 0$.*

Our aim is to derive this result from Brown-Halmos and Sarason theorems. Then we will see on examples how to use in practice the condition $I - T_{n+1}^* T_{n+1} \geq 0$. Let us start with the following simple lemma from linear algebra.

Lemma 2.10. *For every finite matrix M, we have $\|M\| \leq 1$ if and only if $I - M^*M \geq 0$.*

Proof. We have

$$\begin{aligned} \|M\| \leq 1 &\iff \forall x: \ \|Mx\| \leq \|x\|, \\ &\iff \forall x: \ \|Mx\|^2 \leq \|x\|^2, \\ &\iff \forall x: \ 0 \leq \|x\|^2 - \|Mx\|^2 = \langle x, x\rangle - \langle Mx, Mx\rangle = \langle x, x\rangle - \langle x, M^*Mx\rangle, \\ &\iff \forall x: \ 0 \leq \langle x, (I - M^*M)x\rangle. \end{aligned}$$

But the last line is just the definition of $I - M^*M \geq 0$. □

The maximum modulus principle for analytic functions imply that a nonconstant analytic function in the open unit disk $\mathbb{D}$, continuous up to the boundary $\mathbb{T} = \{z \in \mathbb{C}: \ |z| = 1\}$, attains its maximum of modulus on the unit circle $\mathbb{T}$. There is a version of this result that does not require continuity assumption.

Lemma 2.11. *Let $f \in H^\infty$, and let $f = \sum_{k=0}^\infty c_k z^k$ be its Tailor decomposition in the open unit disk. Then $\{c_k\} \in \ell^2(\mathbb{Z}_+)$, and the series $\sum_{k=0}^\infty c_k e^{ikt}$ converges in $L^2[0, 2\pi]$ to some function $\tilde{f} = \sum_{k=0}^\infty c_k e^{ikt}$. Morever, we have $\tilde{f} \in L^\infty[0, 2\pi]$. Conversely, every function $\tilde{f} \in L^\infty[0, 2\pi]$ whose Fourier decomposition has the form $\tilde{f} = \sum_{k=0}^\infty c_k e^{ikt}$ determines the function $f \in H^\infty$ by $f = \sum_{k=0}^\infty c_k z^k$. This correspondence $f \to \tilde{f}$ is one to one, and, moreover, we have*

$$\|f\|_{H^\infty} = \|\tilde{f}\|_{L^\infty[0,2\pi]}.$$

For a while, we take this lemma for granted. In what follows, we will meet it in the general theory of Hardy spaces.

Proof of Schur theorem. By Lemma 2.11, Schur problem with initial data $a_0, a_1, \dots, a_n$ has a solution if and only if there exists a function $\varphi \in L^\infty[0, 2\pi]$ that satisfies the following two properties:

1. $\hat{\varphi}(k) = a_k$ for $k = 0, \dots, n$,
2. $\|\varphi\|_{L^\infty[0,2\pi]} \leq 1$.

In terms of Toeplitz operators this can be reformulated as follows: there exists a function $\varphi \in L^\infty[0, 2\pi]$ such that

1. T_{n+1} is the left-upper submatrix of a Toeplitz matrix T with symbol φ,

2. $\|\varphi\|_{L^\infty[0,2\pi]} \le 1$.

Brown-Halmos theorem allows us to further reformulate this in the form

1. T_{n+1} is the left-upper submatrix of a Toeplitz matrix T with symbol φ,

2. $\|T\| \le 1$,

and Sarason theorem eventually implies that this is equivalent to the bound $\|T_{n+1}\| \le 1$. Finally, Lemma 2.10 completes the proof. □

2.5 Some examples

The condition $I - T_{n+1}^* T_{n+1} \ge 0$ from Schur theorem can be checked by the Sylvester criterion, which we recall below.

Let M be an $n \times n$ Hermitian matrix. We are interested whether it is positive definite (positive semi-definite), that is, for all $x \ne 0$ holds $\langle Mx, x\rangle_{\mathbb{C}^n} > 0$ (≥ 0), in which case we write $M > 0$ ($M \ge 0$). Sylvester criterion gives a list of necessary and sufficient conditions to answer this question.

Definition 2.12. Let us write $M = (a_{i,j})_{1\le i,j\le n}$ for the entries of M. Let $M_k := (a_{i,j})_{1\le i,j\le k}$ be the unique upper left $k\times k$ submatrix of M. We call its determinant $\mu_k = \det M_k$ the k-the leading principal minor of M. More generally, given $I \subset \{1, \dots, n\}$, we define the $|I| \times |I|$ submatrix $M_I = (a_{i,j})_{i,j\in I}$ and denote the corresponding determinant by $\mu_I := \det M_{(I,J)}$. Note that $M_k = M_I$ for $I = \{1, \dots, k\}$.

Theorem 2.13 (Sylvester criterion for positive definiteness)**.** *We have $M > 0$ if and only if $\mu_k > 0$ for every $k = 1, \dots, n$.*

Theorem 2.14 (Sylvester criterion for positive semi-definiteness)**.** *We have $M \ge 0$ if and only if $\mu_I \ge 0$ for every $I \subset \{1, \dots, n\}$.*

Next, we illustrate the usage of Schur theorem in the cases $n = 0$ and $n = 1$.

Example 2.15. For $n = 0$, we have $T_1 = (a_0)$ and hence condition $I - T_{n+1}^* T_{n+1} \ge 0$ takes the form $(1 - |a_0|^2) \ge 0$ which is equivalent to $|a_0| \le 1$.

Example 2.16. For $n = 1$, we have

$$T_2 = \begin{pmatrix} a_0 & 0 \\ a_1 & a_0 \end{pmatrix}$$

and hence $I - T_{n+1}^* T_{n+1} \geq 0$ rewrites in the form

$$I - T_2^* T_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \bar{a}_0 & \bar{a}_1 \\ 0 & \bar{a}_0 \end{pmatrix} \begin{pmatrix} a_0 & 0 \\ a_1 & a_0 \end{pmatrix} = \begin{pmatrix} 1 - |a_0|^2 - |a_1|^2 & -\bar{a}_1 a_0 \\ -a_1 \bar{a}_0 & 1 - |a_0|^2 \end{pmatrix} \geq 0.$$

The minors are $\mu_{\{1\}} = 1 - |a_0|^2 - |a_1|^2$, $\mu_{\{2\}} = 1 - |a_0|^2$ and $\mu_{\{1,2\}} = \det(I - T_2^* T_2) = (1 - |a_0|^2)(1 - |a_0|^2 - |a_1|^2) - |a_0|^2 |a_1|^2 = (1 - |a_0|^2)^2 - |a_1|^2$, which have to be non-negative by Sylvester criterion. These conditions are equivalent to the bound

$$|a_1| + |a_0|^2 \leq 1$$

which is the solution of Schur problem for $n = 1$.

Note that the limit of positive definite matrices is positive semi-definite, hence we can solve $I - T_2^* T_2 \geq 0$ by checking for which a_i we have $I - T_1^* T_2 > 0$. In that case we only need to consider the principal minors $\mu_1 = \mu_{\{1\}}$, $\mu_2 = \mu_{\{1,2\}}$, which reduces to $|a_1| + |a_2|^2 < 1$. In the limit we obtain that a solution exists when $|a_1| + |a_0|^2 \leq 1$. For large n the same idea allows to consider $n + 1$ inequalities $\mu_k > 0$, $k = 1, \ldots, n + 1$ instead of much larger set of inequalities $\mu_I \geq 0$, $I \subset \{1, \ldots, n\}$.

3 Hardy spaces, I

3.1 Definitions and basic theory

Let $L^p(\mathbb{T})$, $1 \le p \le \infty$, denote the standard Lebesgue space on the unit circle $\mathbb{T} = \{z : |z| = 1\}$ with respect to the Lebesgue measure m on $\mathbb{T}$ normalized by $m(\mathbb{T}) = 1$. For finite p, the norm in $L^p(\mathbb{T})$ is given by

$$\|f\|_p = \left(\int_{\mathbb{T}} |f(\xi)|^p \, dm \right)^{\frac{1}{p}}.$$

Recall that the integration with respect to m is given by

$$\int_{\mathbb{T}} g(\xi) \, dm = \frac{1}{2\pi} \int_0^{2\pi} g(e^{it}) \, dt, \qquad g \in L^p(\mathbb{T}).$$

Using Cauchy-Schwarz inequality, it is not difficult to check that $L^{p_1}(\mathbb{T}) \supset L^{p_2}(\mathbb{T})$ if $p_1 \le p_2$. The space $L^\infty(\mathbb{T})$ consists of all Borel functions $f : \mathbb{T} \to \mathbb{C}$ whose absolute value is bounded by a constant almost everywhere on $\mathbb{T}$, the smallest such a constant is called the essential supremum of f. This defines the norm in $L^\infty(\mathbb{T})$,

$$\|f\|_\infty = \operatorname*{ess\,sup}_{\xi \in \mathbb{D}} |f(\xi)|.$$

The Fourier coefficients of a function $f \in L^1(\mathbb{T})$ will be denoted by $\widehat{f}(k)$, so that

$$\hat{f}(k) = \langle f, z^k \rangle_{L^2(\mathbb{T})} = \int_{\mathbb{T}} f(\xi) \overline{\xi}^k \, dm = \frac{1}{2\pi} \int_0^{2\pi} f(e^{it}) e^{-ikt} \, dt.$$

Definition 3.1 (Hardy spaces on the unit circle)**.** Let $1 \le p \le \infty$. The Hardy space $H^p(\mathbb{T})$ is the subspace in $L^p(\mathbb{T})$ formed by functions for which $\widehat{f}(k) = 0$ for all $k < 0$. In other words,

$$H^p(\mathbb{T}) = \{f \in L^p(\mathbb{T}) : \widehat{f}(k) = 0, \ k < 0\}.$$

The norm in $H^p(\mathbb{T})$ coincides with the norm in $L^p(\mathbb{T})$.

Example 3.2. In the special case $p = 2$, the space $H^p(\mathbb{T})$ is Hilbert. It admits a simple description in terms of power series. Indeed, from basic Fourier analysis we know that

$$L^2(\mathbb{T}) = \left\{ \sum_{n \in \mathbb{Z}} a_n z^n : \sum_{n \in \mathbb{Z}} |a_n|^2 < \infty \right\},$$

where the series converge in the norm of $L^2(\mathbb{T})$. Therefore,

$$H^2(\mathbb{T}) = \left\{ \sum_{n\in\mathbb{Z}_+} a_n z^n : \sum_{n\in\mathbb{Z}_+} |a_n|^2 < \infty \right\},$$

with the convergence again with respect to the norm of $L^2(\mathbb{T})$.

One of key initial observations related to Hardy spaces is that functions from $H^2(\mathbb{T})$ admit analytic continuation to the open unit disk $\mathbb{D}$.

Lemma 3.3. *For every $f \in H^2(\mathbb{T})$, the series $\sum_{k\in\mathbb{Z}_+} \widehat{f}(k) z^k$ converges in the open unit disk to an analytic function.*

Proof. Take $f \in H^2(\mathbb{T})$. Since $\sum_{k\in\mathbb{Z}_+} |\widehat{f}(k)|^2 < \infty$, we have $\sup_k |\widehat{f}(k)| < C < \infty$. Therefore,

$$\left| \sum_{k\in\mathbb{Z}_+} \widehat{f}(k) z^k \right| \le C \sum_{k\in\mathbb{Z}_+} |z|^k = \frac{C}{1-|z|}, \qquad |z| < 1.$$

Thus, the series $\sum_{n\in\mathbb{Z}_+} a_n z^n$ converges uniformly on compact subsets of $\mathbb{D}$ to an anlytic function. □

The previous observation indicates that functions in $H^p(\mathbb{T})$, initially defined almost everywhere on $\mathbb{T}$, might have an analytic description. This is indeed the case. Let us define the Hardy spaces $H^p(\mathbb{D})$ in the open unit disk $\mathbb{D}$.

Definition 3.4 (Hardy spaces in the unit disk)**.** Let $1 \le p < \infty$. Define the space

$$H^p(\mathbb{D}) = \{ f : \mathbb{D} \to \mathbb{C} : \ f \text{ analytic in } \mathbb{D},\ \|f\|_{H^p(\mathbb{D})} < \infty \},$$

whith the norm

$$\|f\|_{H^p(\mathbb{D})} = \sup_{0<r<1} \left(\int_{\mathbb{T}} |f(r\xi)|^p \, dm \right)^{1/p}.$$

For $p = +\infty$, define $H^\infty(\mathbb{D})$ to be the space of all bounded analytic functions in $\mathbb{D}$ with the norm $\|f\|_{H^\infty(\mathbb{D})} = \sup_{|z|<1} |f(z)|$.

The spaces $H^p(\mathbb{D})$ and $H^p(\mathbb{T})$ are essentially the same due to the following deep result.

Theorem 3.5. *Let $f \in H^p(\mathbb{D})$ for some $1 \le p \le \infty$. Then for Lebesgue almost every $\xi \in \mathbb{T}$ there exists the limit*

$$\widetilde{f}(\xi) := \lim_{r\to 1} f_r(\xi), \qquad f_r(\xi) := f(r\xi).$$

Moreover, we have $\widetilde{f} \in H^p(\mathbb{T})$. For every $1 \le p < \infty$, we have

$$\lim_{r \to 1} \|f_r - \widetilde{f}\|_{L^p(\mathbb{T})} = 0.$$

The operator $U : f \mapsto \widetilde{f}$ is the isomorphism (a linear isometric bijection) between Banach spaces $H^p(\mathbb{D})$ and $H^p(\mathbb{T})$. In particular, we have $\|Uf\|_{H^p(\mathbb{T})} = \|f\|_{H^p(\mathbb{D})}$. We also have

$$U\left(\sum_{k \in \mathbb{Z}_+} a_k z^k\right) = \sum_{k \in \mathbb{Z}_+} a_k \xi^k,$$

where the series in the left hand side is the Tailor series of a function $f \in H^p(\mathbb{D})$ on $\mathbb{D}$, and the series in the right hand side is the Fourier series of the function $\tilde{f} = Uf$ on $\mathbb{T}$.

In this course, we take Theorem 3.5 for granted. Based on this result, we will identify functions in $H^p(\mathbb{D})$ and $H^p(\mathbb{T})$. We will use the same letter f for the boundary value $\widetilde{f}$ of a function $f \in H^2(\mathbb{D})$. Since $H^p(\mathbb{D})$ and $H^p(\mathbb{T})$ are isomorphic, in many situations we will simply write H^p to denote the Hardy space.

Example 3.6. In the case $p = 2$ one can easily check identity $\|Uf\|_{H^p(\mathbb{T})} = \|f\|_{H^p(\mathbb{D})}$ from the statement of Theorem 3.5 provided that the rest of this theorem is known. By definition,

$$\|f\|_{H^2(\mathbb{D})} = \sup_{0<r<1} \|f_r\|_{L^2(\mathbb{T})}.$$

Consider the Taylor expansion of $f_r(z) = \sum_{n=0}^{\infty} a_n r^n z^n$. The harmonics $\{z^k\}_{k \in \mathbb{Z}_+}$ form the orthonormal basis in $L^2(\mathbb{T})$. It follows that

$$\|f_r\|^2_{L^2(\mathbb{T})} = \left\|\sum_{n=0}^{\infty} a_n r^n z^n\right\|^2_{L^2(\mathbb{T})} = \sum_{n=0}^{\infty} |a_n|^2 r^{2n}.$$

Thus, we have

$$\|f\|_{H^2(\mathbb{D})} = \sup_{0<1<r} \sum_{n=0}^{\infty} |a_n|^2 r^{2n} = \sum_{n=0}^{\infty} |a_n|^2 = \|f\|_{H^2(\mathbb{T})},$$

as claimed.

3.2 Riesz projectors. Hankel and Toeplitz operators

Definition 3.7 (Riesz projectors)**.** Let $f \in L^2(\mathbb{T})$, and let $f(z) = \sum_{n=-\infty}^{\infty} a_n z^n$ be its Fourier series with respect to the orthonormal system $\{z^n\}_{n \in \mathbb{Z}}$ in $L^2(\mathbb{T})$. We define two mappings

P_+, P_- by

$$P_+ f = \sum_{n\in\mathbb{Z}_+} a_n z^n, \qquad P_- f = \sum_{n\in\mathbb{Z}\setminus\mathbb{Z}_+} a_n z^n.$$

These mappings P_+, P_- are orthogonal projectors in the space $L^2(\mathbb{T})$. They are called Riesz projectors.

Observe that the image of P_+ is the space of those $L^2(\mathbb{T})$-functions that have vanishing Fourier coefficients $\widehat{f}(k) = 0$ for $k < 0$. In other words, we have $P_+ L^2(\mathbb{T}) = H^2(\mathbb{T})$. On the other hand, since $P_- f = \overline{z \sum_{n=0}^{\infty} a_{-n-1} z^n}$, we have $\operatorname{Ran} P_- = \overline{zH^2(\mathbb{T})}$, where the line denotes the complex conjugation. It is clear that $L^2(\mathbb{T})$ decomposes as

$$L^2(\mathbb{T}) = H^2(\mathbb{T}) \oplus \overline{zH^2(\mathbb{T})} = \operatorname{Ran} P_+ \oplus \operatorname{Ran} P_-.$$

It is also clear that $P_+ \oplus P_- = I$. Moreover, since $P_\pm$ are orthogonal projectors, we have $\|P_+\| = \|P_-\| = 1$.

Polynomials are the simplest bounded analytic functions that are dense in the space $H^2(\mathbb{T})$. For a function $f = \sum_{n=0}^{\infty} a_n z^n$ in $H^2(\mathbb{T})$, the polynomials $f_n = \sum_{k=0}^{k} a_k z^k$ converge to f in norm: $\lim_{n\to\infty} \|f - f_n\|_{L^2(\mathbb{T})} = 0$. Denote the set of polynomials by $\mathcal{P}$.

Definition 3.8 (Hankel operators)**.** The Hankel operator $H_\varphi : H^2(\mathbb{T}) \to \overline{zH^2(\mathbb{T})}$ with symbol $\varphi \in L^2(\mathbb{T})$ is densely defined by

$$H_\varphi : f \mapsto P_-(\varphi f), \qquad f \in \mathcal{P}.$$

Definition 3.9 (Toeplitz operators)**.** The Toeplitz operator $T_\varphi : H^2(\mathbb{T}) \to H^2(\mathbb{T})$ with symbol $\varphi \in L^2(\mathbb{T})$ is densely defined by

$$T_\varphi f \mapsto P_+(\varphi f), \qquad f \in \mathcal{P}.$$

In the above two definitions we used the fact that $\varphi f \in L^2(\mathbb{T})$ for every $\varphi \in L^2(\mathbb{T})$, $f \in \mathcal{P}$. In particular, the Riesz projectors $P_\pm(\varphi f)$ are defined correctly. Using bilinear forms, one can define Hankel and Toeplitz operators for more general symbols than $\varphi \in L^2(\mathbb{T})$ (say, for $\varphi \in L^1(\mathbb{T})$). However, it can be shown that such operators will not be bounded, so we skip their consideration here.

The following definition and the forthcoming results will be used to relate Hankel and Toeplitz matrices with Hankel and Toeplitz operators.

Definition 3.10 (Unitary operators)**.** Let $\mathcal{H}_1$, $\mathcal{H}_2$ be Hilbert spaces. A linear operator $U : \mathcal{H}_1 \to \mathcal{H}_2$ is called unitary if U is an isometric bijection.

The following two lemmas are elementary.

Lemma 3.11. *Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces with orthonormal bases $\{\vec{e}_k\}_{k\in\mathbb{Z}_+}$ and $\{\vec{v}_k\}_{k\in\mathbb{Z}_+}$ respectively. Then the operator*

$$U : \sum_{k=0}^{\infty} a_k \vec{e}_k \to \sum_{k=0}^{\infty} a_k \vec{v}_k$$

is unitary.

Lemma 3.12. *Let $U : \mathcal{H}_1 \to \mathcal{H}_2$ be a unitary operator. Then $U^* = U^{-1}$.*

Lemma 3.13. *If $T : \mathcal{H}_1 \to \mathcal{H}_2$, $\widetilde{T} : \widetilde{\mathcal{H}_1} \to \widetilde{\mathcal{H}_2}$ are densely defined linear operators, and there exist unitary operators U_1, U_2 such that $\widetilde{T} = U_1 T U_2$ on the domain of definition of $\widetilde{T}$, then T, $\widetilde{T}$ are bounded or not simultaneously and their operator norms coincide: $\|T\| = \|\widetilde{T}\|$.*

Theorem 3.14. *There exist unitary operators $U_1 : H^2(\mathbb{T}) \to \ell^2(\mathbb{Z}_+)$, $U_2 : \overline{zH^2(\mathbb{T})} \to \ell^2(\mathbb{Z}_+)$, such that for every Hankel operator $H_\varphi : H^2(\mathbb{T}) \to \overline{zH^2(\mathbb{T})}$ and for every Toeplitz operator $T_\varphi : H^2(\mathbb{T}) \to H^2(\mathbb{T})$ we have*

$$H_\varphi = U_2^{-1} H_{\{c_k\}} U_1, \qquad T_\varphi = U_1^{-1} T_{\{c_k\}} U_1,$$

for a suitable choice of sequence $\{c_k\}$. More precisely, in the Hankel case $c_k = \widehat{\varphi}(-k-1)$ for $k \in \mathbb{Z}_+$, while in the Toeplitz case we have $c_k = \widehat{\varphi}(-k)$ for $k \in \mathbb{Z}$.

Proof. The argument is based on the following key observation. Fix $\varphi \in L^2(\mathbb{T})$, $k, j \in \mathbb{Z}_+$. Consider the following inner product in $L^2(\mathbb{T})$:

$$\begin{aligned}
\langle H_\varphi z^k, \overline{z}^{j+1}\rangle &= \langle P_- \varphi z^k, \overline{z}^{j+1}\rangle \\
&= \langle \varphi z^k, P_-^* \overline{z}^{j+1}\rangle \quad \text{as } P_- \text{ is the orthogonal projector,} \\
&= \langle \varphi z^k, P_- \overline{z}^{j+1}\rangle \\
&= \langle \varphi z^k, \overline{z}^{j+1}\rangle \qquad \text{as there is no need to project,} \\
&= \int_{\mathbb{T}} \varphi z^{k+j+1}\, dm \\
&= \widehat{\varphi}(-k-j-1).
\end{aligned}$$

Similarly for Toeplitz operators,

$$\langle T_\varphi z^k, z^j\rangle = \widehat{\varphi}(j-k).$$

Next, observe that

$$\langle H_{\{c_k\}}\vec{e}_k, \vec{e}_j\rangle_{\ell^2} = c_{k+j}, \qquad \langle T_{\{c_k\}}\vec{e}_k, \vec{e}_j\rangle_{\ell^2} = c_{k-j}.$$

Thus if $c_k = \widehat{\varphi}(-k-1)$ for $k \in \mathbb{Z}_+$, then

$$\langle H_\varphi z^k, \overline{z}^{j+1}\rangle_{L^2(\mathbb{T})} = \langle H_{\{c_k\}}\vec{e}_k, \vec{e}_j\rangle_{\ell^2(\mathbb{Z}_+)}.$$

Similarly, if $c_k = \widehat{\varphi}(-k)$ for $k \in \mathbb{Z}$, then

$$\langle T_\varphi z^k, z^j\rangle_{L^2(\mathbb{T})} = \langle T_{\{c_k\}}\vec{e}_k, \vec{e}_j\rangle_{\ell^2(\mathbb{Z}_+)}.$$

This motivates us to define $U_1 : \sum_{k\in\mathbb{Z}_+} a_k z^k \mapsto \{a_k\}_{k\in\mathbb{Z}_+}$ and $U_2 : \overline{z\sum_{k\in\mathbb{Z}_+} a_k z^k} \mapsto \{a_k\}_{k\in\mathbb{Z}_+}$. With these definitions at hand, we have

$$\begin{aligned}\langle H_\varphi z^k, \overline{z}^{j+1}\rangle_{L^2(\mathbb{T})} &= \langle H_{\{c_k\}}U_1(z^k), U_2(\overline{z}^{j+1})\rangle_{\ell^2(\mathbb{Z}_+)},\\ \langle T_\varphi z^k, z^j\rangle_{L^2(\mathbb{T})} &= \langle T_{\{c_k\}}U_1(z^k), U_1(z^j)\rangle_{\ell^2(\mathbb{Z}_+)}.\end{aligned}$$

We see that $H_\varphi = U_2^{-1}H_{\{c_k\}}U_1$ and $T_\varphi = U_1^{-1}T_{\{c_k\}}U_1$ for our choice of sequences $\{c_k\}_k$. The theorem is proven by observing that polynomials are linear combinations of harmonics z^k, $k \in \mathbb{Z}_+$, and the inner product is bilinear. $\square$

4 Proofs of Nehari, Brown-Halmos and Sarason theorems

The goal of this section is clear from its title. We start with a "warm-up" result that demonstrates the main idea of the proofs that will follow next.

4.1 Warm-up: the operator norm of a multiplication operator

Theorem 4.1. *Let $\varphi \in L^\infty(\mathbb{T})$. Then the norm of the multiplication operator $M_\varphi : f \mapsto \varphi f$ on $L^2(\mathbb{T})$ is $\|M_\varphi\| = \|\varphi\|_{L^\infty(\mathbb{T})}$.*

Proof. For every $f \in L^2(\mathbb{T})$ we have

$$\|M_\varphi f\|_{L^2(\mathbb{T})} = \|\varphi f\|_{L^2(\mathbb{T})} \le \|\varphi\|_{L^\infty(\mathbb{T})} \|f\|_{L^2(\mathbb{T})}.$$

It follows that M_φ is a bounded operator on $L^2(\mathbb{T})$ with $\|M_\varphi\| \le \|\varphi\|_{L^\infty(\mathbb{T})}$. To prove that the last inequality is, in fact, equality, we write

$$\|M_\varphi\| = \sup_{\|f\|,\|g\| \le 1} |\langle M_\varphi f, g\rangle| = \sup_{h \in E} \left| \int_{\mathbb{T}} \varphi h \, dm \right|, \tag{1}$$

where $E = \{h : h = f\bar{g}, \text{ for } f, g \in B_{L^2(\mathbb{T})}[0,1]\}$. Here and below $B_{L^p(\mathbb{T})}[0,1]$ stands for the unit ball in the Banach space $L^p(\mathbb{T})$, $1 \le p \le \infty$. The crucial step of the proof is to show that $E = B_{L^1(\mathbb{T})}[0,1]$. In other words, we are going to prove that

$$B_{L^1(\mathbb{T})}[0,1] = B_{L^2(\mathbb{T})}[0,1] \cdot B_{L^2(\mathbb{T})}[0,1].$$

Inclusion $\supseteq$ is evident from Cauchy-Schwarz inequality. To prove the opposite inclusion $\subseteq$, we take arbitrary $h \in B_{L^1(\mathbb{T})}[0,1]$ and define the function $s \in B_{L^\infty(\mathbb{T})}[0,1]$ by

$$s = \begin{cases} \dfrac{h}{|h|} & \text{if } h \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $h = s \cdot |h|$. Then for functions $f, g \in B_{L^2(\mathbb{T})}[0,1]$ defined by

$$f = s\sqrt{|h|}, \qquad g = \sqrt{|h|},$$

we have $h = sfg = sf\bar{g}$, and equality $E = B_{L^1(\mathbb{T})}[0,1]$ follows. Let us now use duality relation $L^1(\mathbb{T})^* = L^\infty(\mathbb{T})$ to complete the proof. Define the linear functional $\Phi_\varphi : h \mapsto \int_{\mathbb{T}} \varphi \cdot h \, dm$ on the Banach space $L^1(\mathbb{T})$. We have

$$\|M_\varphi\| = \sup_{h \in B_{L^1(\mathbb{T})}[0,1]} \left| \int_{\mathbb{T}} \varphi \cdot h \, dm \right| = \|\Phi_\varphi\|_{L^1(\mathbb{T})^*} = \|\varphi\|_{L^\infty(\mathbb{T})}.$$

The result follows. □

4.2 Proof of Nehari theorem

In the equivalent setting of Hankel operators, Nehari theorem takes the following form:

Theorem 4.2 (Nehari). *Let $\varphi \in L^2(\mathbb{T})$ and consider the Hankel operator $H_\varphi : f \mapsto P_-(\varphi f)$ from $H^2(\mathbb{T})$ onto $\overline{zH^2(\mathbb{T})}$ densely defined on the set of polynomials $\mathcal{P} = span\{z^k : k \in \mathbb{Z}_+\}$. The following assertions are equivalent:*

1. *H_φ admits a bounded continuation to the whole space $H^2(\mathbb{T})$, i.e., $H_\varphi \in \mathcal{B}(H^2(\mathbb{T}), \overline{zH^2(\mathbb{T})})$,*

2. *H_φ has a bounded symbol $\tilde{\varphi}$, that is, $H_\varphi = H_{\tilde{\varphi}}$ for some $\tilde{\varphi} \in L^\infty(\mathbb{T})$.*

Moreover, we have $\|H_\varphi\| \le \|\tilde{\varphi}\|_{L^\infty(\mathbb{T})}$ for every symbol $\tilde{\varphi} \in L^\infty(\mathbb{T})$ and one can choose $\tilde{\varphi}$ so that $\|H_\varphi\| = \|\tilde{\varphi}\|_{L^\infty(\mathbb{T})}$. In other words,

$$\|H_\varphi\| = \inf\{\|\tilde{\varphi}\|_{L^\infty(\mathbb{T})} : \ \tilde{\varphi} \text{ is a symbol of } H_\varphi\}. \tag{2}$$

If $\varphi \in L^\infty(\mathbb{T})$, we also have $\|H_\varphi\| = \operatorname{dist}_{L^\infty(\mathbb{T})}(\varphi, H^\infty)$.

Proof. Implication (2) ⇒ (1) is easy. Indeed, if $H_\varphi = H_{\tilde{\varphi}}$ for some $\tilde{\varphi} \in L^\infty(\mathbb{T})$, then

$$\|H_{\tilde{\varphi}} f\|_{\overline{zH^2(\mathbb{T})}} = \|P_- M_{\tilde{\varphi}} f\|_{\overline{zH^2(\mathbb{T})}} \le \|M_{\tilde{\varphi}} f\|_{L^2(\mathbb{T})} \le \|\tilde{\varphi}\|_{L^\infty(\mathbb{T})} \|f\|_{H^2(\mathbb{T})},$$

because the norm of the orthogonal projection P_- is equal to one. Thus, H_φ admits a bounded continuation to the whole space $H^2(\mathbb{T})$, and moreover, $\|H_\varphi\| \le \|\tilde{\varphi}\|_{L^\infty(\mathbb{T})}$ for every symbol $\tilde{\varphi} \in L^\infty(\mathbb{T})$. To prove implication (1) ⇒ (2) and formula (2), we will take an arbitrary Hankel operator H_φ that admits bounded continuation to the whole space $H^2(\mathbb{T})$ and find a symbol $\tilde{\varphi}$ such that $\|H_\varphi\| = \|\tilde{\varphi}\|_{L^\infty(\mathbb{T})}$. The proof consists of two steps.

Step 1. Let us show that the bounded extension of H_φ to the whole space $H^2(\mathbb{T})$ acts on functions $f \in C(\overline{\mathbb{D}}) \cap H^2(\mathbb{T})$ by the same formula $H_\varphi(f) = P_-(\varphi f)$. Take $f \in C(\overline{\mathbb{D}}) \cap H^2(\mathbb{T})$ and find a sequence $\{f_n\} \subset \mathcal{P}$ such that $f_n \to f$ in $L^\infty(\mathbb{T})$. This sequence can be found, e.g., as follows. By Fejér theorem on Cesàro convergence of Fourier series, we have

$$\left\| f(z) - \frac{1}{n}\sum_{k=0}^{n-1} S_k(f,z) \right\|_{L^\infty(\mathbb{T})} \to 0, \qquad S_k(f,z) = \sum_{j=-k}^{k} \hat{f}(j) z^j.$$

Since $f \in H^2(\mathbb{T})$, we in fact have $S_k(f,z) = \sum_{j=0}^{k} \hat{f}(j) z^j$ and hence $\frac{1}{n}\sum_{k=0}^{n-1} S_k(f,z) \in \mathcal{P}$ for every n. Then, for the sequence $f_n = \frac{1}{n}\sum_{k=0}^{n-1} S_k(f,z)$, we can estimate

$$\begin{aligned} \|H_\varphi f - P_-(\varphi f)\|_{L^2(\mathbb{T})} &\le \|H_\varphi (f - f_n)\|_{L^2(\mathbb{T})} + \|P_-(\varphi(f - f_n))\|_{L^2(\mathbb{T})} \\ &\le \|H_\varphi\| \|f - f_n\|_{L^2(\mathbb{T})} + \|\varphi\|_{L^2(\mathbb{T})} \|f - f_n\|_{L^\infty(\mathbb{T})} \to 0. \end{aligned}$$

It follows that $H_\varphi f = P_-(\varphi f)$ for every function $f \in C(\overline{\mathbb{D}}) \cap H^2(\mathbb{T})$ and the first step is completed.

Step 2. Consider the linear subset $H^1_e(\mathbb{T})$ in $H^1(\mathbb{T})$ of all functions f that admit analytic extension to the open disk $|z| < 1 + \varepsilon_f$ for some $\varepsilon_f > 0$. Denote by $B_{H^1_e(\mathbb{T})}[0,1]$ the unit ball in this linear normed space (with the norm inherited from $H^1(\mathbb{T})$). Let also $B_{H^2_c(\mathbb{T})}[0,1]$ be the unit ball in the linear normed space $C(\overline{\mathbb{D}}) \cap H^2(\mathbb{T})$ with the norm inherited from $H^2(\mathbb{T})$. We claim that $B_{H^1_e(\mathbb{T})}[0,1] \subset B_{H^2_c(\mathbb{T})}[0,1] \cdot B_{H^2_c(\mathbb{T})}[0,1]$. To this end, take $h \in B_{H^1_e(\mathbb{T})}[0,1]$ and let $\{\lambda_i\}_{k=1}^n$ be the zeros of h in $\mathbb{D}$ counted with multiplicities (note that n is finite). Define the Blaschke product

$$B(z) = \prod_{k=1}^{n} \frac{z - \lambda_k}{1 - \overline{\lambda_k} z}$$

and observe that $\frac{h}{B}$ is holomorphic in a neighborhood of $\overline{\mathbb{D}}$ and has no zeroes in $\mathbb{D}$. Then, functions

$$f = B \cdot \sqrt{\frac{h}{B}}, \qquad g = \sqrt{\frac{h}{B}}$$

belong to $B_{H^2_c(\mathbb{T})}[0,1]$ and satisfy $h = fg$, as required.

Step 3. We now ready to prove that arbitrary Hankel operator H_φ that admits bounded continuation to the whole space $H^2(\mathbb{T})$ has a symbol $\tilde{\varphi}$ such that $\|H_\varphi\| = \|\tilde{\varphi}\|_{L^\infty(\mathbb{T})}$. For this, we define the linear functional

$$\Phi_\varphi : h \mapsto \int_{\mathbb{T}} \varphi z h \, dm, \qquad h \in H^1_e(\mathbb{T}).$$

For every $h \in B_{H^1_e(\mathbb{T})}[0,1]$, we have

$$\Phi_\varphi(h) = \int_{\mathbb{T}} \varphi z f g \, dm = \langle \varphi f, \overline{zg} \rangle_{L^2(\mathbb{T})} = \langle P_-(\varphi f), \overline{zg} \rangle_{L^2(\mathbb{T})} = \langle H_\varphi f, \overline{zg} \rangle_{L^2(\mathbb{T})},$$

where f, g are the functions constructed at Step 2. Note also that we have used Step 1 when wrote $P_-(\varphi f) = H_\varphi f$. Now, we see that

$$\|\Phi_\varphi\| = \sup_{h \in B_{H^1_e(\mathbb{T})}[0,1]} |\Phi_\varphi(h)| \le \sup_{f,g \in B_{H^2_e(\mathbb{T})}[0,1]} |\langle H_\varphi f, \overline{zg}\rangle_{L^2(\mathbb{T})}| \le \|H_\varphi\|.$$

By Hahn-Banach theorem, there exists a continuation $\tilde{\Phi}_\varphi$ of Φ_φ to $L^1(\mathbb{T})$ having the same norm. Since $L^1(\mathbb{T})^* = L^\infty(\mathbb{T})$, we see that there exists $\tilde{\varphi} \in L^\infty(\mathbb{T})$ such that

$$\tilde{\Phi}_\varphi(h) = \int_{\mathbb{T}} \tilde{\varphi} h \, dm, \qquad \|\tilde{\varphi}\|_{L^\infty(\mathbb{T})} = \|\tilde{\Phi}\| = \|\Phi_\varphi\| \le \|H_\varphi\|.$$

For this $\tilde{\varphi}$ we have $H_\varphi = H_{\bar{z}\tilde{\varphi}}$ because

$$\langle H_{\bar{z}\tilde{\varphi}} f, \overline{zg}\rangle_{L^2(\mathbb{T})} = \int_{\mathbb{T}} \tilde{\varphi} f g \, dm = \tilde{\Phi}(fg) = \Phi_\varphi(fg) = \langle H_\varphi f, \overline{zg}\rangle_{L^2(\mathbb{T})}$$

for all $f, g \in H^2_e(\mathbb{T})$. Hence, the operators H_φ, $H_{\bar{z}\tilde{\varphi}}$ have the same values of bilinear forms on a dense subset $H^2(\mathbb{T})$, which (using continuity H_φ, $H_{\bar{z}\tilde{\varphi}}$) implies that they coincide on the whole space $H^2(\mathbb{T})$. Thus, we found $\bar{z}\tilde{\varphi} \in L^\infty(\mathbb{T})$ such that $H_\varphi = H_{\bar{z}\tilde{\varphi}}$ and $\|H_{\bar{z}\tilde{\varphi}}\| \le \|\bar{z}\tilde{\varphi}\|_{L^\infty(\mathbb{T})} \le \|H_\varphi\| = \|H_{\bar{z}\tilde{\varphi}}\|$. In particular, $\bar{z}\tilde{\varphi}$ is a symbol of H_φ and $\|H_\varphi\| = \|\bar{z}\tilde{\varphi}\|_{L^\infty(\mathbb{T})}$.

It remains to verify the formula $\|H_\varphi\| = \operatorname{dist}_{L^\infty(\mathbb{T})}(\varphi, H^\infty)$ in the case $\varphi \in L^\infty(\mathbb{T})$. Note that $H_\psi = 0$ for $\psi \in L^2(\mathbb{T})$ if and only if $\psi \in H^2$. Thus, the bound $\|H_\varphi\| \le \|\varphi\|_{L^\infty(\mathbb{T})}$ can be improved to $\|H_\varphi\| \le \operatorname{dist}_{L^\infty(\mathbb{T})}(\varphi, H^\infty)$. On the other hand, $\varphi - \tilde{\varphi}$ generates zero Hankel operator and hence it belongs to $H^2(\mathbb{T}) \cap L^\infty(\mathbb{T}) = H^\infty(\mathbb{T})$ by adding to φ arbitrary elements from $H^\infty(\mathbb{T})$. Therefore, we have

$$\operatorname{dist}_{L^\infty(\mathbb{T})}(\varphi, H^\infty) \le \|\tilde{\varphi}\| = \|H_\varphi\|,$$

and hence, equality $\|H_\varphi\| = \operatorname{dist}_{L^\infty(\mathbb{T})}(\varphi, H^\infty)$ holds. □

4.3 Proof of Brown-Halmos theorem

Let us start with the restatement of the Brown-Halmos theorem for Toeplitz operators. This result is equivalent to the Brown-Halmos theorem of Toeplitz matrices.

Theorem 4.3 (Brown-Halmos). *Let $\varphi \in L^2(\mathbb{T})$ and let $T_\varphi : f \mapsto P_-(\varphi f)$ be the Toeplitz operator densely defined on $H^2(\mathbb{T})$ on the set of polynomials $\mathcal{P} = span\{z^k : k \in \mathbb{Z}_+\}$. The operator T_φ admits a bounded extension to the whole space $H^2(\mathbb{T})$ if and only if $\varphi \in L^\infty(\mathbb{T})$.*

Moreover, we have

$$\|T_\varphi\| = \|\varphi\|_{L^\infty(\mathbb{T})}.$$

We will need the following simple lemma.

Lemma 4.4. *For every point z in the open unit disk $\mathbb{D}$, we have*

$$\frac{1}{1-\overline{z}\xi} \in H^2(\mathbb{D}) \quad \text{and} \quad \Big\|\frac{1}{1-\overline{z}\xi}\Big\|_{H^2(\mathbb{D})} = \frac{1}{\sqrt{1-|z|^2}}.$$

Proof. Consider the function

$$f(\xi) = \frac{1}{1-\overline{z}\xi} = \sum_{k=0}^{\infty} \overline{z}^k \xi^k. \tag{3}$$

Since $|z| < 1$, the sum on the right hand side converges. From equation (3) and Parseval identity, we obtain

$$\|f\|^2_{L^2(\mathbb{T})} = \sum_{k=-\infty}^{\infty} |\hat{f}(k)|^2 = \sum_{k=0}^{\infty} |\overline{z}^k|^2 = \frac{1}{1-|z|^2}.$$

Since $\hat{f}(k) = 0$ for $k < 0$, we also see that $f \in H^2(\mathbb{T})$. □

Proof of Brown-Halmos theorem. We need to check two inequalities:

$$\|T_\varphi\| \le \|\varphi\|_{L^\infty(\mathbb{T})}, \qquad \|\varphi\|_{L^\infty(\mathbb{T})} \le \|T_\varphi\|.$$

The first inequality is almost evident. Indeed, since P_+ is an orthogonal projection, we have $\|P_+\| \le 1$. Thus, for every polynomial f we have

$$\|T_\varphi f\|_{L^2(\mathbb{T})} = \|P_+(\varphi f)\|_{L^2(\mathbb{T})} \le \|\varphi f\|_{L^2(\mathbb{T})} \le \|\varphi\|_{L^\infty(\mathbb{T})} \|f\|_{L^2(\mathbb{T})}.$$

It follows that $\|T_\varphi\| \le \|\varphi\|_{L^\infty(\mathbb{T})}$.

Let us now prove the second inequality. In case $\|T_\varphi\| = +\infty$ there is nothing to prove. So, assume that T_φ admits a bounded extension to the space $H^2(\mathbb{T})$. Then, arguing exactly as in the proof of Nehari theorem (see Step 1), it is easy to check that the bounded extension of T_φ to $H^2(\mathbb{T})$ acts on functions $f \in C(\overline{\mathbb{D}}) \cap H^2(\mathbb{T})$ by the same formula $T_\varphi(f) = P_+(\varphi f)$ as for $f \in \mathcal{P}$. Next, denote

$$B_{H^2_c}[0,1] = \{f \in C(\overline{\mathbb{D}}) \cap H^2(\mathbb{T}) : \ \|f\|_{L^2(\mathbb{T})} \le 1\}.$$

For $f, g \in C(\overline{\mathbb{D}}) \cap H^2(\mathbb{T})$, we have

$$\langle T_\varphi f, g\rangle = \langle P_+(\varphi f), g\rangle = \langle \varphi f, P_+(g)\rangle = \langle \varphi f, g\rangle$$

Consequently,

$$\|T_\varphi\| = \sup_{f,g \,\in\, B_{H^2_c}[0,1]} \left| \int_{\mathbb{T}} \varphi f \,\overline{g}\, dm \right|.$$

Let us take $f = g = k_z$, for k_z defined by

$$k_z(\xi) = \frac{\sqrt{1-|z|^2}}{1-\overline{z}\xi}, \qquad \xi \in \mathbb{T}.$$

By Lemma 4.4, for each $z \in \mathbb{D}$ we have $k_z \in B_{H^2_c}[0,1]$. It follows that

$$\|T_\varphi\| \geq \sup_{z \,\in\, \mathbb{D}} \left| \int_{\mathbb{T}} \varphi |k_z|^2 \, dm \right|.$$

Since $|k_z|^2$ is the standard Poisson kernel, for almost every $\zeta \in \mathbb{T}$ we have

$$\lim_{\substack{r\to 1\\ r<1}} \int_{\mathbb{T}} \varphi |k_{r\zeta}|^2 \, dm = \varphi(\zeta). \tag{4}$$

It follows that $\varphi \in L^\infty(\mathbb{T})$ and $\|\varphi\|_{L^\infty(\mathbb{T})} \leq \|T_\varphi\|$. $\square$

4.4 Proof of Sarason theorem

We again start with rewriting the statement in the language of operators of function spaces. For this, we need the following definition.

Definition 4.5. Let $\mathcal{P}_n = \operatorname{span}\{z^k,\, 0 \leq k \leq n-1\}$ be the space of polynomials of degree at most $n-1$. We define the orthogonal projection $P_n : L^2(\mathbb{T}) \to \mathcal{P}_n$ by

$$P_n : \sum_{k=-\infty}^{\infty} a_k z^k \mapsto \sum_{k=0}^{n-1} a_k z^k,$$

and the truncated Toeplitz operator $A_\varphi : \mathcal{P}_n \to \mathcal{P}_n$ with symbol $\varphi \in L^2(\mathbb{T})$ by

$$A_\varphi : f \mapsto P_n(\varphi f), \quad f \in \mathcal{P}_n.$$

With this definition, A_φ is a linear operator on the space Hilbert space $\mathcal{P}_n$ (that we equip with the inner product and norm inherited from $L^2(\mathbb{T})$). Since $\mathcal{P}_n$ is finitely-dimensional, we have $A_\varphi \in \mathcal{B}(\mathcal{P}_n)$.

Let us compute the matrix of A_φ in the orthonormal basis $\{z^k\}_{k=0}^{n-1}$ of $\mathcal{P}_n$:

$$\langle A_\varphi z^k, z^j\rangle = \langle P_n(\varphi z^k), z^j\rangle = \langle \varphi z^k, P_n(z^j)\rangle = \langle \varphi z^k, z^j \langle = \langle \varphi, z^{j-k}\rangle = \hat{\varphi}(j-k).$$

Hence, the matrix of A_φ is

$$A_\varphi = \begin{bmatrix} \hat{\varphi}(0) & \hat{\varphi}(-1) & \hat{\varphi}(-2) & \cdots & \hat{\varphi}(-(n-1)) \\ \hat{\varphi}(1) & \hat{\varphi}(0) & \hat{\varphi}(-1) & \cdots & \hat{\varphi}(-(n-2)) \\ \hat{\varphi}(2) & \hat{\varphi}(1) & \hat{\varphi}(0) & \cdots & \hat{\varphi}(-(n-3)) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{\varphi}(n-1) & \hat{\varphi}(n-2) & \hat{\varphi}(n-3) & \cdots & \hat{\varphi}(0) \end{bmatrix}. \tag{5}$$

We see that A_φ is a truncated Toeplitz matrix of size $n \times n$. Note also that symbols $\varphi \in H^\infty(\mathbb{T})$ generate lower triangular matrices A_φ (indeed, we have $\hat{\varphi}(-k) = 0$, $\forall k > 1$). Conversely, any lower triangular matrix of the form

$$T_{n+1} = \begin{bmatrix} a_0 & 0 & 0 & \dots & 0 \\ a_1 & a_0 & 0 & \dots & 0 \\ a_2 & a_1 & a_0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_n & a_{n-1} & a_{n-2} & \dots & a_0 \end{bmatrix}$$

is the truncated Toeplitz matrix A_φ for some $\varphi \in H^\infty(\mathbb{T})$. One can take, say,

$$\varphi = a_0 + a_1 z + \ldots + a_n z^n.$$

With this remarks, we can state the operator form of Sarason theorem.

Theorem 4.6. *Let $\varphi \in H^\infty(\mathbb{T})$, then there exist $\tilde{\varphi} \in H^\infty(\mathbb{T})$ such that*

$$A_\varphi = P_n T_{\tilde{\varphi}}|_{\mathcal{P}_n} \quad \text{and} \quad \|A_\varphi\| = \|T_{\tilde{\varphi}}\| = \|\tilde{\varphi}\|.$$

Remark. This version of Sarason theorem is for lower triangular matrices/operators. The passage from lower to upper triangular matrices/operators can be achieved by considering adjoint operators, which preserves both the Toeplitz structure and the operator norm.

Lemma 4.7. *For $f \in H^2(\mathbb{T})$ we have $P_n f = z^n P_-(\overline{z}^n f)$.*

Proof. Let $f = \sum_{k=0}^{\infty} a_k z^k$. Then,

$$\overline{z}^n f = \overline{z}^n a_0 + \overline{z}^{n-1} a_1 + \ldots + \overline{z} a_{n-1} + a_n + z a_{n+1} + \ldots$$

Consequently,

$$P_-(\overline{z}^n f) = \overline{z}^n a_0 + \overline{z}^{n-1} a_1 + \cdots + \overline{z} a_{n-1},$$

and

$$z^n P_-(\overline{z}^n f) = a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} = P_n f,$$

which yields the statement. □

Proof of Sarason theorem. Let $\varphi \in H^\infty(\mathbb{T})$. Then for every $g \in H^2(\mathbb{T})$, the first $n-1$ Fourier coefficients of $\varphi z^n g$ are equal to 0. In other words, we have

$$P_n(\varphi z^n g) = 0.$$

It follows that $P_n(\varphi f) = P_n(\varphi(f + z^n g))$. Thus,

$$\|A_\varphi\| = \sup_{\substack{f \in \mathcal{P}_n \\ \|f\| \le 1}} \|P_n(\varphi f)\| = \sup_{\substack{f \in \mathcal{P}_n,\, g \in H^2(\mathbb{T}), \\ \|f\|^2 + \|z^n g\|^2 \le 1}} \|P_n(\varphi(f + z^n g))\| = \sup_{\substack{\tilde{f} \in H^2(\mathbb{T}) \\ \|\tilde{f}\| \le 1}} \|P_n(\varphi \tilde{f})\|.$$

We used here the fact that every function $\tilde{f} \in H^2(\mathbb{T})$ has the orthogonal decomposition $\tilde{f} = (P_n)(\tilde{f}) + (I - P_n)(\tilde{f})$, which can be rewritten in the form $\tilde{f} = f + z^n g$ for some $f \in \mathcal{P}_n$, $g \in H^2(\mathbb{T})$ such that $\|\tilde{f}\|^2 = \|f\|^2 + \|z^n g\|^2$. By previous lemma and in view of the fact that $z \in \mathbb{T}$, we can write

$$\|A_\varphi\| = \sup_{\substack{\tilde{f} \in H^2(\mathbb{T}) \\ \|\tilde{f}\| \le 1}} \|P_n(\varphi \tilde{f})\| = \sup_{\substack{\tilde{f} \in H^2(\mathbb{T}) \\ \|\tilde{f}\| \le 1}} \|z^n P_-(\overline{z}^n \varphi \tilde{f})\| = \sup_{\substack{\tilde{f} \in H^2(\mathbb{T}) \\ \|\tilde{f}\| \le 1}} \|P_-(\overline{z}^n \varphi \tilde{f})\| = \|H_\psi\|,$$

where we set $\psi = \overline{z}^n \varphi$. By Nehari theorem, there exists $\psi_1 \in L^\infty(\mathbb{T})$ such that $H_\psi = H_{\psi_1}$ and $\|\psi_1\|_{L^\infty(\mathbb{T})} = \|H_\psi\| = \|H_{\psi_1}\|$. In particular, we have

$$\|A_\varphi\| = \|\psi_1\|_{L^\infty(\mathbb{T})}. \tag{6}$$

Since $\psi = \overline{z}^n \varphi$ and $\psi - \psi_1 \in H^2(\mathbb{T}) \cap L^\infty(\mathbb{T}) = H^\infty(\mathbb{T})$, hence for some $h \in H^\infty(\mathbb{T})$ we can write

$$\overline{z}^n \varphi = \psi_1 + h \quad \Longrightarrow \quad \varphi = z^n \psi_1 + z^n h. \tag{7}$$

Equation (7) implies $\widehat{(z^n\psi_1)}(k) = \hat{\varphi}(k)$, for all $-(n-1) \leq k \leq (n-1)$. In particular, we have

$$A_\varphi = A_{z^n\psi_1} = P_n T_{z^n\psi_1}|_{\mathcal{P}_n}.$$

Taking Brown-Halmos theorem and equation (6) into consideration, we get

$$\|T_{z^n\psi_1}\| = \|z^n\psi_1\|_{L^\infty(\mathbb{T})} = \|\psi_1\|_{L^\infty(\mathbb{T})} = \|A_\varphi\|.$$

Thus, one can take $\tilde{\varphi} = z^n\psi_1$ in the statement of Sarason theorem. $\square$

5 Shift operator

This section is devoted to the shift operator – the most important operator in function-theoretic operator theory.

Definition 5.1. Operator $S : H^2 \to H^2$ defined by

$$S : f \mapsto zf$$

is called **the shift operator** on the Hardy space H^2.

Note that $zH^2 \subset H^2$, hence, the shift operator S is just the restriction of the multiplication operator on $L^2(\mathbb{T})$ to one of its invariant subspaces, H^2. Let us also note that S coincides with the Toeplitz operator with symbol $\varphi = z$, i.e., $S = T_z$. Our goal in this section is to describe the lattice of invariant subspaces of $S : H^2 \to H^2$. We will see that the structure of invariant subspaces is very rich from the point of view of function theory. The operator S can be considered as *a functional model* for the operator

$$\hat{S} : \{x_n\}_{n \in \mathbb{Z}_+} \mapsto \{x_{n-1}\}_{n \in \mathbb{Z}_+}, \qquad x_{-1} = 0,$$

on the Hilbert space $\ell^2(\mathbb{Z}_+)$. This model, in particular, describes invariant subspaces of $S \cong \hat{S}$ in a very explicit form. It is possible to construct a functional model (a unitary equivalent operator that acts on some Hilbert space of functions by a concrete formula) for every linear bounded operator on a Hilbert space. While we do not pursue this direction here, let us mention that this model can be also given in terms of some shift operator.

5.1 Wiener theorem

We start with describing reducing subspaces of the unitary operator $M_z : f \mapsto zf$ on the Hilbert space $L^2(\mu)$.

Theorem 5.2 (Wiener)**.** *Let μ be a finite Borel measure on the unit circle $\mathbb{T}$ and let E be a closed subspace of $L^2(\mu)$ such that $zE = E$. Then $E = \chi_F L^2(\mu)$ for some Borel subset $F \subset \mathbb{T}$.*

Proof. Denote by χ the orthogonal projection to E of the constant function $1 \in L^2(\mu)$, so that $\chi = P_E 1$. We have

$$\int_{\mathbb{T}} \chi(1 - \overline{\chi}) z^k \, d\mu = (z^k \chi, 1 - \chi) = 0 \quad \text{for every } k \geq 0,$$

because $1-\chi = 1-P_E 1 = (I-P_E)1 = P_{E^\perp}1 \in E^\perp$ and $z^k\chi \in E$. For $k<0$, the integral above is also 0 because $z^{-k}\chi$ is in $z^{-k}E = E$. So, the function $g := \chi(1-\overline{\chi})$ satisfies

$$\int_{\mathbb{T}} gp\,d\mu = 0 \quad \text{for every trigonometric polynomial } p.$$

By Cauchy–Schwartz inequality one can easily check that $g \in L^1(\mu)$. Therefore if $p_n \to \varphi$ uniformly on $\mathbb{T}$, then

$$\int_{\mathbb{T}} g\varphi\,d\mu = 0. \tag{8}$$

Since every continious function can be approximated with polynomials, equation (8) holds for every $\varphi \in C(\mathbb{T})$. It follows that the signed measure $g\,d\mu$ is identically zero, i.e., $g=0$ μ-almost everywhere on $\mathbb{T}$. It follows that $\chi(\xi)$ is either 0 or 1 for μ-almost every $\xi \in \mathbb{T}$. Then $\chi = \chi_F$, where $F = \chi^{-1}(\{1\})$ is the Borel subset of $\mathbb{T}$. Since $zE = E$ and $\chi_F \in E$, we see that

$$\operatorname{clos}_{L^2(\mu)}\{\chi_F p :\ p \text{ is trigonometric polynomial}\} \subseteq E,$$

that is, $\chi_F L^2(\mu) \subseteq E$. It remains to check that $E = \chi_F L^2(\mu)$. Assume that $e \in E$ is orthogonal to $\chi_F L^2(\mu)$. It is enought to show that $e = 0$. Since $e \perp \chi_F L^2(\mu)$, we have

$$\int_{\mathbb{T}} e\chi_F z^k\,d\mu = 0, \quad k \in \mathbb{Z}.$$

Since $ez^k \in E$ for every $k \in \mathbb{Z}$ and $1-\chi_F \in E^\perp$, we also have

$$\int_{\mathbb{T}} e(1-\chi_F)z^k\,d\mu = 0, \quad k \in \mathbb{Z}.$$

Summing up these two relations, we obtain $e \perp z^k$ in $L^2(\mu)$ for every $k \in \mathbb{Z}$. Therefore, we have $e = 0$ in $L^2(\mu)$. □

5.2 Beurling theorem

Theorem 5.3 (Beurling)**.** *A proper nonzero subspace $E \subset L^2(\mathbb{T})$ satisfies $zE \subsetneq E$ if and only if $E = \theta H^2$ for some measurable function θ on $\mathbb{T}$ such that $|\theta| = 1$ almost everywhere on $\mathbb{T}$.*

Proof. Assume first $E = \theta H^2$ for some θ on $\mathbb{T}$ such that $|\theta| = 1$ almost everywhere on $\mathbb{T}$. Then

$$zE = z\theta H^2 = \theta(zH^2) \subsetneq \theta H^2 = E,$$

where the inclusion is proper because $1 \in H^2 \ominus zH^2$. Thus, we have proved inclusion $zE \subsetneq E$.

Now assume that $zE \subsetneq E$ and take $\theta \in E \ominus zE$ with $\|\theta\| = 1$. For every $k \geq 1$ we have

$$\int_{\mathbb{T}} |\theta|^2 z^k \, dm = \langle \underbrace{z^k\theta}_{\substack{\cap \\ zE}}, \underbrace{\theta}_{\substack{\cap \\ E}} \rangle_{L^2} = 0.$$

Since for every $k \leq -1$ we have $\int_{\mathbb{T}} |\theta|^2 z^k \, dm = 0$ by complex conjugation, it follows that the function $|\theta|^2 \in L^1(\mathbb{T})$ satisfies

$$\widehat{|\theta|^2}(k) = 0, \qquad k \in \mathbb{Z} \setminus \{0\}.$$

So, $|\theta|^2$ is a constant function. Since $\|\theta\|_{L^2(\mathbb{T})} = 1$, we have $|\theta|^2 = 1$ almost everywhere on $\mathbb{T}$. Using $\theta \in E$ and $z^k E \subset E$, $k \geq 1$, we obtain

$$E \supseteq \theta \operatorname{clos}_{L^2(\mathbb{T})}(\mathcal{P}) = \theta H^2.$$

To prove that $E = \theta H^2$, consider an element $e \in E \ominus \theta H^2$. We need to show that $e = 0$. Directly from the orthogonality $e \perp \theta H^2$ we get

$$\int_{\mathbb{T}} e \, \overline{\theta} \, \overline{z}^k \, dm = \langle e, z^k\theta \rangle = 0, \qquad k \geq 0.$$

Furthermore, since, by definition, $\theta \perp zE$, we also have

$$\int_{\mathbb{T}} e \, \overline{\theta} \, z^k \, dm = \langle ez^k, \theta \rangle = 0, \qquad k \geq 1.$$

Therefore, $\widehat{(e\overline{\theta})}(k) = 0$ for every $k \in \mathbb{Z}$. From here we get $e\overline{\theta} = 0$ almost everywhere on $\mathbb{T}$. Since $|\theta|^2 = 1$ almost everywhere on $\mathbb{T}$, this means that $e = 0$ in $L^2(\mathbb{T})$, and the result follows. $\square$

Theorem 5.4 (Beurling theorem for the shift operator). *A nonzero subspace $E \subset H^2$ is invariant under the shift operator S on H^2 if and only if $E = \theta H^2$ for some function $\theta \in H^\infty$ such that $|\theta| = 1$ almost everywhere on $\mathbb{T}$.*

Proof. Assume that a nonzero subspace $E \subset H^2$ is invariant under the shift operator S on H^2. We may have two possibilities: $S(E) = E$ or $S(E) \subsetneq E$. In the case $S(E) = E$ by Wiener theorem there exists a Borel subset $F \subset \mathbb{T}$ such that $E = \chi_F L^2(\mathbb{T})$. In particular, for every $h \in E$, we have $h(\xi) = 0$ almost everywhere on F. By the uniqueness theorem for functions in the Hardy space, $h = 0$. It follows that $E = \{0\}$, contradicting to the assumption of the theorem. In other words, the case $S(E) = E$ does not arise. Consider now the case

$S(E) \subsetneq E$. By Theorem 5.3, we have $E = \theta H^2$ for some measurable function θ such that $|\theta| = 1$ almost everywhere on $\mathbb{T}$. In particular, since $\theta \in E$ and $E \subset H^2$, we have $\theta \in H^2$. Assumption $|\theta| = 1$ implies $\theta \in H^\infty$. $\square$

5.3 Some examples

Example 5.5. Take $z_0 \in \mathbb{D}$ and consider the subspace $E = \{f \in H^2 : f(z_0) = 0\}$. Clearly, $E \neq \{0\}$ and $S(E) \subset E$, so by Beurling theorem there exists $\theta_{z_0} \in H^\infty$, such that $|\theta_{z_0}| = 1$ almost everywhere on $\mathbb{T}$, $E = \theta_{z_0} H^2$. Let us find this function θ_{z_0}. Note that $\theta_{z_0}(z_0) = 0$. Moreover, if $\theta_{z_0}(z_1) = 0$ for some $z_1 \in \mathbb{D}\backslash\{z_0\}$, then $f(z_1) = 0$ for every $f \in E$, which is not the case. Summarizing, $\theta_{z_0}(z_0) = 0$, θ_{z_0} does not vanish in $\mathbb{D}\backslash\{z_0\}$, and $|\theta_{z_0}| = 1$ almost everywhere on $\mathbb{T}$. It is easy to gess such a function:

$$\theta_{z_0} = \frac{z - z_0}{1 - \overline{z_0} z} \quad \text{(Möbius transformation)}.$$

Note that θ_{z_0} satisfies all above assumptions. We actually have

$$E = \theta_{z_0} H^2 \qquad \Longleftrightarrow \qquad \left(\begin{cases} f \in H^2 \\ f(z_0) = 0 \end{cases} \quad \Longrightarrow \frac{f}{\theta_{z_0}} \in H^2 \right).$$

The right hand side of this equivalence indeed holds true, because $\frac{f}{\theta_{z_0}}$ is analytic in $\mathbb{D}$ and satisfies

$$\sup_{0<r<1} \int_{\mathbb{D}} \left| \frac{f(rz)}{\theta_{z_0}(rz)} \right|^2 dm < \infty,$$

where we used the fact that $|\theta_{z_0}(z)| \to 1$ as $|z| \to 1$. Thus, we have shown that $E = \theta_{z_0} H^2$.

Example 5.6. Let Λ be a finite subset of $\mathbb{D}$ and let $E = \{f \in H^2(\mathbb{D}) : f(z) = 0 \ \forall z \in \Lambda\}$. Then $E \neq \{0\}$ and $S(E) \subset E$. In this case,

$$E = \theta_\Lambda H^2, \qquad \theta_\Lambda = \prod_{z_k \in \Lambda} \frac{z - z_k}{1 - \overline{z_k} z}.$$

The function θ_Λ is called **a finite Blaschke product**.

Example 5.7. Let $z_0 \in \mathbb{D}$ and let $E = \{f \in H^2(\mathbb{D}) : f(z_0) = f'(z_0) = 0\}$. Then $E \neq \{0\}$ and $S(E) \subset E$. In this case,

$$E = \theta_{z_0}^2 H^2, \qquad \theta_{z_0}^2 = \left(\frac{z - z_0}{1 - \overline{z_0} z} \right)^2.$$

This example can be easily modificated to produce invariant subspaces $E = \theta_{z_0}^n H^2$.

Example 5.8. Let $z_0 \in \mathbb{T}$, $a > 0$, and let

$$E = \theta_{a,z_0} H^2, \qquad \theta_{a,z_0} = \exp\left(-a \frac{1 + \overline{z_0} z}{1 - \overline{z_0} z}\right).$$

Then $E \subset H^2$, $E \neq \{0\}$ and $S(E) \subset E$. The space $H^2 \ominus E$ sometimes is called the Paley-Wiener space in the open unit disk.

6 Hardy spaces, II

6.1 Convergence test for Blaschke products

Definition 6.1. Let $\{\lambda_k\}_{k=1}^{\infty}$ be a sequence of points in $\mathbb{D}$. We do not require $\lambda_k \neq \lambda_j$ for $k \neq j$. Thus, $\{\lambda_k\}_{k=1}^{\infty}$ might contain repeating elements. The function

$$B = \prod_{k=1}^{\infty} b_{\lambda_k}, \qquad b_\lambda = \begin{cases} \frac{|\lambda|}{\lambda}\frac{\lambda - z}{1-\bar{\lambda}z}, & \lambda \neq 0, \\ z & \lambda = 0, \end{cases}$$

is called **the infinite Blaschke product with zeros** $\{\lambda_k\}_{k=1}^{\infty}$. Note that factor $\frac{|\lambda|}{\lambda}$ is chosen so that $b_\lambda(0) = |\lambda| \geq 0$ for any $\lambda \in \mathbb{D}$.

Remark. Later on, we will often interpret $\frac{\lambda - z}{1-\bar{\lambda}z}$ as z if $\lambda = 0$.

Next theorem gives us exact conditions, under which Blaschke product converges to a not identically zero analytic function.

Theorem 6.2. *Blaschke product with zeros $\{\lambda_k\}_{k=1}^{\infty}$ converges to a not identically zero analytic function in $\mathbb{D}$ if and only if $\sum_{k=1}^{\infty}(1-|\lambda_k|) < \infty$.*

Proof. We may assume that $0 \notin \{\lambda_k\}_{k=1}^{\infty}$. Denote $\varepsilon_n = 1 - |\lambda_n|$ and let us compute

$$B(0) = \prod_{k=1}^{\infty} \frac{|\lambda_k|}{\lambda_k}\lambda_k = \prod_{k=1}^{\infty} |\lambda_k| = \prod_{k=1}^{\infty} e^{\log(1-(1-|\lambda_k|))} = e^{\sum_{k=1}^{\infty}\log(1-\varepsilon_k)}.$$

The sum $\sum_{k=1}^{\infty}\log(1-\varepsilon_k)$ converges to a finite number if and only if $\sum_{k=1}^{\infty}\varepsilon_k < \infty$. Thus, if the Blaschke product B converges at $z=0$ to $B(0) \neq 0$, then $\sum_{k=1}^{\infty}(1-|\lambda_k|) < \infty$.

For the opposite direction, we write $B = \prod_{k=1}^{\infty}(1+R_k)$, where

$$\begin{aligned} R_k &= \frac{|\lambda_k|}{\lambda_k}\frac{\lambda_k - z}{1-\overline{\lambda_k}z} - 1 = \frac{|\lambda_k|(\lambda_k - z) - \lambda_k + |\lambda_k|^2 z}{\lambda_k(1-\overline{\lambda_k}z)} = \\ &= \frac{\lambda_k(|\lambda_k|-1) + z|\lambda_k|(|\lambda_k|-1)}{\lambda_k(1-\overline{\lambda_k}z)}. \end{aligned}$$

We have

$$|R_k(z)| \leq \frac{(1-|\lambda_k|)(|\lambda_k| + |z||\lambda_k|)}{|\lambda_k|(1-|\lambda_k||z|)} \leq \frac{2(1-|\lambda_k|)}{\inf_{k\geq 1}|\lambda_k|(1-|z|)} = \varepsilon_k(z).$$

The sum $\sum_{k=1}^{\infty}\varepsilon_k(z)$ converges uniformly in $\{|z| < r\}$ for every $r < 1$. Consequently, the series in formula

$$B = e^{\sum_{k=1}^{\infty}\log(1+R_k)}$$

converges uniformly to an analytic function in each disk $|z| \le r < 1$. $\square$

6.2 Boundary values of Blaschke products

When B is a finite Blaschke product, we already know that $|B| \equiv 1$ on the unit circle $\mathbb{T}$. The following lemma shows that essentially the same holds for infinite convergent Blaschke products.

Lemma 6.3. *For every convergent Blaschke product B, we have*

$$\lim_{r\to 1} |B(rz)| = 1$$

for almost every $z \in \mathbb{T}$. In other words, $|B| = 1$ almost everywhere on $\mathbb{T}$.

Proof. Let B have zeroes $\{\lambda_k\}_{k=1}^{\infty}$, counted with multiplicity and such that $|\lambda_k| \le |\lambda_{k+1}|$ for each k. For a positive integer N, consider the finite Blaschke product

$$B_N = \prod_{k=1}^{N} \frac{|\lambda_k|}{\lambda_k} \frac{\lambda_k - z}{1 - \overline{\lambda_k} z}.$$

Note that B/B_N is the finite Blaschke product with zeroes $\{\lambda_k\}_{k=N+1}^{\infty}$, counted with their multiplicities. Consequently, it is bounded and holomorphic on $\mathbb{D}$. We can thus write

$$\frac{B}{B_N}(z) = \sum_{k=1}^{\infty} c_k z^k, \quad \text{and} \quad \frac{B}{B_N}(0) = \widehat{\frac{B}{B_N}}(0) = \int_{\mathbb{T}} \frac{B}{B_N}(\xi)\, dm(\xi).$$

Next, by definition we have

$$\left|\frac{B}{B_N}(0)\right| = \prod_{k=N+1}^{\infty} |\lambda_k|.$$

Since B/B_N is a convergent Blaschke product, the Blaschke condition in Theorem 6.2 implies that the product on the right hand side is nonzero for large N. Furthermore, as $N \to \infty$, this product converges to 1.

On the other hand, recalling $|B_N| \equiv 1$ on $\mathbb{T}$, we have

$$\left|\frac{B}{B_N}(0)\right| = \left|\int_{\mathbb{T}} \frac{B}{B_N}(\xi)\, dm(\xi)\right| \le \int_{\mathbb{T}} \left|\frac{B}{B_N}(\xi)\right|\, dm(\xi) = \int_{\mathbb{T}} |B(\xi)|\, dm(\xi) \le 1,$$

where the last equality follows from the inequality $|B(\xi)| \le 1$ that holds almost everywhere on $\mathbb{T}$. We thus have the chain of inequalities

$$1 = \lim_{N\to\infty} \prod_{k=N+1}^{\infty} |\lambda_k| \leq \int_{\mathbb{T}} |B(\xi)|\, dm(\xi) \leq 1,$$

implying

$$\int_{\mathbb{T}} |B(\xi)|\, dm(\xi) = 1.$$

Finally, we get $\lim_{r\to1} |B(r\xi)| = |B(\xi)| = 1$ for almost every $\xi \in \mathbb{T}$. □

6.3 Blaschke factors of functions in Hardy spaces

We will next aim to prove the following theorem, describing a decomposition of a function in $H^p(\mathbb{T})$ into a Blaschke product and a function without zeroes in $\mathbb{D}$.

Theorem 6.4. *Let $f \in H^p(\mathbb{T})$, where $1 \leq p \leq \infty$. Then, the set of zeroes of f, $\{\lambda_k\}_{k=1}^N$, $0 \leq N \leq \infty$, counted with multiplicities, satisfies Blaschke condition $\sum_{k=1}^N (1-|\lambda_k|) < \infty$, and f has the representation $f = B\widetilde{f}$, where $B = \prod_{k=1}^N b_{\lambda_k}$, $\widetilde{f} \in H^p(\mathbb{T})$ and $\|f\|_{H^p(\mathbb{T})} = \|\widetilde{f}\|_{H^p(\mathbb{T})}$. The function $\widetilde{f}$ has no zeroes in $\mathbb{D}$.*

We begin with some preliminaries.

Lemma 6.5. *For $f \in H^1(\mathbb{T})$, we have $\|f_r\|_{L^1(\mathbb{T})} \leq \|f\|_{L^1(\mathbb{T})}$ where $f_r(z) = f(rz)$, $0 < r < 1$.*

Proof. As we know from Homework 1,

$$f(rz) = \int_{\mathbb{T}} f(\xi) \frac{1-|rz|^2}{|1-r\overline{z}\xi|^2}\, dm(\xi), \qquad z \in \mathbb{T}.$$

Using Fubini theorem, we thus obtain

$$\begin{aligned} \int_{\mathbb{T}} |f(rz)|\, dm(z) &\leq \int_{\mathbb{T}} \int_{\mathbb{T}} |f(\xi)| \frac{1-|r\xi|^2}{|1-r\overline{z}\xi|^2}\, dm(\xi)\, dm(z) \\ &= \int_{\mathbb{T}} |f(\xi)| \int_{\mathbb{T}} \frac{1-|r\xi|^2}{|1-r\overline{z}\xi|^2}\, dm(z)\, dm(\xi) \\ &= \int_{\mathbb{T}} |f(\xi)|\, dm(\xi). \end{aligned}$$

This proves the claim. □

Corollary 6.6. For every $f \in H^1(\mathbb{T})$, the family $\|f_r\|_{L^1(\mathbb{T})}$ is increasing and, moreover, we have

$$\sup_{0<r<1} \|f_r\|_{L^1(\mathbb{T})} = \lim_{r\to1} \|f_r\|_{L^1(\mathbb{T})} = \|f\|_{L^1(\mathbb{T})}.$$

Proof. Let $r_1 < r_2$. Since $f_{r_2} \in H^1(\mathbb{T})$, we can apply Lemma 6.5 with $f = f_{r_2}$ and $r = r_1/r_2$ to obtain $||f_{r_1}||_{L_1(\mathbb{T})} \leq ||f_{r_2}||_{L_1(\mathbb{T})}$, so the family is indeed increasing and the first equality holds. Finally, since $\sup_{0<r<1} \|f_r\|_{L^1(\mathbb{T})} = \|f\|_{L^1(\mathbb{T})}$, we obtain the remaining equality. □

We are now ready to prove Theorem 6.4.

Proof. Consider the functions $\widetilde{f_M} = f/B_M$, where $B_M = \prod_{k=1}^{M} b_{\lambda_k}$. For a fixed M, $|B_M(z)|$ converges to 1 uniformly as $|z| \to 1$, so we have

$$\|\widetilde{f_M}\|_{H^p(\mathbb{T})}^p = \lim_{r\to 1} \int_{\mathbb{T}} \frac{|f(rz)|^p}{|B_M(rz)|^p}\, dm(z) = \int_{\mathbb{T}} |f(z)|^p\, dm(z) = \|f\|_{H^p(\mathbb{T})}^p.$$

Next, we define the function $\widetilde{f} = f/B$. For $0 < r < 1$, we observe

$$\|\widetilde{f}_r\|_{H^p(\mathbb{T})} = \lim_{M\to\infty} \left\| \frac{f_r}{B_{M,r}} \right\|_{L^p(\mathbb{T})} = \lim_{M\to\infty} \|\widetilde{f_{M,r}}\|_{H^p(\mathbb{T})} \leq \|\widetilde{f_M}\|_{H^p(\mathbb{T})} = \|f\|_{H^p(\mathbb{T})}.$$

It follows that $\|\widetilde{f}_r\|_{H^p(\mathbb{T})} \leq \|f\|_{H^p(\mathbb{T})}$ for all $0 < r < 1$, so Corollary 6.6 implies $\widetilde{f} \in H^p(\mathbb{T})$. In particular, the negative Fourier coefficients of $\widetilde{f}$ all equal 0 and Lemma 6.3 implies $|\widetilde{f}| \equiv |f|$ almost everywhere on $\mathbb{T}$. Finally, this implies $\widetilde{f} \in H^p(\mathbb{T})$ and $\|\widetilde{f}\|_{H^p(\mathbb{T})} = \|\widetilde{f}\|_{H^p(\mathbb{T})}$. □

6.4 Logarithmic integrability for functions in the Hardy spaces

Theorem 6.7. *Let $f \in H^p(\mathbb{T}) \setminus \{0\}$, $1 \leq p \leq \infty$. Then $\log|f| \in L^1(\mathbb{T})$. In particular, $f \neq 0$ almost everywhere on $\mathbb{T}$.*

Proof. We have $f \in H^1(\mathbb{T})$ because $H^p(\mathbb{T}) \subset H^1(\mathbb{T})$ for every p. By Theorem 6.4, there exists a convergent Blaschke product B and a function $\widetilde{f} \in H^1(\mathbb{T})$ such that $f = B\widetilde{f}$ and $\|f\|_{H_1(\mathbb{T})} = \|\widetilde{f}\|_{H_1(\mathbb{T})}$. Since $|f| = |\widetilde{f}|$ almost everywhere on $\mathbb{T}$, we note that it is sufficient to prove $\log|\widetilde{f}| \in L^1(\mathbb{T})$. Moreover, dividing by a constant, we can assume that

$$\widetilde{f}_r(0) = \widetilde{f}(0) = 1, \qquad 0 \leq r < 1.$$

Under these assumptions, we define $g_r(z) = \log|\widetilde{f}_r(z)|$ for $0 < r < 1$. Let also $C = \|f\|_{L_1(\mathbb{T})}$ and note that Lemma 6.5 implies $C \geq \|f_r\|_{L_1(\mathbb{T})}$. Consequently, we obtain

$$\int_{\mathbb{T}} e^{g_r(\xi)}\, dm(\xi) = \int_{\mathbb{T}} e^{\log|\widetilde{f}_r(\xi)|}\, dm(\xi) = \int_{\mathbb{T}} \left|\widetilde{f}_r(\xi)\right| dm(\xi) \leq C$$

for every $0 < r < 1$. By definition, the function $\widetilde{f}_r$ is nonzero on $\frac{1}{r}\mathbb{D}$, and $\log \widetilde{f}_r$ is defined and holomorphic on $\frac{1}{r}\mathbb{D}$. It follows that the function $\log|\widetilde{f}_r| = \mathrm{Re}\,(\log \widetilde{f}_r)$ is harmonic, implying

$$\int_{\mathbb{T}} g_r(\xi)\,dm(\xi) = \int_{\mathbb{T}} \log|\widetilde{f}_r(\xi)|\,dm(\xi) = \log|\widetilde{f}_r(0)| = 0.$$

We now split the function g_r into its positive and negative part:

$$(g_r)_+(z) = \max\{g_r(z), 0\} \quad \text{and} \quad (g_r)_-(z) = \max\{-g_r(z), 0\}.$$

It is straightforward to check that $g_r = (g_r)_+ - (g_r)_-$ and $|g_r| = (g_r)_+ + (g_r)_-$. Using the inequality $e^x > \max(x, 0)$ that holds for every $x \in \mathbb{R}$, we conclude

$$\begin{aligned}\int_{\mathbb{T}} |g_r(\xi)|\,dm(\xi) &= \int_{\mathbb{T}} (g_r)_+\,dm(\xi) + \int_{\mathbb{T}} (g_r)_-\,dm(\xi)\\ &= 2\int_{\mathbb{T}} (g_r)_+\,dm(\xi)\\ &\le 2\int_{\mathbb{T}} e^{g_r}\,dm(\xi) \le 2C.\end{aligned}$$

This immediately implies the inequality

$$\int_{\mathbb{T}} |\log|\widetilde{f}_r(\xi)||\,dm(\xi) \le 2C. \tag{9}$$

Finally, we are ready to bound the $L^1(\mathbb{T})$-norm of $\log|\widetilde{f}|$:

$$\begin{aligned}\int_{\mathbb{T}} |\log|\widetilde{f}(\xi)||\,dm(\xi) &= \int_{\mathbb{T}} \lim_{r\to 1} |\log|f_r(\xi)||\,dm(\xi)\\ &= \int_{\mathbb{T}} \liminf_{r\to 1} |\log|f_r(\xi)||\,dm(\xi)\\ &\le \liminf_{r\to 1} \int_{\mathbb{T}} |\log|f_r(\xi)||\,dm(\xi) \le 2C,\end{aligned}$$

where we used Fatou lemma in the first inequality. □

An immediate corollary is the following.

Corollary 6.8. Let f, g be two functions in $H^p(\mathbb{T})$, $1 \le p \le \infty$, such that $f = g$ on a set $E \subseteq \mathbb{T}$ with $m(E) > 0$. Then, $f = g$ in $H^1(\mathbb{T})$. In particular, $f(z) = g(z)$ for all $z \in \mathbb{D}$.

Proof. We have $f - g \in H^1(\mathbb{T})$, so either $f - g = 0$ or $\log|f-g| \in L^1(\mathbb{T})$. As $\log|f-g| = -\infty$ on E, the latter cannot hold. □

Corollary 6.9. Let $S\colon H^2(\mathbb{T}) \to H^2(\mathbb{T})$ be the shift operator, and let $E \subseteq H^2(\mathbb{T})$ be a set satisfying $SE \subseteq E$. Then $SE \neq E$.

Proof. Assume that $SE = E$ holds. Then $zE = E$ in $L^2(\mathbb{T})$. By Wiener theorem, we have

$$E = \chi_F L^2(\mathbb{T}) \quad \text{for a Borel subset } F \subseteq \mathbb{T}.$$

Since $1 \in L^2(\mathbb{T})$, it follows $\chi_F \in \chi_F L^2(\mathbb{T}) = E \subseteq H^2(\mathbb{T})$. However, it would then follow that $\log \chi_F \in L^1(\mathbb{T})$, leading to a contradiction. □

6.5 Inner-outer factorization

We recall that the absolute value of a finite or infinite convergent Blaschke product equals 1 almost everywhere on $\mathbb{T}$. In this section, we discuss functions with this property.

Definition 6.10. A function $\theta \in H^2$ is called an **inner function** if $|\theta| = 1$ almost everywhere on $\mathbb{T}$.

By applying Beurling theorem together with Corollary 6.9, we observe that a subspace $E \subseteq H^2(\mathbb{T})$ is invariant under the shift operator if and only if $E = \theta H^2(\mathbb{T})$ for some inner function θ. We know that Blaschke products are examples of inner functions. Next theorem gives an integral representation of functions without zeros in $\mathbb{D}$. It will be used to describe general inner functions.

Theorem 6.11. *Let $f \in H^1(\mathbb{T})$ be such that $f(z) \neq 0$ for every $z \in \mathbb{D}$. Then, there exists a real-valued measure μ of finite variation (i.e., a difference of two positive finite measures) on $\mathbb{T}$ such that*

$$f(z) = \exp\left(-\int_{\mathbb{T}} \frac{1+\bar{\xi}z}{1-\bar{\xi}z}\, d\mu(\xi)\right), \qquad z \in \mathbb{D}.$$

Moreover, if f is an inner function, then $\mu \geq 0$ and $\mu \perp m$. Conversely, any such μ generates an inner function by the above equality.

Before proving this theorem, we recall two important theorems of functional analysis.

Theorem 6.12 (Banach-Alaoglu theorem for separable Banach spaces)**.** *Let X be a separable Banach space and $\varphi_n \in X^*$ functionals satisfying $\|\varphi_n\| \leq c$ for all $n \in \mathbb{N}$. Then, there exists a subsequence $\{\varphi_{n_k}\}$ converging pointwise to some functional $\varphi \in X^*$.*

Theorem 6.13 (Riesz-Markov theorem)**.** *Let K be a compact Hausdorff space. Then, we have $(C_{\mathbb{R}}(K))^* = \mathcal{M}_{\mathbb{R}}(K)$, where $\mathcal{M}_{\mathbb{R}}(K)$ is the set of real-valued Borel measures on K with finite variation.*

We are ready to prove Theorem 6.11.

Proof. We first note that it suffices to find a measure μ satisfying the equality

$$\log f(z) = -\int_{\mathbb{T}} \frac{1+\bar{\xi}z}{1-\bar{\xi}z}\, d\mu(\xi) \quad \text{for } z \in \mathbb{D}.$$

As holomorphic functions are defined up to a constant by their real parts, it will suffice to prove the equality

$$\log |f(z)| = -\int_{\mathbb{T}} \frac{1-|z|^2}{|1-\bar{\xi}z|^2}\, d\mu(\xi) \quad \text{for } z \in \mathbb{D}. \tag{10}$$

We consider a sequence $0 < r_n < 1$ converging to 1 as $n \to \infty$ and measures

$$\mu_{r_n} = -\log |f_{r_n}|\, dm.$$

We can bound the total variation of μ_{r_n}:

$$\|\mu_{r_n}\|_{\mathcal{M}_{\mathbb{R}}(\mathbb{T})} = \|\log |f_{r_n}|\|_{L^1(\mathbb{T})} \le C < \infty,$$

where the uniform (in r) bound follows from (9). By Riesz-Markov, we can thus consider $\{\mu_{r_n}\}$ as a uniformly bounded sequence in $(C_{\mathbb{R}}(\mathbb{T}))^*$. Since $C_{\mathbb{R}}(\mathbb{T})$ is separable (the polynomials with coefficients in $\mathbb{Q}$ form a dense subset), Banach-Alaoglu theorem implies that there exists a measure $\mu \in \mathcal{M}_{\mathbb{R}}(\mathbb{T})$ and a subsequence $\{r_{n_k}\}$ of $\{r_n\}$ such that

$$\lim_{k\to\infty} \int_{\mathbb{T}} \varphi\, d\mu_{r_{n_k}} = \int_{\mathbb{T}} \varphi\, d\mu \quad \text{for all} \quad \varphi \in C_{\mathbb{R}}(\mathbb{T}). \tag{11}$$

Finally, by Homework 1, the equality

$$\log |f_{r_{n_k}}(z)| = \int_{\mathbb{T}} \frac{1-|z|^2}{|1-\bar{\xi}z|^2} \log |f_{r_{n_k}}(\xi)|\, dm(\xi)$$

holds for all k. As $k \to \infty$, we have

$$\log |f_{r_{n_k}}(z)| \to \log |f(z)| \quad \text{and} \quad \int_{\mathbb{T}} \frac{1-|z|^2}{|1-\bar{\xi}z|^2} \log |f_{r_{n_k}}(\xi)|\, dm(\xi) \to -\int_{\mathbb{T}} \frac{1-|z|^2}{|1-\bar{\xi}z|^2}\, d\mu(\xi),$$

proving the desired equality (10).

Let us now prove that if f is inner, then $\mu \ge 0$ and $\mu \perp m$. Recall that by construction μ is the weak-* limit of the measures $\mu_{r_{n_k}} = -\log |f_{r_{n_k}}|\, dm$. Since for inner functions we have $|f(z)| \le 1$ for $z \in \mathbb{D}$, the measures $\mu_{r_{n_k}}$ are nonegative. Hence, both sides (11) are nonnegative for every nonnegative $\varphi \in C(\mathbb{T})$. It follows that $\mu \ge 0$. Next, we want to prove that $\mu \perp m$. By Radon-Nikodym theorem, we have $\mu = w\, dm + \mu_s$ for some $w \in L^1(\mathbb{T})$ and

$\mu_s \perp m$. We need to check that $w = 0$ almost everywhere on $\mathbb{T}$. Note that $w \geq 0$, $\mu_s \geq 0$ because $\mu \geq 0$. Assumption $|f(\zeta)| = 1$ for almost every $\zeta \in \mathbb{T}$ implies

$$\begin{aligned} 0 &= \lim_{\substack{z=r\zeta,\\ r\to 1}} \log|f(r\zeta)| \\ &= \lim_{\substack{z=r\zeta,\\ r\to 1}} \log\left|\exp\left(-\int_{\mathbb{T}} \frac{1+\bar{\xi}z}{1-\bar{\xi}z}\, d\mu(\xi)\right)\right| \\ &= -\lim_{\substack{z=r\zeta\\ r\to 1}} \int_{\mathbb{T}} \frac{1-|z|^2}{|1-\bar{\xi}z|^2} d\mu(\xi), \end{aligned}$$

where we used the elementary formulas $\log|e^\lambda| = \operatorname{Re}\lambda$; $\operatorname{Re}\frac{1+\bar{\xi}z}{1-\bar{\xi}z} = \frac{1-|z|^2}{|1-\bar{\xi}z|^2}$. It follows that

$$\underbrace{\int_{\mathbb{T}} \frac{1-|z|^2}{|1-\bar{\xi}z|^2}\, d\mu(\xi)}_{\downarrow\; 0} \geq \underbrace{\int_{\mathbb{T}} \frac{1-|z|^2}{|1-\bar{\xi}z|^2} w(\xi) dm(\xi)}_{\downarrow\; w(\zeta) \text{ for almost every } \zeta\in\mathbb{T}} \geq 0.$$

Therefore, we have $w = 0$ almost everywhere on $\mathbb{T}$, and $\mu = \mu_s \perp m$.

Turning to the last part of the statement, we note that if

$$f(z) = \exp\left(-\int_{\mathbb{T}} \frac{1+\bar{\xi}z}{1-\bar{\xi}z}\, d\mu(\xi)\right) \quad \text{for } z \in \mathbb{D},$$

for some real measure μ of finite variation, then (10) holds for f and μ, and the general measure theory (more precisely, the differentiation theorem for Lebesgue integral) tells us that

$$\log|f(z)| = \lim_{r\to 1}\left(-\int_{\mathbb{T}} \frac{1-|rz|^2}{|1-\bar{\xi}rz|^2}\, d\mu(\xi)\right) = w_\mu(z)$$

for almost every $z \in \mathbb{T}$, where w_μ denotes the weight of the absolutely continuous part of $\mu = w_\mu\, dm + \mu_s$ in its Radon-Nikodym decomposition. In the case $\mu = \mu_s$ is purely singular, we see that $|f(z)| = 1$ almost everywhere on $\mathbb{T}$. If, moreover, $\mu \geq 0$, then $|f(z)| \leq 1$ for $z \in \mathbb{D}$ by (10), and hence f is inner. $\square$

Corollary 6.14. Every inner function θ in $\mathbb{D}$ has the form $\theta = \alpha B u$, where $\alpha \in \mathbb{T}$ is a constant, B is a Blaschke product, and u has the form

$$u(z) = \exp\left(-\int_{\mathbb{T}} \frac{1+\bar{\xi}z}{1-\bar{\xi}z} d\mu(\xi)\right), \qquad z \in \mathbb{D},$$

for some measure $\mu \geq 0$ with $\mu \perp m$. Conversely, every function of the form $\theta = \alpha B u$ with α, B, u as above is inner.

With a bit more effort, one can deduce from Theorem 6.11 the following result.

Theorem 6.15 (Inner-outer factorization). *Let $f \in H^p(\mathbb{T})$, $1 \leq p \leq \infty$. Then there exists the unique inner function θ such that $f = \theta F$, where $F \in H^p(\mathbb{T})$ is defined by*

$$F(z) = \exp\left(\int_{\mathbb{T}} \log |f(\xi)| \frac{1+\bar{\xi} z}{1-\bar{\xi} z} dm(\xi)\right), \qquad z \in \mathbb{D}.$$

Remark. The function F in the above theorem is called outer. Note that F has no zeroes in $\mathbb{D}$ and is completely determined by its absolute value $|F| = |f|$ on $\mathbb{T}$.

7 Hankel operators of finite rank

7.1 Coinvariant subspaces

According to Beurling theorem, a nonzero proper subspace $E \subset H^2$ is invariant under the shift operator $S : f \mapsto zf$ on H^2 if and only if $E = \theta H^2$ for some inner function θ. The Hilbert adjoint operator to the operator S acts by the formula

$$S^* : f \mapsto \frac{f - f(0)}{z}, \qquad f \in H^2.$$

This operator is often called the backward shift operator on H^2. In the unitary equivalent representation on the Fourier side, it acts as follows:

$$\hat{S}^* : \{x_0, x_1, x_2, \ldots\} \mapsto \{x_1, x_2, \ldots\}, \qquad \{x_k\}_{k \in \mathbb{Z}_+} \in \ell^2(\mathbb{Z}_+).$$

Note that S^* coincides with the Toeplitz operator $T_{\bar{z}}$ with symbol $\bar{z}$. As a direct consequence of the Beurling theorem and general Hilbert space theory, we see that a proper nonzero subspace $E \subset H^2$ is invariant under the backward shift operator S^* if and only it has the form $E = H^2 \ominus \theta H^2$ for some inner function θ.

Definition 7.1. The coinvariant subspace of H^2 generated by an inner function θ is the subspace $K^2_\theta = H^2 \ominus \theta H^2$.

We will prove some simple results on coinvariant subspaces of H^2.

Lemma 7.2. *Assume that $\theta = \theta_1\theta_2$ where θ_j are nonconstant inner functions. Then we have proper inclusions $\theta_j H^2 \supsetneq \theta H^2$, $j = 1, 2$. Equivalently, we have $K^2_{\theta_j} \subsetneq K^2_\theta$, $j = 1, 2$.*

Proof. We have $\theta H^2 = \theta_1\theta_2 H^2 = \theta_1(\theta_2 H^2) \subset \theta_1 H^2$ since $\theta_2 H^2$ is a subspace of H^2. The last inclusion is proper. Indeed, if $\theta H^2 = \theta_1 H^2$, then $\theta_1 = \theta g$ for some $g \in H^2$, so $\theta_1 = \theta_1\theta_2 g$, and $1 = \theta_2 g$. Let us write the inner-outer factorization for the function $g = \alpha BuF$. The Blaschke factor B does not appear because the function $1 = \theta_2 g$ has no roots in $\mathbb{D}$. Also, the outer factor is $F = 1$ because $|g| = 1$ almost everywhere on $\mathbb{T}$. Thus, we have $1 = \alpha\theta_2 u$ for a singular inner function u. However, in case u is nonconstant, we have $|u(0)| < 1$, a contradiction with $1 = \alpha\theta_2(0)u(0)$. Thus, θ_2 must be a constant. $\square$

Lemma 7.3. *If an inner function θ has nonconstant inner divisors $\theta = \theta_1 \cdots \theta_M$ for every $M \geq 1$, then* $\dim K^2_\theta = \infty$.

Proof. Take $M \geq 1$. We have $K^2_{\theta_1} \subsetneq K^2_{\theta_1\theta_2} \subsetneq \cdots \subsetneq K^2_{\theta_1\theta_2\cdots\theta_M} = K^2_\theta$. It follows that $\dim K^2_\theta \geq M-1$ for every $M \geq 1$. □

Lemma 7.4. *We have $\dim K^2_\theta < \infty$ if and only if θ is a finite Blaschke product.*

Proof. Write $\theta = Bu$, where B is some Blaschke product and u is the singular part of θ. If $B(z) = 0$ for $z \in \{\lambda_k\}_{k=1}^\infty$, it has the following divisors:

$$B_k = \frac{z - \lambda_k}{1 - \overline{\lambda_k} z}$$

Then the previous lemma imply $\dim K^2_\theta = \infty$. Next, if the set of zeros $\{\lambda_k\}_{k=1}^N$ is finite, then either $u \equiv 1$ and then θ is a finite Blaschke product (and we are done), or $u \not\equiv 1$, so that the measure μ in the representation of singular inner function is non-zero, and then B has the following divisors:

$$u_k = \exp\left(-\int_{\mathbb{T}} \frac{1+\bar{\xi}z}{1-\bar{\xi}z}\frac{d\mu}{M}\right)$$

for each M and each $1 \leq k \leq M$. We have $u = u_1 \cdots u_M$ for all M. Since u is an inner function, Lemma 7.3 implies that $\dim K_u = \infty$. By Lemma 7.2, we know $K_u \subset K_\theta$. Hence, $\dim K_\theta = \infty$. □

Our next aim is to describe coinvariant subspaces K^2_B generated by finite Blaschke products.

Theorem 7.5. *Let B be a finite Blaschke product with the sequence of zeroes $\{\lambda_k\}_{k=1}^N$ counting multiplicities. Then*

$$K_B = span\left\{\frac{1}{(1-\overline{\lambda_k}z)^{j_k}};\quad k = 1, \ldots, N,\ 1 \leq j_k \leq multiplicity\ of\ \lambda_k\right\}.$$

We will prove this result only in the case of simple zeroes (for some hints regarding the general case, see the the list of exercises). In other words, we will prove that for a finite Blaschke product B with simple zeros $\{\lambda_k\}_{k=1}^N$ we have

$$K^2_B = \operatorname{span}\left\{\frac{1}{1-\overline{\lambda_k}z}, k = 1, \ldots, N\right\}. \tag{12}$$

We will need the notion of the reproducing kernel.

Definition 7.6. The reproducing kernel of H^2 at $\lambda \in \mathbb{D}$ is the function $k_\lambda \in H^2$ such that

$$\langle f, k_\lambda \rangle_{H^2} = f(\lambda)$$

for every $f \in H^2$.

Remark. It is easy to check that the point evaluation functional $f \mapsto f(\lambda)$ is linear and continuous on H^2 for every $\lambda \in \mathbb{D}$. By the Riesz representation theorem for Hilbert spaces, this implies the existence and the uniqueness of a reproducing kernel at λ. Our immediate aim is to find an explicit formula for k_λ.

Lemma 7.7. *We have* $k_\lambda = \frac{1}{1-\overline{\lambda} z}$.

Proof. Write $f = \sum_{k\in\mathbb{Z}_+} c_k z^k$ for a function $f \in H^2$. Note that $\sum_k |c_k|^2 < \infty$. We have:

$$\frac{1}{1-\overline{\lambda} z} = \sum_{k=0}^{\infty} \overline{\lambda}^k z^k.$$

Thus,

$$\left\langle f, \frac{1}{1-\overline{\lambda} z} \right\rangle_{H^2} = \left\langle \{c_k\}, \{\overline{\lambda}^k\} \right\rangle_{\ell^2(\mathbb{Z}_+)} = \sum_{k\in\mathbb{Z}_+} c_k \overline{\left(\overline{\lambda^k}\right)} = \sum_{k\in\mathbb{Z}_+} c_k \lambda^k = f(\lambda).$$

It follows that $k_\lambda = \frac{1}{1-\overline{\lambda} z}$. □

Proof of formula (12). Let us check that K_B^2 and the subspace in the right hand side of (12) have the same orthogonal complement in H^2. By definition, we have $(K_B^2)^\perp = BH^2$. For the right hand side, we have

$$\begin{aligned}\left(\operatorname{span}\left\{\frac{1}{1-\overline{\lambda_k} z},\ k = 1,\dots,N\right\}\right)^\perp &= \left\{f \in H^2 :\ \langle f, k_{\lambda_k}\rangle = 0,\ 1 \le k \le N\right\} \\ &= \{f \in H^2 :\ f(\lambda_k) = 0,\ 1 \le k \le N\} \\ &= BH^2,\end{aligned}$$

as required. □

Lemma 7.8. *Let $\theta \in H^2$ be an inner function. Then, $K_\theta^2 = H^2 \ominus \theta H^2$ coincides on $\mathbb{T}$ with the subspace $H^2(\mathbb{T}) \cap \overline{z}\theta\overline{H^2(\mathbb{T})}$.*

Let us stress that identity $K_\theta^2 = H^2(\mathbb{T}) \cap \overline{z}\theta\overline{H^2(\mathbb{T})}$ holds only on the unit circle. To make situation clear, we discuss a particular case of Lemma 7.8 before its proof.

Example 7.9. Fix some $n \ge 1$ and consider the inner function $\theta_n = z^n$. By definition, we have

$$K_{\theta_n}^2 = H^2 \ominus \theta_n H^2 = \{f \in L^2(\mathbb{T}) :\ \widehat{f}(k) = 0,\ k < 0\} \ominus \{f \in L^2(\mathbb{T}) :\ \widehat{f}(k) = 0,\ k < n\}.$$

Hence,

$$K^2_{\theta_n} = \left\{ f \in L^2(\mathbb{T}),\ \widehat{f}(k) = 0 \text{ for all integers } k \in (-\infty, 0) \cup [n, \infty) \right\},$$

that is, $K^2_{\theta_n}$ coincides with the the set $\mathcal{P}_{n-1}$ of polynomials of degree at most $n-1$. At the same time, on the unit circle $\mathbb{T}$ we have

$$H^2(\mathbb{T}) \cap \overline{z}\theta_n \overline{H^2(\mathbb{T})} = H^2(\mathbb{T}) \cap \frac{1}{z} z^n \overline{H^2(\mathbb{T})} = H^2(\mathbb{T}) \cap z^{n-1} \overline{H^2(\mathbb{T})}$$

which is the restriction to $\mathbb{T}$ of the set $\mathcal{P}_{n-1}$ of polynomials of degree at most $n-1$.

Proof of Lemma 7.8. Write $\widetilde{K^2_\theta} = H^2(\mathbb{T}) \cap \overline{z}\theta\overline{H^2(\mathbb{T})}$. Let us show that $K^2_\theta = \widetilde{K^2_\theta}$. To prove $\widetilde{K^2_\theta} \subset K^2_\theta$, let us check that every function in $\widetilde{K^2_\theta}$ is orthogonal to θH^2. Take $h \in \widetilde{K^2_\theta}$. By construction, there exists $g \in H^2$ such that $h = \overline{z}\theta\overline{g}$ on $\mathbb{T}$. We have

$$h \perp \theta H^2 \iff \overline{z}\theta\overline{g} \perp \theta H^2 \underbrace{\iff}_{\text{since } |\theta|=1 \text{ a.e. on } \mathbb{T}} \overline{z}\,\overline{g} \perp H^2 \iff \int_{\mathbb{T}} \overline{z}\,\overline{g}\,\overline{z^j}\, dm(z) = 0, \ \ \forall j \geq 0,$$

that holds if and only if

$$\widehat{g}(-(j+1)) = 0, \ \ \forall j \geq 0.$$

The latter holds by the definition of H^2. Thus, we have shown that $\widetilde{K^2_\theta} \subset K^2_\theta$.

On the other hand, the opposite inclusion $\widetilde{K^2_\theta} \supset K^2_\theta$ holds if and only if $(\widetilde{K^2_\theta})^\perp \subset (K^2_\theta)^\perp$, or, equivalently, $(\widetilde{K^2_\theta})^\perp \subset \theta H^2$. So, take $h \in (\widetilde{K^2_\theta})^\perp$ and observe that $h \perp \overline{z}\theta\overline{H^2}$, from where we get $\overline{\theta}h \perp \overline{z}\overline{H^2}$ which yields $h \in \theta H^2$. This ends the proof. $\square$

7.2 Kronecker theorem

Theorem 7.10 (Kronecker theorem)**.** *Let H_φ be the Hankel operator with symbol $\varphi \in L^2(\mathbb{T})$. Denote by $\varphi_{st} = P_-\varphi$ the standard symbol of H_φ. The operator H_φ has finite rank if and only if $\varphi \in \overline{B}H^2$ for some finite Blaschke product B. Equivalently, H_φ has a finite rank if and only if $\varphi_{st} = \overline{z}\,\overline{g}$ for some rational function g such that*

$$g \in span\left\{ \frac{1}{(1-\overline{\lambda_k}z)^{j_k}};\ \ k = 1, \dots, N,\ 1 \leq j_k \leq \textit{multiplicity of } \lambda_k \right\},$$

where $\{\lambda_k\}_{k=1}^N$ is a finite sequence in $\mathbb{D}$.

Proof. Take arbitrary $\varphi \in L^2(\mathbb{T})$ and observe that $\operatorname{Ker} H_\varphi$ is an invariant subspace of the shift operator $S : f \mapsto zf$ on H^2:

$$H_\varphi f = 0 \iff P_-(\varphi f) = 0 \iff \varphi f \in H^2 \iff z\varphi f \in H^2.$$

By Beurling theorem, $\operatorname{Ker} H_\varphi = \theta H^2$ for an inner function θ. From Lemma 7.4, we now see that the subspace $\operatorname{Ker} H_\varphi$ has finite codimension in H^2 (equivalently, H_φ has a finite rank) if and only if θ is a finite Blaschke product. Thus, the following assertions are equivalent:

- $\operatorname{rank} H_\varphi < \infty$
- $H_\varphi(BH^2) = 0$ for some finite Blaschke product B.

The second assertion can be rewritten as $P_-(\varphi BH^2) = 0$, or equivalently,

$$\varphi BH^2 \subset H^2 \iff \varphi \in \overline{B}H^2.$$

This proves the first part of the theorem. Next, note that $\overline{B}H^2 = \overline{B}(K_B^2 \oplus BH^2) = \overline{B}K_B^2 \oplus H^2$ because the multiplication operator $f \mapsto \overline{B}f$ is unitary on $L^2(\mathbb{T})$ (in particular, it preserves orthogonality). For the standard symbol φ_{st}, we then have

$$\varphi \in \overline{B}H^2 \iff \varphi_{st} \in P_-(\overline{B}H^2) = \overline{B}K_B^2 = \overline{B}H^2 \cap \overline{zH^2} = \bar{z} \cdot \overline{(\bar{z}B\overline{H^2} \cap H^2)} = \overline{zK_B^2},$$

where we used twice Lemma 7.8. It remains to apply Theorem 7.5. □

8 Compact Hankel operators

Our aim in this section is to prove the following theorem.

Theorem 8.1 (Hartman theorem)**.** *A Hankel operator is compact if and only if it has a continuous symbol. In other words, we have $H_\varphi \in \mathcal{S}_\infty(H^2, \bar{z}\overline{H^2})$ for $\varphi \in L^2(\mathbb{T})$ if and only if $H_\varphi = H_\psi$ for some $\psi \in C(\mathbb{T})$.*

The proof is based on some function theory that is interesting in themselves.

8.1 Sarason algebra

The set of analytic functions in $\mathbb{D}$ that admit a continuous extension to the closed unit disk $\overline{\mathbb{D}}$ is called the disk algebra. We will use notation $A(\mathbb{D})$ for it.

Lemma 8.2 (Sarason lemma)**.** *For $\varphi \in C(\mathbb{T})$, we have* $\operatorname{dist}_{L^\infty}(\varphi, H^\infty) = \operatorname{dist}_{L^\infty}(\varphi, A(\mathbb{D}))$.

Remark. The space H^∞ is much larger than $A(\mathbb{D})$. In particular, $A(\mathbb{D})$ is separable while H^∞ is not. Thus, Sarason lemma says that regularity assumption $\varphi \in C(\mathbb{T})$ forces the best approximant to φ in H^∞ be much more regular than a typical function in H^∞.

Proof of Lemma 8.2. We trivially have $\operatorname{dist}(\varphi, H^\infty) \le \operatorname{dist}(\varphi, A(\mathbb{D}))$ because $A(\mathbb{D}) \subset H^\infty$. To prove the converse inequality, we take $\varphi \in C(\mathbb{T})$, $h \in H^\infty$ and define $h_r(z) = h(rz)$,

$$\varphi_r(z) = \int_{\mathbb{T}} \varphi(\xi) \frac{1-r^2}{|1-r\bar{\xi}z|^2}\, dm(\xi), \qquad z \in \mathbb{T}.$$

We have

$$\|\varphi - h\|_{L^\infty(\mathbb{T})} \overset{(\star)}{=} \lim_{r\to 1} \|(\varphi - h)_r\|_{L^\infty(\mathbb{T})} = \lim_{r\to 1} \|\varphi_r - h_r\|_{L^\infty(\mathbb{T})},$$

where in $(\star)$ we used the fact that for every $u \in L^\infty(\mathbb{T})$ its harmonic continuation

$$u(z) = \int_{\mathbb{T}} u(\xi) \frac{1-|z|^2}{|1-\bar{\xi}z|^2} dm(\xi), \qquad z \in \mathbb{D}$$

satisfies $\lim_{r\to 1} \|u_r\|_{L^\infty} = \|u\|_{L^\infty}$. In turn, this holds because $u(r\zeta) \to u(\zeta)$ for almost every $\zeta \in \mathbb{T}$ (see Homework I). Hence,

$$\begin{aligned} \|\varphi - h\|_{L^\infty(\mathbb{T})} &= \lim_{r\to 1} \|\varphi_r - \varphi + \varphi - h_r\|_{L^\infty(\mathbb{T})} \\ &\ge \lim_{r\to 1} \|\varphi - h_r\|_{L^\infty(\mathbb{T})} - \underbrace{\lim_{r\to 1} \|\varphi - \varphi_r\|_{L^\infty(\mathbb{T})}}_{\to 0} \\ &= \lim_{r\to 1} \|\varphi - h_r\|_{L^\infty(\mathbb{T})}. \end{aligned}$$

Here, $\lim_{r\to 1}\|\varphi-\varphi_r\|_{L^\infty(\mathbb{T})}=0$ because φ is continuous (see Homework I). Since $h_r\in A(\mathbb{D})$ for every $r<1$, we have

$$\lim_{r\to 1}\|\varphi-h_r\|_{L^\infty(\mathbb{T})}\geq \operatorname{dist}(\varphi, A(\mathbb{D})).$$

It follows that $\|\varphi-h\|_{L^\infty(\mathbb{T})}\geq \operatorname{dist}(\varphi, A(\mathbb{D}))$. Since $h\in H^\infty$ is arbitrary, the proof of lemma is completed. $\square$

Definition 8.3. The set $C(\mathbb{T})+H^\infty(\mathbb{T})$ is called Sarason algebra.

Theorem 8.4. *Sarason algebra is a Banach algebra with respect to the norm* $\|\cdot\|_{L^\infty(\mathbb{T})}$.

Proof. We first check that $C(\mathbb{T})+H^\infty(\mathbb{T})$ is closed in $L^\infty(\mathbb{T})$. To this end, consider the map

$$\begin{aligned} j:\ L^\infty(\mathbb{T}) &\longrightarrow L^\infty(\mathbb{T})/H^\infty(\mathbb{T})\\ f &\longmapsto j(f)=[f]=f+H^\infty(\mathbb{T}). \end{aligned}$$

Note that $C(\mathbb{T})+H^\infty(\mathbb{T})=j^{-1}([C(\mathbb{T})])$, and j is continuous. So, to prove that $C(\mathbb{T})+H^\infty(\mathbb{T})$ is closed in $L^\infty(\mathbb{T})$ we only need to check that $[C(\mathbb{T})]$ is closed in $L^\infty(\mathbb{T})/H^\infty(\mathbb{T})$. Let $\{[\varphi_n]\}\subset[C(\mathbb{T})]$ be such that $[\varphi_n]\to[u]$. Without loss of generality, $\{\varphi_n\}\subset C(\mathbb{T})$. By definition of the quotient norm, we have $\operatorname{dist}(\varphi_n-u, H^\infty)\to 0$, so there exists $\{h_n\}\subset H^\infty(\mathbb{T})$ such that $\|\varphi_n-u-h_n\|\to 0$, i.e. $\|(\varphi_n-h_n)-u\|\to 0$. In particular, $\{\varphi_n-h_n\}$ is a Cauchy sequence in $L^\infty(\mathbb{T})$, i.e., $\|(\varphi_n-h_n)-(\varphi_m-h_m)\|\to 0$, hence $\operatorname{dist}(\varphi_n-\varphi_m, H^\infty(\mathbb{T}))\to 0$. By Sarason lemma, we get $\operatorname{dist}(\varphi_n-\varphi_m, A(\mathbb{D}))\to 0$, that is, $\{\varphi_n+A(\mathbb{D})\}$ is a Cauchy sequence in $C(\mathbb{T})/A(\mathbb{D})$. The latter linear space is complete as a quotient of a Banach space and its closed subspace, therefore, $\{\varphi_n+A(\mathbb{D})\}$ converges in $C(\mathbb{T})/A(\mathbb{D})$ and hence in $L^\infty/H^\infty(\mathbb{T})$ to an element $\{\varphi+A(\mathbb{D})\}$, $\varphi\in C(\mathbb{T})$. By the uniqueness of the limit, we get $\varphi=u$ and the proof of closedness of $C(\mathbb{T})+H^\infty(\mathbb{T})$ is complete. It remains to observe that

$$\begin{aligned} C(\mathbb{T})+H^\infty(\mathbb{T}) &= \operatorname{clos}_{L^\infty}(\text{trig. polynomials})+H^\infty(\mathbb{T})\\ &= \operatorname{clos}_{L^\infty}\Big(\underbrace{\text{trig. polynomials }+H^\infty(\mathbb{T})}_{\text{subalgebra of } L^\infty(\mathbb{T})}\Big) \end{aligned}$$

It follows that $C(\mathbb{T})+H^\infty(\mathbb{T})$ is a closed subalgebra of $L^\infty(\mathbb{T})$. $\square$

8.2 Proof of Hartman theorem

Lemma 8.5. *For every* $\varphi\in L^\infty(\mathbb{T})$, *we have* $T_\varphi^*=T_{\bar\varphi}$.

Proof. For every $h_{1,2} \in H^2$, we have

$$\langle T_\varphi h_1, h_2 \rangle_{H^2} = \langle \varphi h_1, h_2 \rangle_{L^2(\mathbb{T})} = \langle h_1, \overline{\varphi} h_2 \rangle_{L^2(\mathbb{T})} = \langle h_1, T_{\overline{\varphi}} h_2 \rangle_{H^2},$$

which proves the statement. □

Lemma 8.6. *Let $K \in S_\infty(H^2, \overline{zH^2})$. Then $\|KT_{z^n}\| \to 0$ as $n \to +\infty$.*

Proof. Since $\|T_{z^n}\| \leq 1$ and every compact operator can be approximated in norm by a linear combination of rank one operators, we can assume that $K = \langle \cdot, v \rangle_{H^2} u$ for some $v \in H^2$, $u \in \overline{zH^2}$. Then

$$\|KT_{z^n}\| = \|\langle \cdot, T^*_{z^n} v \rangle_{H^2} u\| = \|T^*_{z^n} v\| \cdot \|u\|.$$

The latter tends to zero as $n \to +\infty$. Indeed, we know from Lemma 8.5 that $T^*_{z^n} = T_{\bar{z}^n}$. So, if $v = \sum_{k=0}^{\infty} \hat{v}(k) z^k$, then

$$T^*_{z^n} v = P_+ \left(\sum_{k=0}^{\infty} \hat{v}(k) z^{k-n} \right) = \sum_{k=n}^{\infty} \hat{v}(k) z^{k-n},$$

hence $\|T^*_{z^n} v\|^2 = \sum_{k=n}^{\infty} |\hat{v}(k)|^2$ by Parseval formula. It remains to use $v \in H^2$, hence $\sum_{k=n}^{\infty} |\hat{v}(k)|^2 \to 0$ as $n \to \infty$. □

Everything is ready for the proof of Hartman theorem.

Proof of Hartman theorem. Assume that $H_\varphi = H_\psi$ for some $\psi \in C(\mathbb{T})$. Then there exists a sequence of trigonometric polynomials $\{p_n\}$ such that $\|\psi - p_n\| \to 0$ as $n \to \infty$. By Kronecker theorem, H_{p_n} has finite rank for each n. It follows that H_φ can be approximated in norm by finite rank operators:

$$\|H_\varphi - H_{p_n}\| = \|H_\psi - H_{p_n}\| = \|H_{\psi - p_n}\| \leq \|\psi - p_n\|_{L^\infty(\mathbb{T})} \to 0.$$

It follows that H_φ is compact. Conversely, assume that H_φ is compact. By Nehari theorem, we might assume that $\varphi \in L^\infty(\mathbb{T})$. We will prove that

$$\operatorname{dist}_{\mathcal{B}(H^2, \overline{zH^2})}(H_\varphi, S_\infty(H^2, \overline{zH^2})) = \operatorname{dist}_{L^\infty(\mathbb{T})}(\varphi, C(\mathbb{T}) + H^\infty).$$

Since both $S_\infty(H^2, \overline{zH^2})$, $C(\mathbb{T}) + H^\infty$ are closed (see Theorem 8.4), it will imply the claim of the Hartman theorem. Observe that the inequality

$$\operatorname{dist}_{\mathcal{B}(H^2, \overline{zH^2})}(H_\varphi, S_\infty(H^2, \overline{zH^2})) \leq \operatorname{dist}_{L^\infty(\mathbb{T})}(\varphi, C(\mathbb{T}) + H^\infty)$$

is trivial: one can consider Hankel operators with symbols $\psi \in C(\mathbb{T})+H^\infty$ (they are compact by the first part of the proof) and use the estimate $\|H_{\varphi-\psi}\| \le \|\varphi-\psi\|_{L^\infty(\mathbb{T})}$. To prove the converse inequality, we take arbitrary $K \in S_\infty(H^2, \overline{zH^2})$ and estimate

$$\|H_\varphi - K\| \ge \limsup_{n\to\infty} \|(H_\varphi - K)T_{z^n}\| \ge \limsup_{n\to\infty} \|H_\varphi T_{z^n}\| - \liminf_{n\to\infty} \|KT_{z^n}\| \ge \limsup_{n\to\infty} \|H_\varphi T_{z^n}\|.$$

By definition Hankel operators, we have $H_\varphi T_{z^n} = H_{\varphi\cdot z^n}$ for each $n \ge 0$. Then, from Nehari theorem we obtain

$$\|H_{\varphi\cdot z^n}\| = \operatorname{dist}_{L^\infty}(\varphi\cdot z^n, H^\infty) = \operatorname{dist}_{L^\infty(\mathbb{T})}(\varphi, \bar{z}^n H^\infty) \ge \operatorname{dist}_{L^\infty(\mathbb{T})}(\varphi, C(\mathbb{T})+H^\infty).$$

In the last inequality we have used the fact that $\bar{z}^n H^\infty \subset C(\mathbb{T}) + H^\infty$ by Theorem 8.4 or from the direct consideration. Combining last estimates, we see that

$$\|H_\varphi - K\| \ge \operatorname{dist}_{L^\infty(\mathbb{T})}(\varphi, C(\mathbb{T})+H^\infty),$$

which completes the proof. □

9 Asymptotic behavior of Toeplitz determinants

In this section we study the asymptotic behavior of Toeplitz determinants. We fix a measure $\mu \geq 0$ with infinite support on $\mathbb{T}$ and consider the truncated $n \times n$ Toeplitz matrices

$$T_{\mu,n} = \begin{bmatrix} \hat{\mu}(0) & \hat{\mu}(-1) & \hat{\mu}(-2) & \dots & \hat{\mu}(-n+1) \\ \hat{\mu}(1) & \hat{\mu}(0) & \hat{\mu}(-1) & \dots & \hat{\mu}(-n+2) \\ \hat{\mu}(2) & \hat{\mu}(1) & \hat{\mu}(0) & \dots & \hat{\mu}(-n+3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{\mu}(n-1) & \hat{\mu}(n-2) & \hat{\mu}(n-3) & \dots & \hat{\mu}(0) \end{bmatrix}$$

generated by moments of μ. Here, as before, $\hat{\mu}(k) = \langle 1, z^k \rangle_{L^2(\mu)} = \int_{\mathbb{T}} \bar{z}^k \, d\mu$. We would like to determine the behavior of $\det T_{\mu,n}$ as $n \to \infty$.

9.1 Szegő theorem on Toeplitz determinants

Theorem 9.1. *Let $\mu = w \, dm + \mu_s$ be the Radon-Nikodym decomposition of a finite positive measure μ on $\mathbb{T}$ into the absolutely continuous and singular parts. Then*

$$\lim_{n \to +\infty} \frac{\log \det T_{\mu,n}}{n} = \int_{\mathbb{T}} \log w \, dm \tag{13}$$

where both sides are finite or equal to $-\infty$ simultaneously.

We remark that $\log w \leq \log(1 + w) \leq w$ and $w \in L^1(\mathbb{T})$ because μ is nonnegative and finite. Therefore, the integral $\int_{\mathbb{T}} \log w \, dm$ is correctly defined and

$$-\infty \leq \int_{\mathbb{T}} \log w \, dm < +\infty.$$

Szegő theorem in particular says that this integral equals $-\infty$ if and only $\frac{\log \det T_n}{n} \to -\infty$ as $n \to +\infty$. Another remark is that the statement is invariant under the scaling $\mu \mapsto \lambda\mu$, so we can (and will) assume that $\mu(\mathbb{T}) = 1$. By Jensen inequality, the right hand side in (13) will then be non-positive.

One can guess Szegő formula (13) in the following way. Suppose $\mu = |\varphi|^2 \, dm$ for some function $\varphi \in H^\infty$. Then the definition of Toeplitz operators together with assumption $\varphi \in H^\infty$ give $T_\mu = T_{|\varphi|^2} = T_{\bar{\varphi}} T_\varphi$. If φ is very nice (say, analytic in the disk $(1 + \varepsilon)\mathbb{D}$), one might hope that

$$T_{\mu,n} = T_{\bar{\varphi},n} T_{\varphi,n} + \text{ small error term,}$$

so that

$$\det T_{\mu,n} = |\det T_{\varphi,n}|^2(1+o(1)), \qquad n \to +\infty.$$

From (5) we see that the matrix $T_{\varphi,n}$ is lower-triangular, hence its determinant can be easily computed:

$$\det T_{\varphi,n} = (\hat{\varphi}(0))^n = \exp\left(n \log \hat{\varphi}(0)\right) = \exp\left(n \int_{\mathbb{T}} \log \varphi \, dm\right).$$

In the last formula we assumed that φ has no zeroes in $(1+\varepsilon)\mathbb{D}$, so that $\log \varphi$ is a correctly defined harmonic function in $(1+\varepsilon)\mathbb{D}$ and the mean value formula holds. From here, we get

$$\det T_{\mu,n} \sim |\det T_{\varphi,n}|^2 = \exp\left(n \int_{\mathbb{T}} \log |\varphi|^2 \, dm\right) = \exp\left(n \int_{\mathbb{T}} \log w \, dm\right).$$

Taking the logarithm and dividing by n, we arrive at Szegő formula (13). With some effort, the above reasoning can be made a rigorous proof that works for nice enough measures μ. However, the most general case, as stated in Theorem 9.1, requires a completely different approach.

9.2 Orthogonal polynomials

Definition 9.2. Let μ be a probability measure on $\mathbb{T}$ with infinite support. Then monomials $\{z^k\}_{k\geq 0}$ are linearly independent in $L^2(\mu)$ and therefore the sequence of monic orthogonal polynomials

$$\Phi_n = z^n + \ldots, \qquad \langle \Phi_k, \Phi_j \rangle_{L^2(\mu)} = 0 \text{ if } k \neq j$$

is defined, say, by Gram-Schmidt orthogonalization process. We also will deal with the orthonormal polynomials

$$\varphi_n = \frac{\Phi_n}{\|\Phi_n\|_{L^2(\mu)}}, \qquad n \geq 0.$$

Definition 9.3. For a polynomial p of degree at most n, we define the reflected polynomial p^* by $p^* = z^n \overline{p(1/\bar{z})}$.

Note that the operation $p \mapsto p^*$ depends on n. It will be clear from the context which n we use in each concrete case. On the unit circle, we have $p^* = z^n \overline{p(z)}$. In particular, $\langle p, q \rangle_{L^2(\mu)} = \langle q^*, p^* \rangle_{L^2(\mu)}$ for any pair of polynomials p, q of degree at most n.

Example 9.4. Let $\mu = m$ be the normalized Lebesgue measure on $\mathbb{T}$. Then $\Phi_n = \varphi_n = z^n$, and $\Phi_n^* = \varphi_n^* = 1$ for every $n \geq 0$.

Proposition 9.5. We have $\det T_{\mu,n} = \prod_0^{n-1} \|\Phi_k\|^2_{L^2(\mu)}$ for every $n \geq 1$.

Proof. Let $\mathcal{P}_n$ be the Hilbert space consisting of polynomials of degree at most $n-1$ with the inner product inherited from $L^2(\mu)$. Consider the operator $V_n : \mathcal{P}_n \to \mathcal{P}_n$ defined by $V_n : \varphi_k \mapsto z^k$. We have

$$\det T_{\mu,n} = \det[\hat{\mu}(k-j)]_{0\le k,j\le n-1} = \det[\langle V_n\varphi_j, V_n\varphi_k\rangle_{L^2(\mu)}]_{0\le k,j\le n-1}.$$

Thus, $\det T_{\mu,n} = \det V_n^* V_n = |\det V_n|^2$. At the same time, V_n acts as

$$V_n\varphi_k = z^k = \sum_{j=0}^{k} c_{j,k}\varphi_j,$$

therefore, the matrix of V_n in the orthonormal basis $\{\varphi_s\}_{s=0}^{n-1}$ of $\mathcal{P}_n$ is upper-triangular with numbers $c_{k,k}$ on the diagonal. In particular, we have

$$\det V_n = \prod_{k=0}^{n-1} c_{k,k}.$$

Comparing the coefficients in front of leading terms in the relation

$$z^k = \sum_{j=0}^{k} c_{j,k}\varphi_j$$

we see that $c_{kk} = \|\Phi_k\|_{L^2(\mu)}$. The lemma follows. □

The next result is of great importance for the theory of orthogonal polynomials.

Proposition 9.6 (Szegő recursion)**.** For some complex numbers $\{a_n\}_{n\ge 0}$ we have

$$\Phi_{n+1} = z\Phi_n - \overline{a_n}\Phi_n^*$$

for every $n \ge 0$.

Proof. By construction, the polynomials Φ_{n+1}, $z\Phi_n$ have the same degree $n+1$ and leading term z^{n+1}. It follows that $p = \Phi_{n+1} - z\Phi_n$ is a polynomial of degree at most n. Let us show that $p^* = z^n\overline{p}$ is orthogonal to all polynomials of degree at most $n-1$ (that is, p^* is a constant multiple of Φ_n), then the statement will follow by taking the reflection. Take $q \in \mathcal{P}_{n-1}$, set $q^* = z^{n-1}\overline{q}$. We have

$$\langle p^*, q\rangle_{L^2(\mu)} = \langle zq^*, p\rangle_{L^2(\mu)} = \langle zq^*, \Phi_{n+1}\rangle_{L^2(\mu)} - \langle zq^*, z\Phi_n\rangle_{L^2(\mu)}.$$

The first summand in the right hand side is zero because $zq^* \in \mathcal{P}_n$. The second summand vanishes because $\langle zq^*, z\Phi_n\rangle_{L^2(\mu)} = \langle q^*, \Phi_n\rangle_{L^2(\mu)}$ and $q^* \in \mathcal{P}_{n-1}$. $\square$

Definition 9.7. Numbers $\{a_n\}_{n\geq 0}$ in Proposition 9.6 are called the recurrence coefficients of μ.

Proposition 9.8. We have $|a_n| < 1$ for every $n \geq 0$.

Proof. Using orthogonality, we get

$$\|z\Phi_n\|^2_{L^2(\mu)} = \|\Phi_{n+1} + \overline{a_n}\Phi_n^*\|^2_{L^2(\mu)} = \|\Phi_{n+1}\|^2_{L^2(\mu)} + |a_n|^2\|\Phi_n^*\|^2_{L^2(\mu)}.$$

Since $\|\Phi_n^*\|^2_{L^2(\mu)} = \|\Phi_n\|^2_{L^2(\mu)}$ and $\|z\Phi_n\|^2_{L^2(\mu)} = \|\Phi_n\|^2_{L^2(\mu)}$, we then have

$$\|\Phi_{n+1}\|^2_{L^2(\mu)} = (1 - |a_n|^2)\|\Phi_n\|^2_{L^2(\mu)}. \tag{14}$$

Since both sides are strictly positive, we obtain $|a_n| < 1$. $\square$

Proposition 9.9. We have $\|\Phi_k\|^2_{L^2(\mu)} = \prod_{j=0}^{k-1}(1 - |a_j|^2)$ for every $k \geq 0$, where the empty product (for $k = 0$) is defined as 1.

Proof. The claim follows by induction from (14). $\square$

Next result allows us to rewrite Szegő theorem in terms of the recurrence coefficients of orthogonal polynomials.

Theorem 9.10. *Under assumptions of Theorem 9.1, we have*

$$\lim_{n\to+\infty} \frac{\log\det T_{\mu,n}}{n} = \log\prod_{k=0}^{\infty}(1 - |a_k|^2),$$

where both sides are finite or equal $-\infty$ simultaneously.

Proof. We already know that

$$\frac{\log\det T_{\mu,n}}{n} = \frac{\log\prod_0^{n-1}\|\Phi_k\|^2_{L^2(\mu)}}{n} = \frac{\log\prod_{k=0}^{n-1}\prod_{j=0}^{k-1}(1 - |a_j|^2)}{n}.$$

In other words,

$$\frac{\log\det T_{\mu,n}}{n} = \frac{n-1}{n}\log(1 - |a_0|^2) + \frac{n-2}{n}\log(1 - |a_1|^2) + \ldots + \frac{1}{n}\log(1 - |a_{n-2}|^2).$$

Then, since all terms are non-positive, we have

$$\liminf_{n\to\infty} \frac{\log\det T_{\mu,n}}{n} \geq \liminf_{n\to\infty} \Big(\log(1-|a_0|^2) + \ldots + \log(1-|a_{n-2}|^2) \Big) = \sum_0^\infty \log(1-|a_k|^2),$$

and at the same time for a fixed N we have

$$\begin{aligned} \limsup_{n\to\infty} \frac{\log\det T_{\mu,n}}{n} &\leq \limsup_{n\to\infty} \frac{n-1}{n} \log(1-|a_0|^2) + \ldots + \frac{n-N}{n} \log(1-|a_{N-1}|^2) \\ &\leq \sum_0^{N-1} \log(1-|a_k|^2), \quad \text{for every} \;\; N \geq 1, \end{aligned}$$

from where we see that there exists the limit

$$\lim_{n\to\infty} \frac{\log\det T_{\mu,n}}{n} = \sum_0^\infty \log(1-|a_k|^2) = \log \prod_{k=0}^\infty (1-|a_k|^2),$$

as claimed. □

Lemma 9.11. *We have* $(1-|z|^2)\sum_{k=0}^n |\varphi_k(z)|^2 = |\varphi_{n+1}^*(z)|^2 - |\varphi_{n+1}(z)|^2$ *for every* $z \in \mathbb{D}$. *The polynomials* φ_n^* *do not vanish in* $\overline{\mathbb{D}}$.

Proof. For $n = 0$, we have

$$(1-|z|^2)\sum_{k=0}^n |\varphi_k(z)|^2 = 1 - |z|^2$$

and

$$|\varphi_{n+1}^*(z)|^2 - |\varphi_{n+1}(z)|^2 = \frac{|1-a_0 z|^2 - |z-\bar{a}_0|^2}{1-|a_0|^2} = 1 - |z|^2.$$

So, for $n = 0$ we have $(1-|z|^2)\sum_{k=0}^n |\varphi_k(z)|^2 = |\varphi_{n+1}^*(z)|^2 - |\varphi_{n+1}(z)|^2$. Next, suppose that this formula holds for n, and let us check that it holds for $n+1$. This is equivalent to the claim

$$|\varphi_{n+1}^*(z)|^2 - |\varphi_{n+1}(z)|^2 + (1-|z|^2)|\varphi_{n+1}(z)|^2 = |\varphi_{n+2}^*(z)|^2 - |\varphi_{n+2}(z)|^2,$$

that can be rewritten in the form

$$|\varphi_{n+1}^*(z)|^2 - |z\varphi_{n+1}(z)|^2 = |\varphi_{n+2}^*(z)|^2 - |\varphi_{n+2}(z)|^2. \tag{15}$$

Writing the recursion in the form

$$\sqrt{1-|a_{n+1}|^2}\varphi_{n+2}(z) = z\varphi_{n+1}(z) - \bar{a}_{n+1}\varphi^*_{n+1}(z)$$
$$\sqrt{1-|a_{n+1}|^2}\varphi^*_{n+2}(z) = \varphi^*_{n+1}(z) - a_{n+1}z\varphi_{n+1}(z),$$

and substituting these expressions into the right hand side of (15), we see that (15) indeed holds true. So, it remains to show that the polynomial φ^*_{n+1} does not vanish in $\overline{\mathbb{D}}$, $n \geq 0$. From the identity

$$(1-|z|^2)\sum_{k=0}^{n}|\varphi_k(z)|^2 = |\varphi^*_{n+1}(z)|^2 - |\varphi_{n+1}(z)|^2$$

we see that $|\varphi^*_{n+1}(z)|^2 \neq 0$ for $z \in \mathbb{D}$. Looking at the same identity at a possible node $z_0 \in \mathbb{T}$ of φ^*_{n+1}, we see that the left hand side behaves as $\gtrsim |z-z_0|$ as $z \to z_0$ in $\mathbb{D}$, while the right hand side is $O(|z-z_0|^2)$ under the same regime $z \to z_0$. So, φ^*_{n+1} does not vanishes on $\mathbb{T}$ as well. □

9.3 Schur algorithm

To proceed further towards the proof of Szegő theorem, we will need a piece of function theory. The Schur class $\mathcal{S}(\mathbb{D})$ in the open unit disk $\mathbb{D}$ consists of analytic functions f in $\mathbb{D}$ such that

$$\sup_{z\in\mathbb{D}}|f(z)| \leq 1.$$

For $f \in \mathcal{S}(\mathbb{D})$, we write $f \in \mathcal{S}_*(\mathbb{D})$ if f is not a finite Blaschke product. Take $f \in \mathcal{S}_*(\mathbb{D})$, set $f_0 = f$, and define the sequence $\{f_n\}_{n\geq 0}$ using Schur algorithm:

$$zf_{n+1} = \frac{f_n - f_n(0)}{1-\overline{f_n(0)}f_n}, \qquad n \geq 0. \tag{16}$$

The correctness of these procedure is justified in the lemma below.

Lemma 9.12. *Let $f \in \mathcal{S}_*(\mathbb{D})$. Then for every $n \in \mathbb{Z}_+$, we have $f_n(0) \in \mathbb{D}$ and $f_n \in \mathcal{S}_*(\mathbb{D})$. In particular, the fraction $\frac{f_n-f_n(0)}{1-\overline{f_n(0)}f_n}$ is correctly defined and belongs to $\mathcal{S}_*(\mathbb{D})$ for every $n \geq 0$.*

Proof. The proof is based on induction. We have $f_0(0) \in \mathbb{D}$ by the maximum modulus principle for nonconstant analytic functions. Then, the function $g_1 = \frac{f_0-f_0(0)}{1-\overline{f_0(0)}f_0}$ is correctly defined and analytic in $\mathbb{D}$, it maps $\mathbb{D}$ into $\mathbb{D}$. Since f_0 is not a finite Blaschke product, the function g_1 is also not a finite Blaschke product. So, $g_1 \in \mathcal{S}_*(\mathbb{D})$. Moreover, we have $g_1(0) = 0$ by construction. Then by Schwarz lemma for bounded analytic functions in the open unit disk, we have $g_1 = zf_1$ for some $f_1 \in \mathcal{S}_*(\mathbb{D})$. We thus have checked the inductive transition step. $\square$

Remark. In the case where $f \in \mathcal{S}(\mathbb{D}) \setminus \mathcal{S}_*(\mathbb{D})$ is a Blaschke product of order $N \geq 0$, the same Schur algorithm determines the finite sequence of Blaschke products $f_0,\ f_1,\ \dots f_N$ of orders $N,\ N-1,\ \dots 0$, correspondingly. In particular, f_N is a constant of unit modulus and the Schur algorithm stops to avoid division by zero.

Definition 9.13. Let $f \in \mathcal{S}_*(\mathbb{D})$. The sequence $\{f_n(0)\}_{n\geq 0} \subset \mathbb{D}$ is called the sequence of Schur coefficients of f.

Lemma 9.14. *Let $f, g \in \mathcal{S}_*(\mathbb{D})$ be such that $f_n(0) = g_n(0)$ for all $0 \leq n \leq N$. Then f, g have the same Tailor coefficients at the origin of order $0 \leq n \leq N$. Moreover, we have $|f(z) - g(z)| \leq 2|z|^{N+1}$ for all $z \in \mathbb{D}$.*

Proof. Let us first check that $|f(z) - g(z)| = O(|z|^{N+1})$ in $\mathbb{D}$. The proof uses induction. Note that for every $j \geq 0$ the Schur coefficients of the function f_j are exactly $\{f_{j+n}(0)\}_{n\geq 0}$. Then, since $f_N(0) = g_N(0)$, we have

$$|f_N - g_N| = O(|z|)$$

which corresponds the step $k = 0$ for the induction claim

$$|f_{N-k} - g_{N-k}| = O(|z|^{k+1}), \qquad 0 \leq k \leq N.$$

Assuming induction claim holds for some k, we write

$$f_{N-k-1} - g_{N-k-1} = \frac{zf_{N-k} + f_{N-k-1}(0)}{1 + zf_{N-k}\overline{f_{N-k-1}(0)}} - \frac{zg_{N-k} + g_{N-k-1}(0)}{1 + zg_{N-k}\overline{g_{N-k-1}(0)}}.$$

Since $f_{N-k-1}(0) = g_{N-k-1}(0)$ by the assumption of the lemma, the right hand side in the above formula is of order $O(|z| \cdot |f_{N-k} - g_{N-k}|) = O(|z|^{k+2})$ by computing the difference and identifying its asymptotics at zero. Thus, we have shown that $|f(z) - g(z)| = O(|z|^{N+1})$ in $\mathbb{D}$. Since both $|f|$, $|g|$ are bounded by 1 in $\mathbb{D}$, the Schwarz lemma argument tells us that

$$|f(z) - g(z)| \leq \sup_{z\in\mathbb{D}} |f(z) - g(z)| \cdot |z|^{N+1} \leq 2|z|^{N+1},$$

as required. $\square$

Lemma 9.15. *The mapping $\Phi : f \mapsto \{f_n(0)\}_{n\ge 0}$ is a homeomorphism from $\mathcal{S}_*(\mathbb{D})$ with the topology of convergence on compact subsets of $\mathbb{D}$ onto the space $\mathbb{D}^\infty$ of sequences $q : \mathbb{Z}_+ \to \mathbb{D}$ with the topology of elementwise convergence.*

Proof. Let us first show that $\Phi(\mathcal{S}_*(\mathbb{D})) = \mathbb{D}^\infty$. By construction, $\Phi(\mathcal{S}_*(\mathbb{D})) \subset \mathbb{D}^\infty$. Next, take a sequence $\{a_k\}_{k=0}^\infty \subset \mathbb{D}^\infty$ such that $a_n = 0$ for $n \ge N$. Put $f_N = 0$ and define inductively

$$f_{N-k-1} = \frac{zf_{N-k} + a_{N-k-1}}{1 + zf_{N-k}\overline{a_{N-k-1}}}, \qquad 0 \le k \le N.$$

Then, set $f = f_0$. It is straightforward to check that $a_k = f_k(0)$. In particular, we have constructed $f \in \mathcal{S}_*(\mathbb{D})$ with a given set of Schur coefficients $\{a_k\}_{k=0}^\infty \subset \mathbb{D}^\infty$ provided $a_n = 0$ for $n \ge N$. Then, if $(f)_n$ are functions in $\mathcal{S}_*(\mathbb{D})$ with Schur coefficients $\{\chi_{[0,n]}(k)a_k\}$, we see from Lemma 9.14 that

$$|(f)_{j_1}(z) - (f)_{j_2}(z)| \le 2|z|^{\min(j_1,j_2)+1}, \qquad z \in \mathbb{D}.$$

Hence, $\{(f)_j(z)\}_{j\ge 0}$ is a Cauchy sequence on any compact subset of $\mathbb{D}$, i.e., it converges in the topology $\mathcal{S}_*(\mathbb{D})$ to some function f. One can easily check that $f_k(0) = a_k$ for every $k \in \mathbb{D}$. Thus, $\Phi(\mathcal{S}_*(\mathbb{D})) \supset \mathbb{D}^\infty$. Lemma 9.14 implies that Φ is a bijection. Verification of continuity of Φ, Φ^{-1} is left to the reader. □

The following lemma is a piece of the theory of Hardy spaces. We take it for granted.

Lemma 9.16. *Let $g \in \mathcal{S}_*(\mathbb{D})$. Then $1 + g$ is an outer function. In particular, we have*

$$\log|1 + g(0)| = \int_{\mathbb{T}} \log|1 + g(\xi)|\, dm(\xi).$$

where the integral converges absolutely.

Lemma 9.17. *Let $f \in \mathcal{S}_*(\mathbb{D})$ be a function with finitely many nonzero recurrence coefficients. Then*

$$\int_{\mathbb{T}} \log(1 - |f(\xi)|^2)\, dm(\xi) = \log \prod_{n=0}^{\infty} (1 - |f_n(0)|^2). \tag{17}$$

Proof. The proof is based on the following algebraic relation

$$1 - |zf_0|^2 = \frac{(1 - |f_0(z)|^2)(1 - |f_1(z)|^2)}{|1 - \overline{f_0(0)}f_1(z)|^2}$$

which follows from the Schur algorithm by a direct computation. Integrating this relation over $\mathbb{T}$ and using Lemma 9.16, we get

$$\int_{\mathbb{T}} \log(1-|f_0|^2)\, dm = \log(1-|f_0(z)|^2) + \int_{\mathbb{T}} \log(1-|f_1|^2)\, dm.$$

Then, we can iterate this equation $N-1$ times to obtain

$$\int_{\mathbb{T}} \log(1-|f(\xi)|^2)\, dm(\xi) = \log \prod_{n=0}^{N-1} (1-|f_n(0)|^2) + \int_{\mathbb{T}} \log(1-|f_N|^2)\, dm.$$

By our assumption, there is a number N such that $f_N = 0$, $f_n(0) = 0$, $n \geq N$. This yields formula (17). □

9.4 Proof of Szegő theorem

The proof of Theorem 9.1 will be based on the following beautiful result.

Theorem 9.18 (Geronimus theorem). *Let μ be a probability measure on $\mathbb{T}$ with infinite support and let f be the function defined by*

$$\frac{1+zf(z)}{1-zf(z)} = \int_{\mathbb{T}} \frac{1+\bar{\xi}z}{1-\bar{\xi}z}\, d\mu(\xi), \qquad z \in \mathbb{D}.$$

Then $f \in \mathcal{S}_*(\mathbb{D})$ and Schur coefficients of f coincide with recurrence coefficients of μ:

$$f_n(0) = a_n \tag{18}$$

for all $n \geq 0$.

There is no "simple explanation" for the formula (18). Its proof, being an elementary computation based on Szegő recursion, occupies about 2 pages. We skip it and derive Szegő theorem from Geronimus formula (18).

Proof of Theorem 9.1. We already reduced Theorem 9.1 to the following statement:

$$\int_{\mathbb{T}} \log w\, dm = \log \prod_{k=0}^{\infty} (1-|a_k|^2),$$

see Theorem 9.10. For any $n \geq 0$, we have

$$0 = \log 1 = \log\left(\int_{\mathbb{T}} |\varphi_n|^2\, d\mu\right) \geq \log\left(\int_{\mathbb{T}} |\varphi_n|^2 w\, dm\right).$$

By Jensen inequality,

$$\log\left(\int_{\mathbb{T}} |\varphi_n|^2 w\, dm\right) \ge \int_{\mathbb{T}} \log|\varphi_n|^2 dm + \int_{\mathbb{T}} \log w\, dm.$$

Recall that $|\varphi_n| = |\varphi_n^*|$ on $\mathbb{T}$ and φ_n^* has no zeroes in $\overline{\mathbb{D}}$ by Lemma 9.11. Hence,

$$\int_{\mathbb{T}} \log|\varphi_n|^2 dm = \int_{\mathbb{T}} \log|\varphi_n^*|^2 dm = \log|\varphi_n^*(0)|^2.$$

It follows that $-\log|\varphi_n^*(0)|^2 \ge \int_{\mathbb{T}} \log w\, dm$. At the same time,

$$\log|\varphi_n^*(0)|^2 = -\log\|\Phi_n\|^2_{L^2(\mu)} = -\log\prod_0^{n-1}(1-|a_k|^2), \qquad n \ge 0,$$

from where we get

$$-\int_{\mathbb{T}} \log w\, dm \ge -\log\prod_0^{\infty}(1-|a_k|^2).$$

To prove the converse inequality, we observe that Lemma 9.16 implies

$$-\int_{\mathbb{T}} \log(1-|f|^2)\, dm = -\int_{\mathbb{T}} \log\frac{1-|f(z)|^2}{|1-zf(z)|^2}\, dm = -\int_{\mathbb{T}} \log w\, dm. \tag{19}$$

The last identity here follows from the relation

$$\begin{aligned} w(z) &= \lim_{r\to 1}\int_{\mathbb{T}} \frac{1-|rz|^2}{|1-rz\bar{\xi}|^2}\, d\mu(\xi) \\ &= \lim_{r\to 1}\operatorname{Re}\left(\int_{\mathbb{T}} \frac{1+\bar{\xi}rz}{1-\bar{\xi}rz}\, d\mu(\xi)\right) \\ &= \lim_{r\to 1}\operatorname{Re}\left(\frac{1+rzf(rz)}{1-rzf(rz)}\right) \\ &= \lim_{r\to 1}\frac{1-|rzf(rz)|^2}{|1-rzf(rz)|^2} \\ &= \frac{1-|f(z)|^2}{|1-zf(z)|^2} \end{aligned}$$

that holds for almost every $z \in \mathbb{T}$. Denote by $(f)_j$ the function in $\mathcal{S}_*(\mathbb{D})$ with Schur coefficients $\{\chi_{[0,j]}(k) f_k(0)\}_{k=1}^{\infty}$. Using Fatou lemma, Beppo Levy theorem, Lemma 9.14, monotonicity of the function $r \mapsto \|g(r\cdot)\|_{L^2(\mathbb{T})}$ for $g \in H^2$, the nonlinear Parseval formula (17), and Geronimus theorem, we obtain from (19)

$$
\begin{aligned}
-\int_{\mathbb{T}} \log w \, dm &= \int_{\mathbb{T}} \liminf_{r\to 1} (-\log(1-|f(rz)|^2)) \, dm \\
&\leq \liminf_{r\to 1} \int_{\mathbb{T}} (-\log(1-|f(rz)|^2)) \, dm \\
&\leq \liminf_{r\to 1} \lim_{N\to\infty} \int_{\mathbb{T}} \sum_{0}^{N} \frac{|f(rz)|^{2n}}{n} \, dm \\
&\leq \liminf_{r\to 1} \lim_{N\to\infty} \lim_{j\to\infty} \int_{\mathbb{T}} \sum_{0}^{N} \frac{|(f)_j(rz)|^{2n}}{n} \, dm \\
&\leq \liminf_{r\to 1} \lim_{N\to\infty} \lim_{j\to\infty} \int_{\mathbb{T}} \sum_{0}^{N} \frac{|(f)_j(z)|^{2n}}{n} \, dm \\
&\leq \liminf_{r\to 1} \lim_{N\to\infty} \lim_{j\to\infty} \int_{\mathbb{T}} \sum_{0}^{\infty} \frac{|(f)_j(z)|^{2n}}{n} \, dm \\
&= \lim_{j\to\infty} \int_{\mathbb{T}} (-\log(1-|(f)_j(z)|^2)) \, dm \\
&= -\lim_{j\to\infty} \log \prod_{k=0}^{j} (1-|f_k(0)|^2) \\
&= -\log \prod_{k=0}^{\infty} (1-|f_k(0)|^2), \\
&= -\log \prod_{k=0}^{\infty} (1-|a_k|^2).
\end{aligned}
$$

To je to! □

10 Homeworks

Solutions of problems below should be discussed with lecturer.

10.1 Homework I

In the course, we use basic theory of Hardy spaces. This list of problems is devoted to the proof of the embedding $H^p(\mathbb{T}) \subset H^p(\mathbb{D})$ and the fact that every function in $H^p(\mathbb{T})$ has radial boundary values almost everywhere on $\mathbb{T}$ when considered as an element of $H^p(\mathbb{D})$. In reality, $H^p(\mathbb{T}) = H^p(\mathbb{D})$, while we prove just one inclusion here. Recall that

$$\int_{\mathbb{T}} \varphi(\xi)\, dm(\xi) = \frac{1}{2\pi} \int_0^{2\pi} \varphi(e^{it})\, dt$$

for every function $\varphi \in L^1(\mathbb{T})$. A continuous function u in a domain $\Omega \subset \mathbb{C}$ is called harmonic if it satisfies the mean value property

$$u(z) = \frac{1}{2\pi} \int_0^{2\pi} u(z + re^{it})\, dt$$

for any point $z \in \Omega$ and radius $r > 0$ such that the closed disk $B[z,r]$ is contained in Ω. In the case where Ω is simply connected, this property is equivalent to $u = \operatorname{Re} f$ for some analytic function f in Ω.

1. Prove that the function $z \mapsto \frac{1-|z|^2}{|1-\bar{\xi}z|^2}$ is harmonic in $\mathbb{D}$ for every $\xi \in \mathbb{T}$.

2. Prove that the function $U : z \mapsto \int_{\mathbb{T}} u(\xi) \frac{1-|z|^2}{|1-\bar{\xi}z|^2}\, dm(\xi)$ is harmonic in $\mathbb{D}$ if $u \in C(\mathbb{T})$.

In both cases above it is instructive to search for an analytic function f to prove harmonicity from analyticity.

3. Prove that

$$\int_{\mathbb{T}} \frac{1-|z|^2}{|1-\bar{\xi}z|^2}\, dm(\xi) = 1$$

for every $z \in \mathbb{D}$.

4. In the setting of the previous problem, prove that

$$\lim_{r\to 1} U(r\zeta) = u(\zeta)$$

for every $\zeta \in \mathbb{T}$.

The above problems can be summarized as follows: the operator of taking boundary values is a continuous bijection from the set $HM(\overline{\mathbb{D}})$ of harmonic functions in $\mathbb{D}$ continuous up to the boundary, to the set $C(\mathbb{T})$ of all continuous functions on $\mathbb{T}$. Morever, harmonic functions in $\mathbb{D}$ and their boundary values on $\mathbb{T}$ are related by the Poisson formula

$$U(z) = \int_{\mathbb{T}} u(\xi) \frac{1 - |z|^2}{|1 - \bar{\xi} z|^2} \, dm(\xi), \qquad z \in \mathbb{D}.$$

The identity $H^p(\mathbb{T}) = H^p(\mathbb{D})$ is the statement of the same sort. To prove it, we will need the notion of the Lebesgue point. A point $\zeta \in \mathbb{T}$ is a Lebesgue point for a function $f \in L^1(\mathbb{T})$ if

$$\lim_{\varepsilon \to 0} \frac{1}{m(B(\zeta, \varepsilon) \cap \mathbb{T})} \int_{B(\zeta,\varepsilon)\cap\mathbb{T}} |f(\xi) - f(\zeta)| \, dm(\xi) = 0.$$

It is known that for every $f \in L^1(\mathbb{T})$, almost every point of $\mathbb{T}$ is the Lebesgue point. We will use this fact in the proofs below.

5. Let $\zeta \in \mathbb{T}$ be a Lebesgue point for a function $u \in L^1(\mathbb{T})$. Prove that

$$\lim_{r \to 1} U(r\zeta) = u(\zeta), \qquad U(z) = \int_{\mathbb{T}} u(\xi) \frac{1 - |z|^2}{|1 - \bar{\xi} z|^2} \, dm(\xi).$$

In the previous problem it is instructive to approximate the Poisson kernel $\frac{1-|z|^2}{|1-\bar{\xi}z|^2}$ with a finite convex combination of kernels $\frac{1}{m(B(\zeta,\varepsilon)\cap\mathbb{T})}\chi_{B(\zeta,\varepsilon)\cap\mathbb{T}}$, where χ_B is the indicator function of a subset B of the unit circle.

Turning to the Hardy spaces, we continue as follows:

6. Prove that $H^1(\mathbb{T}) \supset H^p(\mathbb{T})$ for all $p \geq 1$ using Cauchy-Schwarz inequality.

7. Prove that for every $f \in H^1(\mathbb{T})$, if we define

$$F(z) = \sum_{n \geq 0} \hat{f}(n) z^n, \qquad \hat{f}(n) = \int_{\mathbb{T}} f(\xi) \overline{\xi}^n \, dm(\xi),$$

then the series converges and defines an analytic function in $\mathbb{D}$.

8. Prove that

$$F(z) = \int_{\mathbb{T}} f(\xi) \frac{1}{1 - \bar{\xi} z} \, dm(\xi) = \int_{\mathbb{T}} f(\xi) \frac{1 - |z|^2}{|1 - \bar{\xi} z|^2} \, dm(\xi), \qquad z \in \mathbb{D},$$

in the setting from previous problem.

9. Using basic estimate for integrals (absolute value of an integral is not smaller than the integral of the absolute value), prove that $F \in H^1(\mathbb{D})$ and $\|F\|_{H^1(\mathbb{D})} \le \|f\|_{H^1(\mathbb{T})}$ for every $f \in H^1(\mathbb{T})$.

10. Using Jensen inequality for integrals, prove that $F \in H^p(\mathbb{D})$ and $\|F\|_{H^p(\mathbb{D})} \le \|f\|_{H^p(\mathbb{T})}$ for every $f \in H^p(\mathbb{T})$, $p \ge 1$.

11. Prove that the mapping $f \mapsto F$ determines a continuous embedding $H^p(\mathbb{T}) \subset H^p(\mathbb{D})$ for every $p \ge 1$. Moreover, for almost every point $\zeta \in \mathbb{T}$ we have $F(r\zeta) \to f(\zeta)$ as $r \to 1$.

10.2 Homework II

This set of problems is mainly devoted to finite rank and compact Hankel operators. In particular, some excercises concern results that were considered as "obvious" during the proof of Kronecker theorem.

Recall that a bounded operator $T : H_1 \to H_2$ between two Hilbert spaces H_1, H_2 has rank n if $\dim \operatorname{Ran} T = n$.

1. Given a pair of separable infinitely dimensional Hilbert spaces H_1, H_2, prove that there exists a unitary operator $U : H_1 \to H_2$.

2. Prove that $\dim \operatorname{Ran} T = \dim(H_1 \ominus \operatorname{Ker} T)$ for every operator $T : H_1 \to H_2$ of finite rank. Here, as usual, $H_1 \ominus \operatorname{Ker} T$ denotes the set of vectors in H_1 that are orthogonal to $\operatorname{Ker} T$.

3. Prove that operators $T : H_1 \to H_2$ of rank n are precisely those that can be represented in the form $T = \sum_{k=1}^{n} c_k \cdot v_k \otimes u_k$ for some orthonormal sequences $\{u_k\}_1^n \subset H_1$, $\{v_k\}_1^n \subset H_2$ and some non-zero constants $c_k \in \mathbb{C} \setminus \{0\}$. Here $(v \otimes u)h = (h, u)_{H_1} v$.

Let K_B^2 be the subspace of the Hardy space H^2 of functions that are orthogonal to BH^2.

4. Using Tailor series decomposition, prove that for every $\lambda \in \mathbb{D}$ the functional $f \mapsto f'(\lambda)$ is continuous on H^2. Prove that there is a vector $\partial k_\lambda \in H^2$ such that

$$f'(\lambda) = (f, \partial k_\lambda)$$

for every $f \in H^2$.

5. Find a formula for the vector ∂k_λ. Make sure that ∂k_λ a rational function.

6. Similarly, find a formula for the vector $\partial^n k_\lambda$ such that $(f, \partial^n k_\lambda)$ is the n-th derivative of any function $f \in H^2$ at a point $\lambda \in \mathbb{D}$.

7. For which Blaschke products B we have $\partial^n k_\lambda \in K_B$?

8. Prove that rank $\dim H_\varphi \le n$ if and only if $\varphi \in \overline{B}H^2$ for some Blashke product B of degree at most n (here, the degree of a Blaschke product is the number of its zeroes counting multiplicities).

Remaining problems concerns best approximation by analytic functions.

9. Using a compactness argument (Montel theorem for analytic functions and/or Banach-Alaoglu theorem for $L^1(\mathbb{T})^* = L^\infty(\mathbb{T})$ Banach space), prove that for every function $\varphi \in L^\infty(\mathbb{T})$ there is at least one function $f \in H^\infty$ such that

$$\|\varphi - f\|_{L^\infty(\mathbb{T})} = \operatorname{dist}_{L^\infty(\mathbb{T})}(\varphi, H^\infty).$$

10. Prove that for every compact operator $K : H_1 \to H_2$ there are vectors $h_1 \in H_1$, $h_2 \in H_2$ of unit norm such that

$$\|K\| = (Kh_1, h_2).$$

11. Prove that for every $\varphi \in C(\mathbb{T})$ there is a function $f \in H^\infty$ and a function h from the unit ball of zH^1 such that

$$\|\varphi - f\|_{L^\infty(\mathbb{T})} = \operatorname{dist}_{L^\infty(\mathbb{T})}(\varphi, H^\infty) = \int_{\mathbb{T}} \varphi h \, dm = \int_{\mathbb{T}} (\varphi - f) h \, dm.$$

12. Prove that the same h works for any $f \in H^\infty$ such that $\|\varphi - f\|_{L^\infty(\mathbb{T})} = \operatorname{dist}_{L^\infty(\mathbb{T})}(\varphi, H^\infty)$.

13. Prove that $(\varphi - f)h > 0$ almost everywhere on $\mathbb{T}$ for φ, f, h as above.

14. Prove that $\varphi - f = \operatorname{dist}_{L^\infty(\mathbb{T})}(\varphi, H^\infty)/h$ almost everywhere on $\mathbb{T}$ (explain why $h \neq 0$ almost everywhere on $\mathbb{T}$).

Previous considerations lead to the following useful theorem.

15. For every function $\varphi \in C(\mathbb{T})$ there exists a **unique** function $f \in H^\infty$ such that

$$\|\varphi - f\|_{L^\infty(\mathbb{T})} = \operatorname{dist}_{L^\infty(\mathbb{T})}(\varphi, H^\infty).$$

Moreover, this function has to be continuous: $f \in H^\infty \cap C(\mathbb{T})$.

Credits

Credits

Dedicated to my first students in Slovenia:

- Eske Bockelmann
- Domen Gradišek
- Ahmad Hussein
- Jianan Peng
- Miloš Puđa
- Jakob Snoj
- Yordan Toshev

who helped me to write this text.

Licence

Last version & source

The last version and the source code of these lecture notes can be found here:
https://www.overleaf.com/read/vqyddmshwnhg#7d053e