Foundations in Signal Processing, Communications and Networking 4

Advanced Topics in System and Signal Theory

A Mathematical Approach

Bearbeitet von
Volker Pohl, Holger Boche

1. Auflage 2009. Buch. vIII, 241 S. Hardcover
ISBN 978 3 642 03638 5
Format (B x L): 15,5 x 23,5 cm
Gewicht: 1170 g

2

Fourier Analysis and Analytic Functions

2.1 Trigonometric Series

One of the most important tools for the investigation of linear systems is Fourier analysis. Let $f \in L^1$ be a complex-valued Lebesgue-integrable function on the unit circle $\mathbb{T}$. Then the *Fourier coefficients* of f are the complex numbers

$$\hat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\, \mathrm{e}^{-\mathrm{i}n\theta}\, \mathrm{d}\theta\,, \qquad n = 0, \pm 1, \pm 2, \dots\,. \tag{2.1}$$

Let $\mathcal{B}$ be an arbitrary subspace of L^1. Then we define by

$$\mathcal{B}_+ := \{\, f \in \mathcal{B} : \hat{f}(n) = 0 \text{ for all } n < 0 \,\}$$

the subspace of all functions in $\mathcal{B}$ for which all Fourier coefficients with negative index are equal to zero.

Definition 2.1 (Conjugate function). *Let $\mathcal{B} \subset L^1$ and $f \in \mathcal{B}$. The function $\widetilde{f} \in \mathcal{B}$ is called the* conjugate function of f in $\mathcal{B}$ *if it satisfies the conditions*

$$f + \mathrm{i}\widetilde{f} \in \mathcal{B}_+ \qquad \textit{and} \qquad \tfrac{1}{2\pi} \textstyle\int_{-\pi}^{\pi} \widetilde{f}(\mathrm{e}^{\mathrm{i}\theta})\, \mathrm{d}\theta = 0\,.$$

In general, a function $f \in \mathcal{B}$ need not possess a conjugate function according to the above definition. However, for every $f \in \mathcal{B} \subset L^1$ it is always possible to define a conjugate functions $\widetilde{f}$ which exists almost everywhere on $\mathbb{T}$ but which does not necessarily belong to $\mathcal{B}$ (see e.g. [41] and also the discussion in Section 5). The second condition on $\widetilde{f}$ in the above definition is only a normalization of $\widetilde{f}$ ensuring that $\widetilde{f}$ is unique if it exists.

The importance of the Fourier coefficients (2.1) originates from the fact that they determine the function f uniquely. In saying this, f is considered as an element of L^1 and functions which differ only on a set of Lebesgue measure zero are identified as equivalent. The interesting question is now, how can we

V. Pohl, H. Boche, *Advanced Topics in System and Signal Theory*, Foundations in Signal Processing, Communications and Networking 4,
DOI 10.1007/978-3-642-03639-2_2,

recapture f and its conjugate function $\widetilde{f}$ from the Fourier coefficients $\hat{f}(n)$. This is usually done by partial sums of the form

$$(S_N^{(w)} f)(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n=-N}^{N} w(n)\, \hat{f}(n)\, \mathrm{e}^{\mathrm{i}n\theta} \tag{2.2}$$

in which the sequence $\{w(n)\}_{n=-N}^{N}$ of complex numbers is an arbitrary *window function* which weights the Fourier coefficients. The series

$$(\widetilde{S}_N^{(w)} f)(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n=-N}^{N} -\mathrm{i}\ \mathrm{sgn}(n)\, w(n)\, \hat{f}(n)\, \mathrm{e}^{\mathrm{i}n\theta} \tag{2.3}$$

with the sign function

$$\mathrm{sgn}(k) = \begin{cases} 1\,, & k > 0 \\ 0\,, & k = 0 \\ -1\,, & k < 0 \end{cases}$$

and with the same window $\{w(n)\}_{n=-N}^{N}$ and the same degree N is called the *series conjugate* to $S_N^{(w)} f$. The usage of the name conjugate series is justified by Definition 2.1 and by the fact that series

$$(S_N^{(w)} f)(\mathrm{e}^{\mathrm{i}\theta}) + \mathrm{i}\,(\widetilde{S}_N^{(w)} f)(\mathrm{e}^{\mathrm{i}\theta}) = w(0)\, \hat{f}(0) + 2\sum_{k=1}^{N} w(n)\, \hat{f}(n)\, \mathrm{e}^{\mathrm{i}n\theta} \tag{2.4}$$

has only nonnegative Fourier coefficients.

The question is whether the series (2.2) and (2.3) converge as N tends to infinity, and if they do converge, do they converge to f and $\widetilde{f}$, respectively? The answer depends on the window $\{w(n)\}_{n=-N}^{N}$ and on the actual topology, i.e. one may ask whether (2.2) and (2.3) converge pointwise, uniformly, or in some type of norm to f and $\widetilde{f}$, respectively. Here, we discuss only some of the most important windows $\{w(n)\}_{n=-N}^{N}$ which are needed in this book. All of these windows will be symmetric, in the sense that $w(-n) = w(n)$ for all n.

To investigate the convergence behavior of (2.2) and (2.3), it is sometimes more convenient to write those series in closed form by inserting the Fourier coefficients (2.1) into (2.2) and (2.3). This gives the following integral representations

$$(S_N^{(w)} f)(\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\, K_N^{(w)}(\theta - \tau)\, \mathrm{d}\tau \tag{2.5}$$

and

$$(\widetilde{S}_N^{(w)} f)(\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\, \widetilde{K}_N^{(w)}(\theta - \tau)\, \mathrm{d}\tau\ . \tag{2.6}$$

with the kernels

$$K_N^{(w)}(\tau) = w(0) + 2\sum_{n=1}^{N} w(n)\cos(n\tau) \tag{2.7}$$

$$\widetilde{K}_N^{(w)}(\tau) = 2\sum_{n=1}^{N} w(n)\sin(n\tau). \tag{2.8}$$

Of course, these kernels depend on the (symmetric) window $\{w(n)\}_{n=-N}^{N}$ which is indicated by the superscript (w). It is immediately clear that the kernels (2.7) and (2.8) are 2π-periodic. Next, three special windows and the corresponding series are discussed.

2.1.1 Fourier Series

The most simple and best known window is the rectangular window given by

$$w(n) = 1\,, \qquad n = 0, \pm 1, \pm 2, \ldots, \pm N\ . \tag{2.9}$$

In this case, the partial sums (2.2) and (2.3) are just the truncated *Fourier* and the *conjugate Fourier series* of f, and will be denoted by s_N and $\widetilde{s}_N$ respectively. The kernel (2.7) of this particular series is called the *Dirichlet kernel* given by

$$\mathcal{D}_N(\tau) = 1 + 2\sum_{n=1}^{N}\cos(n\tau) = \frac{\sin\left([N+\frac{1}{2}]\tau\right)}{\sin(\tau/2)}\ . \tag{2.10}$$

Similarly, the kernel (2.8) of the conjugate series is called *conjugate Dirichlet kernel* given by

$$\widetilde{\mathcal{D}}_N(\tau) = 2\sum_{n=1}^{N}\sin(n\tau) = \frac{1}{\tan(\tau/2)} - \frac{\cos\left([N+\frac{1}{2}]\tau\right)}{\sin(\tau/2)}\ .$$

Therewith the partial Fourier series are given by

$$s_N(f;\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n=-N}^{N}\hat{f}(n)\,\mathrm{e}^{\mathrm{i}n\theta} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\,\mathcal{D}_N(\theta-\tau)\,\mathrm{d}\tau \tag{2.11}$$

and likewise for $\widetilde{s}_N$.

If the function f belongs to L^2 then it is well known that the partial sums s_N and $\widetilde{s}_N$ of the Fourier series converge to f and $\widetilde{f}$ in the L^2-norm, respectively (cf. Example 1.9). However, if f belongs only to L^1, the partial sums s_N and $\widetilde{s}_N$ do not converge to f and $\widetilde{f}$ in the L^1-norm, in general. Moreover if f is a continuous function on $\mathbb{T}$, one might require that s_N and $\widetilde{s}_N$ converge uniformly to f and $\widetilde{f}$, respectively, i.e. that

$$\lim_{N\to\infty}\|f - s_N(f;\cdot)\|_\infty = 0 \qquad \text{and} \qquad \lim_{N\to\infty}\|\widetilde{f} - \widetilde{s}_N(f;\cdot)\|_\infty = 0$$

for all $f \in \mathcal{C}(\mathbb{T})$. However, the partial sums s_N and $\widetilde{s}_N$ do not show such nice behavior. It even happens that for some continuous functions f, the truncated Fourier series $s_n(f;\cdot)$ does not even converge pointwise. More precisely, one has the following result.

Theorem 2.2. *To every $\theta \in [-\pi,\pi)$, there corresponds a set $E(\theta) \subset \mathcal{C}(\mathbb{T})$ of second category which is dense in $\mathcal{C}(\mathbb{T})$ such that*

$$\sup_{N\in\mathbb{N}} |s_N(f;e^{i\theta})| = \infty$$

for every $f \in E(\theta)$.

Proof. Let $\theta \in [-\pi,\pi)$ be arbitrary but fixed. Then for every N (2.11) defines a linear functional $s_N(f;e^{i\theta}) = \mathfrak{D}_N f$ on $\mathcal{C}(\mathbb{T})$, for which holds

$$|\mathfrak{D}_N f| \;\leq\; \|f\|_\infty \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} |\mathcal{D}_N(\tau)|\,d\tau \;=\; \|f\|_\infty\, \|\mathcal{D}_N\|_1$$

which shows that $\|\mathfrak{D}_N\| \leq \|D_N\|_1$. Actually, equality holds. To see this, we define the function

$$g_N(e^{i\tau}) := \begin{cases} 1 & \text{for all } \tau \text{ for which } \mathcal{D}_N(\theta-\tau) \geq 0 \\ -1 & \text{for all } \tau \text{ for which } \mathcal{D}_N(\theta-\tau) < 0 \end{cases}$$

for which certainly holds $\|g_N\|_\infty = 1$ and $|\mathfrak{D}_N\, g_N| = \|\mathcal{D}_N\|_1$ for all $N \in \mathbb{N}$. Moreover, by Lusin's Theorem (see e.g. [70, §2.24]), for every $\epsilon > 0$ there exists an $f_N \in \mathcal{C}(\mathbb{T})$ with $\|f_N\|_\infty \leq 1$ such that

$$\|f_N - g_N\|_\infty = \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} |f_N(\zeta) - g_N(\zeta)| < \epsilon\,.$$

Therewith, one obtains

$$\begin{aligned} |\mathfrak{D}_N\, f_N| &= |\mathfrak{D}_N\, g_N - \mathfrak{D}_N(g_N - f_N)| \geq |\mathfrak{D}_N\, g_N| - |\mathfrak{D}_N(g_N - f_N)| \\ &\geq \|\mathcal{D}_N\|_1 - \epsilon\, \|\mathcal{D}_N\|_1 \end{aligned}$$

such that

$$\|\mathfrak{D}_N\| \;=\; \sup_{\substack{f\in\mathcal{C}(\mathbb{T})\\ \|f\|_\infty\leq 1}} |\mathfrak{D}_N f| \;\geq\; |\mathfrak{D}_N f_N| \;\geq\; (1-\epsilon)\,\|\mathcal{D}_N\|_1$$

which shows that $\|\mathfrak{D}_N\| = \|\mathcal{D}_N\|_1$. Next, it is shown that $\|\mathcal{D}_N\|_1$ diverges as $N \to \infty$. Since the kernel $\mathcal{D}_N$, given by (2.10), is an even function and since $\sin(\tau/2) \leq \tau/2$ for all $0 \leq \tau < \pi$, one obtains

$$\begin{aligned} \|\mathcal{D}_N\|_1 &\geq \frac{2}{\pi}\int_0^{\pi} |\sin([N+1/2]\tau)|\;\frac{d\tau}{\tau} = \frac{2}{\pi}\int_0^{[N+1/2]\pi} |\sin(\tau)|\;\frac{d\tau}{\tau} \\ &> \frac{2}{\pi}\sum_{k=1}^{N}\int_{[k-1]\pi}^{k\pi} \frac{|\sin(\tau)|}{\tau}\,d\tau \geq \frac{2}{\pi}\sum_{k=1}^{N}\frac{2}{k\,\pi} \geq \frac{4}{\pi^2}\,\log(N+1) \quad (2.12) \end{aligned}$$

which shows that $\|\mathfrak{D}_N\| = \|\mathcal{D}_N\|_1 \to \infty$ as $N \to \infty$. This divergence of the operator norm $\|\mathfrak{D}_N\|$ implies by the Banach-Steinhaus theorem (see e.g. [70, §5.8]) that there exists a dense subset $E(\theta) \subset \mathcal{C}(\mathbb{T})$ of second category such that $\sup_{N\in\mathbb{N}} |\mathfrak{D}_N f| = \infty$ for all $f \in E(\theta)$. $\square$

The previous proof showed that the divergence of the Fourier series $s_N(f; \mathrm{e}^{\mathrm{i}\theta})$ at some points $\theta \in [-\pi, \pi)$ is caused by the slow decay of the envelope of the Dirichlet kernel $\mathcal{D}_N(\tau)$ as $\tau \to \infty$. It decreases only proportional to $1/\tau$ which causes the divergence of $\|\mathcal{D}_N\|_1$ as $N \to \infty$ (cf. (2.12) and Fig. 2.2). Thus, to obtain an approximation series (2.2) which converges uniformly, one needs a window $w(n)$ such that the corresponding kernel $K_N^{(w)}(\tau)$ (2.7) decreases faster than $1/\tau$ as $\tau \to \infty$. Two such methods will be discussed in the next subsections.

2.1.2 First arithmetic means – Fejér series

The unfavorable convergence behavior of the partial sums (2.11) is resolved if one considers the *(first) arithmetic means* of the partial sums (2.11)

$$\sigma_N(f; \mathrm{e}^{\mathrm{i}\theta}) := \frac{1}{N} \sum_{k=0}^{N-1} s_k(f; \mathrm{e}^{\mathrm{i}\theta}) \,. \tag{2.13}$$

If (2.11) is inserted into this arithmetic mean, a straight forward calculation gives the representations

$$\sigma_N(f; \mathrm{e}^{\mathrm{i}\theta}) = \sum_{n=-N}^{N} w(n)\, \hat{f}(n)\, \mathrm{e}^{\mathrm{i}n\theta} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\, \mathcal{F}_N(\theta - \tau)\, \mathrm{d}\tau \tag{2.14}$$

with the window function

$$w(n) = 1 - \tfrac{|n|}{N} \,, \qquad n = 0, \pm 1, \pm 2, \cdots \pm N \,. \tag{2.15}$$

and with the kernel

$$\mathcal{F}_N(\tau) = \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{D}_k(\tau) = \frac{1}{N} \frac{\sin^2(N\,\tau/2)}{\sin^2(\tau/2)} \,. \tag{2.16}$$

This first arithmetic mean (2.14) will be called the *Fejér series* of f with the *Fejér kernel* (2.16). For illustration, the window function (2.15) and the corresponding kernel (2.16) are plotted in Fig. 2.1 and 2.2, respectively. Similarly, forming the arithmetic mean of the conjugate partial sums $s_N(f; \mathrm{e}^{\mathrm{i}\theta})$, one obtains the conjugate Fejér mean $\widetilde{\sigma}_N(f; \mathrm{e}^{\mathrm{i}\theta})$ with the same window (2.15) and with the *conjugate Fejér kernel*

$$\widetilde{\mathcal{F}}_N(\tau) = \frac{1}{N} \sum_{k=0}^{N-1} \widetilde{\mathcal{D}}_k(\tau) = \frac{1}{\tan(\tau/2)} - \frac{1}{N} \frac{\sin(N\,\tau)}{2\,\sin^2(\tau/2)} \,.$$

It turns out that the Fejér series posses a much better convergence behavior with respect to continuous functions than the Fourier series. This will follow immediately from the observation that the Fejér kernel is a so called *approximate identity*:

Proposition 2.3. *The Fejér kernel* (2.16) *is an* approximate identity, *i.e. it satisfies for all* $N \in \mathbb{N}$ *the following three properties*

(*a*) $\mathcal{F}_N(\tau) \geq 0$ *for all* $\tau \in [-\pi, \pi)$

(*b*) $\frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{F}_N(\tau)\, \mathrm{d}\tau = 1$

(*c*) $\lim_{N\to\infty} \mathcal{F}_N(\tau) = 0$ *for all* $0 < |\tau| \leq \pi$

Proof. Property (*a*) is obvious from (2.16) and (*b*) follows at once from (2.16) and (2.10). To verify (*c*) choose an arbitrary but fixed $0 < \epsilon < \pi$. Then one obtains from (2.16) that

$$|\mathcal{F}_N(\tau)| \leq \frac{1}{N} \frac{1}{|\sin^2(\tau/2)|} \leq \frac{1}{N} \frac{1}{|\sin^2(\epsilon/2)|} \quad \text{for all } \epsilon \leq |\tau| \leq \pi$$

where the right had side goes to zero as $N \to \infty$. □

Remark. Sometimes, kernels with the three properties (*a*), (*b*), and (*c*) are also called *positive kernels*. However, the proof of the following theorem will show that the definition of an approximate identity seems to be more appropriate. Moreover, we will call a kernel *positive* if it satisfies only property (*a*). With this definitions, a positive kernel need not be an approximate identity, in general, but an approximate identity is always positive.

Theorem 2.4. *Let* $f \in \mathcal{C}(\mathbb{T})$ *be a continuous function on* $\mathbb{T}$. *Then its Fejér series* (2.14) *converges uniformly to* f, *i.e.*

$$\lim_{N\to\infty} \|\sigma_N(f;\cdot) - f\|_\infty = 0 .$$

Proof. Since f is given on the unit circle, the integral representation on the right hand side of (2.14) may be written as

$$\sigma_N(f;\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}(\theta-\tau)})\, \mathcal{F}_N(\tau)\, \mathrm{d}\tau .$$

Therewith and using Properties (a) and (b) of the Fejér kernel, one gets

$$|\sigma_N(f;\mathrm{e}^{\mathrm{i}\theta}) - f(\mathrm{e}^{\mathrm{i}\theta})| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(\mathrm{e}^{\mathrm{i}(\theta-\tau)}) - f(\mathrm{e}^{\mathrm{i}\theta})|\, \mathcal{F}_N(\tau)\, \mathrm{d}\tau . \quad (2.17)$$

We have to show that to every $\epsilon > 0$ there exists an $N_0 \in \mathbb{N}$ such that $|\sigma_N(f;\mathrm{e}^{\mathrm{i}\theta}) - f(\mathrm{e}^{\mathrm{i}\theta})| < \epsilon$ for all $N \geq N_0$ and for all $\theta \in [-\pi, \pi)$. To this end, we fix $\epsilon > 0$ and choose $\delta > 0$ such that

$$\sup_{-\delta\leq\tau\leq\delta} |f(e^{i(\theta-\tau)}) - f(e^{i\theta})| \leq \tfrac{\epsilon}{2} \qquad \text{for each } \theta \in [-\pi,\pi)\ . \tag{2.18}$$

This is always possible since f is continuous at every θ. Next we split up the integral in (2.17) into an integration over the interval $[-\delta,\delta]$ and over the interval $\delta \leq |\tau| \leq \pi$. Using (2.18) to upper bound the integral over $[-\delta,\delta]$, one obtains for an arbitrary $\theta \in [-\pi,\pi)$

$$\begin{aligned} &|\sigma_N(f;e^{i\theta}) - f(e^{i\theta})| \\ &\qquad \leq \frac{\epsilon}{2}\,\frac{1}{2\pi}\int_{-\delta}^{\delta} \mathcal{F}_N(\tau)\,d\tau \;+\; 2\,\|f\|_\infty\,\frac{1}{2\pi}\int_{|\tau|\geq\delta} \mathcal{F}_N(\tau)\,d\tau\ . \end{aligned} \tag{2.19}$$

By properties (a) and (b) of the Fejér kernel, the first term on the right hand side is certainly smaller than $\epsilon/2$. Since $\|f\|_\infty < \infty$, Property (c) of the Fejér kernel shows that there exists an N_0 such that $\mathcal{F}_N(\tau) < \epsilon/(4\,\|f\|_\infty)$ for all $N > N_0$ and all $\delta \leq |\tau| < \pi$. Therewith also the second term in (2.19) is upper bounded by $\epsilon/2$ such that

$$|\sigma_N(f;e^{i\theta}) - f(e^{i\theta})| \leq \epsilon \qquad \text{for all } N \geq N_0 \text{ and all } \theta \in [-\pi,\pi)\ .$$

This is what we wanted to show. □

Since $\mathcal{C}(\mathbb{T})$ is dense in every L^p with $1 \leq p < \infty$, the previous theorem implies that for every $f \in L^p$ the Fejér series $\sigma_N(f;\cdot)$ converges to f in L^p.

2.1.3 Delayed first arithmetic means – de-la-Vallée-Poussin Series

As we will see, it is favorable in some sense, to introduce a delay K in the first arithmetic mean (2.13) and to take the mean of the partial sums $s_K, s_{K+1}, \ldots, s_{K+N-1}$. This gives the so-called *delayed first arithmetic mean*

$$\sigma_{N,K}(f;e^{i\theta}) := \frac{1}{N}\sum_{k=K}^{K+N-1} s_k(f;e^{i\theta})\ . \tag{2.20}$$

wherein $s_k(f;\cdot)$ is again the partial Fourier series (2.11) of f. For $K = 0$, one obtains again the first arithmetic mean (2.13), i.e. $\sigma_{N,0} = \sigma_N$. It is clear that the delayed arithmetic mean can be expressed as the difference of the two (not-delayed) arithmetic means σ_{K+N} and σ_K. Since $N\,\sigma_{K,N} = (K + N)\,\sigma_{K+N} - K\,\sigma_K$ one has the representation

$$\sigma_{N,K}(f;e^{i\theta}) = (1 + \tfrac{K}{N})\,\sigma_{K+N}(f;e^{i\theta}) - \tfrac{K}{N}\,\sigma_K(f;e^{i\theta}) \tag{2.21}$$

for the delayed arithmetic means. Subsequently we consider mainly the particular case where $K = N$. Then the delayed arithmetic mean becomes

$$\sigma_{N,N}(f;e^{i\theta}) = 2\,\sigma_{2N}(f;e^{i\theta}) - \sigma_N(f;e^{i\theta})\ . \tag{2.22}$$

For the particular case ($K = N$), we insert the partial sums (2.11) into (2.20). A straight forward calculation gives the representations

$$\sigma_{N,N}(f;\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n=-2N}^{2N} w(n)\,\hat{f}(n)\,\mathrm{e}^{\mathrm{i}n\theta} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\,\mathcal{V}_N(\theta-\tau)\,\mathrm{d}\tau \quad (2.23)$$

with the trapezoid window function

$$w(n) = \begin{cases} 1 & 0 \le |n| \le N \\ 2\,(1-\frac{n}{2N}) & N+1 < |n| \le 2N \end{cases} \quad (2.24)$$

and with the kernel

$$\mathcal{V}_N(\tau) = 2\,\mathcal{F}_{2N}(\tau) - \mathcal{F}_N(\tau) = \frac{1}{N}\,\frac{\cos(N\tau)-\cos(2N\tau)}{2\,\sin^2(\tau/2)}\,. \quad (2.25)$$

This particular ($K = N$) delayed first arithmetic mean (2.23) is called the *de-la-Vallée-Poussin series* of f and (2.25) is the *de-la-Vallée-Poussin kernel.*

Similarly, forming the delayed arithmetic mean of the conjugate partial sums $\widetilde{s}_N(f;\mathrm{e}^{\mathrm{i}\theta})$, one obtains the *conjugate de-la-Vallée-Poussin series* $\widetilde{\sigma}_{N,N}(f;\mathrm{e}^{\mathrm{i}\theta})$ with the same window (2.24) and with the *conjugate de-la-Vallée-Poussin kernel*

$$\widetilde{\mathcal{V}}_N(\tau) = 2\,\widetilde{\mathcal{F}}_{2N}(\tau) - \widetilde{\mathcal{F}}_N(\tau) = \frac{1}{\tan(\tau/2)} - \frac{1}{N}\,\frac{\sin(N\tau)-\sin(2N\tau)}{2\,\sin^2(\tau/2)}\,.$$

It should be noted, that the de-la-Vallée-Poussin series $\sigma_{N,N}(f;\cdot)$ of a function f involves $4N-1$ Fourier coefficients of f whereas the Fejér series $\sigma_N(f;\cdot)$ uses only $2N-1$ Fourier coefficients of f.

The Fejér series $\sigma_N(f;\cdot)$ of a continuous function f converges uniformly to f. Since by (2.22) the de-la-Vallée-Possin series $\sigma_{N,N}(f;\cdot)$ is just the difference of two Fejér series one has immediately the following corollary of Theorem 2.4.

Corollary 2.5. *Let $f \in \mathcal{C}(\mathbb{T})$ be a continuous function on $\mathbb{T}$. Then its de-la-Vallée-Poussin series* (2.23) *converges uniformly to f, i.e.*

$$\lim_{N\to\infty} \|\sigma_{N,N}(f;\cdot) - f\|_\infty = 0\,.$$

Proof. By Theorem 2.4, the Fejér series $\sigma_N(f;\mathrm{e}^{\mathrm{i}\theta})$ converges to $f(\mathrm{e}^{\mathrm{i}\theta})$, uniformly in θ and therefore, it is a Cauchy sequence (uniformly in θ). Using the representation (2.22) of the de-la-Vallée-Poussin series, one obtains

$$|\sigma_{N,N}(f;\mathrm{e}^{\mathrm{i}\theta}) - f(\mathrm{e}^{\mathrm{i}\theta})| \le |\sigma_{2N}(f;\mathrm{e}^{\mathrm{i}\theta}) - f(\mathrm{e}^{\mathrm{i}\theta})| + |\sigma_{2N}(f;\mathrm{e}^{\mathrm{i}\theta}) - \sigma_N(f;\mathrm{e}^{\mathrm{i}\theta})|$$

where the right hand side goes to zero as $N\to\infty$, uniformly in θ. □

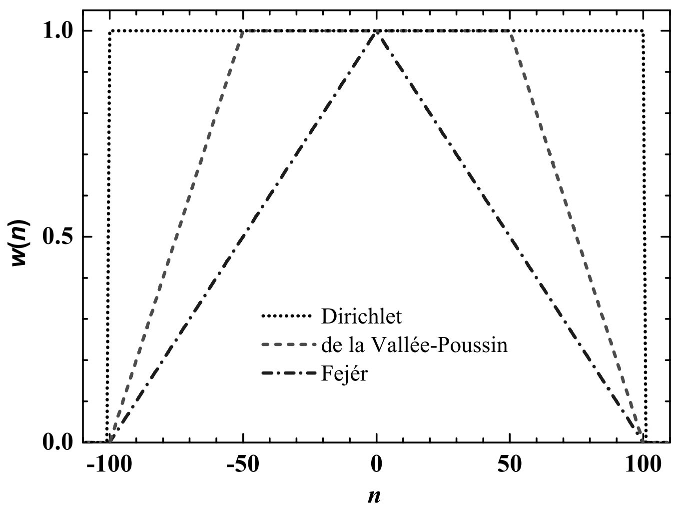

Fig. 2.1. Window functions corresponding to the Fourier series, the first arithmetic means, and the delayed first arithmetic mean for the order $N = 100$.

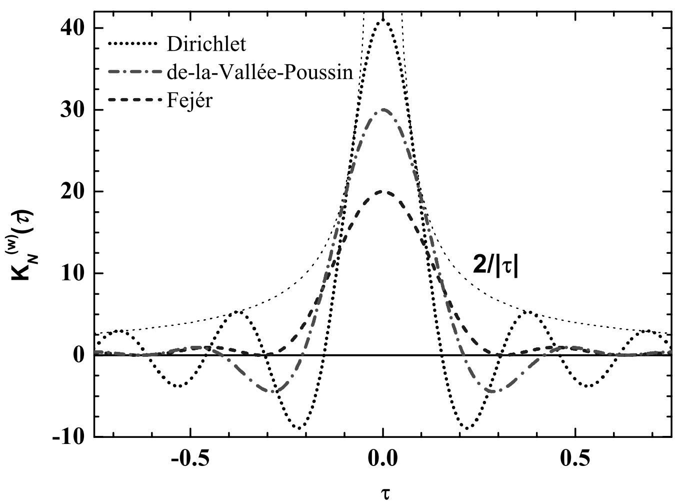

Fig. 2.2. Dirichlet, Fejér, and de-la-Vallée-Poussin kernel for the order $N = 20$.

Figures 2.1 and 2.2 show the window functions and the kernels, respectively, of the three weighted trigonometric series discussed above. The window function $w(n)$ determines primarily the convergence behavior of the partial sums (2.2) and (2.3) as $N \to \infty$. In the next subsection, for example, we will show that the de-la-Vallée-Poussin mean possesses the property that the approximation error decreases almost as fast as possible as $N \to \infty$. However, in applications other properties of the approximation series may also be of some importance. Therefore there does not exist an "optimal" window function, in general, but for different applications, different window functions may be favorable. Consequently, there exists many more possible window functions. In Section 10.5 we will discuss the optimal kernel for the approximation of spectral densities in some detail.

2.1.4 Best approximation by trigonometric polynomials

In what follows, $\mathcal{P}(N)$ denotes the set of all trigonometric polynomials with degree at most N, i.e. the set of all functions of the form

$$f(\mathrm{e}^{\mathrm{i}\theta}) = \frac{a_0}{2} + \sum_{k=1}^{N} a_k \cos(k\,\theta) + b_k \sin(k\,\theta)\,, \qquad \theta \in [-\pi, \pi)$$

with real coefficients $\{a_k\}_{k=0}^N$ and $\{b_k\}_{k=1}^N$. The subset of all $f \in \mathcal{P}(N)$ with a zero constant term a_0 i.e. all $f \in \mathcal{P}(N)$ for which $\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta = 0$ is denoted by $\mathcal{P}_0(N)$.

Theorem 2.4 and Corollary 2.5 imply that every continuous function $f \in \mathcal{C}(\mathbb{T})$ can be uniformly approximated by a trigonometric polynomial, e.g. by the Fejér or de-la-Vallée-Poussin mean. Thus given an $\epsilon > 0$ one always finds a polynomial $p \in \mathcal{P}(N)$ of sufficiently large degree N such that $\|f - p\|_\infty < \epsilon$. Of course, for practical reasons, it is desirable to find the polynomial with the smallest degree N which satisfies the error requirement. We want to show that in a sense the de-la-Vallée-Poussin means are such approximation polynomials with an almost minimal degree.

Given a continuous function $f \in \mathcal{C}(\mathbb{T})$ and a fixed degree $N \geq 0$, the *best approximation* of f of degree N is defined as the number

$$B_N[f] := \inf_{p_N \in \mathcal{P}(N)} \|f - p_N\|_\infty\,. \tag{2.26}$$

It is clear that $B_N[f]$ tends to zero as $N \to \infty$. This follows from letting $\sigma_N(f;\cdot)$ be the Fejér series of f. Then $B_N[f] \leq \|f - \sigma_N(f;\cdot)\|_\infty$ and the right hand side converges to zero by Theorem 2.4. Next we observe that the infimum in (2.26) is attained in $\mathcal{P}(N)$.

Proposition 2.6. *To every $f \in \mathcal{C}(\mathbb{T})$ there exists a polynomial $p_N^* \in \mathcal{P}(N)$ such that $\|f - p_N^*\|_\infty = B_N[f]$.*

Proof. Let $f \in \mathcal{C}(\mathbb{T})$ be arbitrary and fix a degree N. Then by definition (2.26) of $B_N[f]$, there exists a sequence $\{p_N^{(k)}\}_{k=1}^{\infty}$ of polynomials in $\mathcal{P}(N)$ such that to every $\epsilon > 0$ there exists a K_0 such that

$$\|f - p_N^{(k)}\|_\infty \leq B_N[f] + \epsilon \,. \tag{2.27}$$

for all $k \geq K_0$. In particular, all trigonometric polynomials $p_N^{(k)}$ are uniformly bounded which implies that all Fourier coefficients $\hat{p}_N^{(k)}(n)$, $n = 0, \pm 1, \pm 2, \ldots$ of these polynomials are uniformly bounded. By the theorem of Bolzano-Weierstrass, for every n there exists a subsequence of $\hat{p}_N^{(k_i)}(n)$ which converges to a limit $\hat{p}_N^*(n)$. The corresponding subsequence of trigonometric polynomials converges uniformly to the polynomial

$$p_N^*(\mathrm{e}^{\mathrm{i}\theta}) = \textstyle\sum_{n=-N}^{N} \hat{p}_N^*(n)\, \mathrm{e}^{\mathrm{i}n\theta} \,, \qquad \theta \in [-\pi, \pi)$$

for which (2.27) gives $\|f - p_N^*\|_\infty \leq B_N[f]$. Since the reverse inequality is obvious, one gets the desired statement. □

Even though we know that the best approximation is attained, it will be difficult, in general, to determine the optimal polynomial p_N^*. However, the next theorem will show that the de-la-Vallée-Poussin mean $\sigma_{N,N}(f;\cdot)$ of f is always near the optimal polynomial, in the sense that the approximation error $\|f - \sigma_{N,N}(f;\cdot)\|_\infty$ is upper bounded by four times the best approximation.

Theorem 2.7. *Let $f \in \mathcal{C}(\mathbb{T})$ and let $\sigma_{N,K}(f;\cdot)$ with $K \geq N \geq 0$ be its delayed arithmetic mean. Then*

$$\|f - \sigma_{N,K}(f;\cdot)\|_\infty \leq 2\,(1 + \tfrac{K}{N})\, B_N[f]$$

and in particular $\|f - \sigma_{N,N}(f;\cdot)\|_\infty \leq 4\, B_N[f]$.

Proof. Fix the degree N and denote by $p_N \in \mathcal{P}(N)$ the trigonometric polynomial which attains the best approximation according to Proposition 2.6. Write f as

$$f(\mathrm{e}^{\mathrm{i}\theta}) = p_N(\mathrm{e}^{\mathrm{i}\theta}) + r(\mathrm{e}^{\mathrm{i}\theta}) \,, \qquad \theta \in [-\pi, \pi) \,. \tag{2.28}$$

It follows for the rest term that

$$|r(\mathrm{e}^{\mathrm{i}\theta})| = |f(\mathrm{e}^{\mathrm{i}\theta}) - p_N(\mathrm{e}^{\mathrm{i}\theta})| \leq \|f - p_N\|_\infty = B_N[f]$$

for all $\theta \in [-\pi, \pi)$, which implies for the Fejér mean of r that

$$|\sigma_k(r;\mathrm{e}^{\mathrm{i}\theta})| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |r(\mathrm{e}^{\mathrm{i}\tau})|\, \mathcal{F}_k(\theta - \tau)\, \mathrm{d}\tau \leq B_N[f]$$

for all $\theta \in [-\pi, \pi)$ and for every arbitrary $k \geq 0$, using that $\mathcal{F}_k$ is an approximate identity. Applying (2.21) one obtains for the delayed arithmetic mean of r that

$$\begin{aligned}|\sigma_{N,K}(r;\mathrm{e}^{\mathrm{i}\theta})| &\leq (1+\tfrac{K}{N})\,|\sigma_{K+N}(f;\mathrm{e}^{\mathrm{i}\theta})| + \tfrac{K}{N}\,|\sigma_K(f;\mathrm{e}^{\mathrm{i}\theta})| \\ &\leq (1+2\,\tfrac{K}{N})\,B_N[f] \end{aligned} \tag{2.29}$$

for all $\theta \in [-\pi,\pi)$ and arbitrary $K, N \geq 0$.

Consider the partial Fourier series (2.11) of the polynomial p_N and assume that $k \geq N$. Then $s_k(p_N;\mathrm{e}^{\mathrm{i}\theta}) = p_N(\mathrm{e}^{\mathrm{i}\theta})$ for all θ. Therefore, it follows from (2.28) for the partial Fourier series of f that

$$s_k(f;\mathrm{e}^{\mathrm{i}\theta}) = p_N(\mathrm{e}^{\mathrm{i}\theta}) + s_k(r;\mathrm{e}^{\mathrm{i}\theta})\,, \qquad \theta \in [-\pi,\pi)$$

for all $k \geq N$, and for an arbitrary $K \geq N$. The delayed first arithmetic mean of f becomes

$$\frac{1}{N}\sum_{k=K}^{K+N-1} s_k(f;\mathrm{e}^{\mathrm{i}\theta}) = p_N(\mathrm{e}^{\mathrm{i}\theta}) + \frac{1}{N}\sum_{k=K}^{K+N-1} s_k(r;\mathrm{e}^{\mathrm{i}\theta})\,,$$

which is equivalent to

$$\sigma_{N,K}(f;\mathrm{e}^{\mathrm{i}\theta}) = p_N(\mathrm{e}^{\mathrm{i}\theta}) + \sigma_{N,K}(r;\mathrm{e}^{\mathrm{i}\theta})\,, \qquad \theta \in [-\pi,\pi)\,. \tag{2.30}$$

Finally, one obtains for all $K \geq N$

$$\begin{aligned}|f(\mathrm{e}^{\mathrm{i}\theta}) - \sigma_{N,K}(f;\mathrm{e}^{\mathrm{i}\theta})| &\leq |f(\mathrm{e}^{\mathrm{i}\theta}) - p_N(\mathrm{e}^{\mathrm{i}\theta})| + |p_N(\mathrm{e}^{\mathrm{i}\theta}) - \sigma_{N,K}(f;\mathrm{e}^{\mathrm{i}\theta})| \\ &\leq B_N[f] + |\sigma_{N,K}(r;\mathrm{e}^{\mathrm{i}\theta})| \\ &\leq 2\,(1+\tfrac{K}{N})\,B_N[f]\end{aligned}$$

where the second inequality follows from the definition of p_N and from (2.30), whereas the last inequality is a consequence of (2.29). □

We already saw that the best approximation $B_N[f]$ of every continuous function $f \in \mathcal{C}(\mathbb{T})$ converges to zero as $N \to \infty$. However, one would expect that it is easier to approximate a "simple" function than a "complicated" functions. Thus, the best approximation $B_N[f]$ of a "simple" function should converge more rapidly than for a complicated function. The following theorem will show that this is indeed the case, and that in this context a "simple function" is a smooth function, i.e. the smoother the function f, the faster $B_N[f]$ converges to zero.

Theorem 2.8. *Let f be a k-times differentiable function on the unit circle $\mathbb{T}$ whose k-th derivative $f^{(k)}$ has a modulus of continuity of ω. Then there exists a constant C_k, which depends only on k, such that*

$$B_N[f] \leq C_k\,\omega\left(\tfrac{1}{N}\right)\,N^{-k}\,.$$

For functions from the Hölder-Zygmund class $\Lambda_\alpha(\mathbb{T})$ (cf. Section 1.3), we obtain immediately the following corollary as a special case of Theorem 2.8.

Corollary 2.9. *If $f \in \Lambda_\alpha(\mathbb{T})$ for $\alpha > 0$, then there exists a constant C_α such that*

$$B_N[f] \le C_\alpha\, N^{-\alpha}\,.$$

Proof (Theorem 2.8). The proof consists of several steps. Each considers a special case of Theorem 2.8.
1) First it is shown that if $f \in \mathcal{C}(\mathbb{T})$ and has modulus of continuity ω, then $B_N[f] < 12\,\omega(1/N)$. To this end, we use the following approximation

$$f_N(\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\,\mathcal{J}_N(\theta-\tau)\,\mathrm{d}\tau \tag{2.31}$$

with the so-called *Jackson kernel*

$$\mathcal{J}_N(\tau) = \frac{3}{N\,(2\,N^2+1)}\,\left(\frac{\sin(N\,t/2)}{\sin(t/2)}\right)^4\,. \tag{2.32}$$

This approximation method has the following three properties[1] which are used throughout the rest of the proof.

(*a*) $f_N \in \mathcal{P}(2\,N-2)$, i.e. f_N is a trigonometric polynomial of degree $2\,N-2$.
(*b*) $f_N \in \mathcal{P}_0(2\,N-2)$ whenever $\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta = 0$.
(*c*) The kernel (2.32) satisfies the relation

$$\tfrac{1}{2\,\pi}\int_{-\pi}^{\pi}\mathcal{J}_N(\tau)\,\mathrm{d}\tau = 1\,. \tag{2.33}$$

Replacing $\theta-\tau$ by s and splitting up the integral in (2.31) gives

$$\begin{aligned} f_N(\mathrm{e}^{\mathrm{i}\theta}) &= \frac{1}{2\pi}\int_{-\pi}^{0} f(\mathrm{e}^{\mathrm{i}(\theta-s)})\,\mathcal{J}_N(s)\,\mathrm{d}s + \frac{1}{2\pi}\int_{0}^{\pi} f(\mathrm{e}^{\mathrm{i}(\theta-s)})\,\mathcal{J}_N(s)\,\mathrm{d}s \\ &= \frac{1}{2\pi}\int_{0}^{\pi}\Big[f(\mathrm{e}^{\mathrm{i}(\theta+s)}) + f(\mathrm{e}^{\mathrm{i}(\theta-s)})\Big]\,\mathcal{J}_N(s)\,\mathrm{d}s \end{aligned}$$

where the second line follows after the change variable $s \mapsto -s$ in the first integral. Next we use property (2.33) of the kernel $\mathcal{J}_N$. This yields

$$\begin{aligned} \big|f(\mathrm{e}^{\mathrm{i}\theta}) - f_N(\mathrm{e}^{\mathrm{i}\theta})\big| &= \left|\frac{1}{2\pi}\int_0^{\pi}\Big[2\,f(\mathrm{e}^{\mathrm{i}\theta}) - f(\mathrm{e}^{\mathrm{i}(\theta+s)}) - f(\mathrm{e}^{\mathrm{i}(\theta-s)})\Big]\,\mathcal{J}_N(s)\,\mathrm{d}s\right| \\ &\le \frac{1}{2\pi}\int_0^{\pi}\Big(\Big|f(\mathrm{e}^{\mathrm{i}\theta}) - f(\mathrm{e}^{\mathrm{i}(\theta+s)})\Big| + \Big|f(\mathrm{e}^{\mathrm{i}\theta}) - f(\mathrm{e}^{\mathrm{i}(\theta-s)})\Big|\Big)\,\mathcal{J}_N(s)\,\mathrm{d}s \\ &\le \frac{1}{\pi}\int_0^{\pi}\omega(s)\,\mathcal{J}_N(s)\,\mathrm{d}s \qquad \text{for every } \theta \in [-\pi,\pi)\,. \end{aligned}$$

[1] For a proof of these properties, we refer e.g. to [61, vol. 1, Chap. IV], or just note that $\mathcal{J}_N$ is a normalized square of the Fejér kernel (2.16).

By the properties of the modulus of continuity (cf. Section 1.3), it holds $\omega(s) = \omega(N\,s\,\frac{1}{N}) \le (N\,s+1)\,\omega(1/N)$ such that

$$\left|f(\mathrm{e}^{\mathrm{i}\theta}) - f_N(\mathrm{e}^{\mathrm{i}\theta})\right| \le \omega\left(\tfrac{1}{N}\right)\,\left[\frac{N}{\pi}\int_0^\pi s\,\mathcal{J}_N(s)\,\mathrm{d}s + 1\right] \tag{2.34}$$

for every $\theta \in [-\pi,\pi)$. Next, we derive an upper bound for the integral on the right hand side. To do this, the integration interval is divided into two intervals as follows

$$\int_0^\pi s\,\mathcal{J}_N(s)\,\mathrm{d}s = \int_0^{\pi/N} s\,\mathcal{J}_N(s)\,\mathrm{d}s + \int_{\pi/N}^\pi s\,\mathcal{J}_N(s)\,\mathrm{d}s\;.$$

Now, in the first integral we use that $|\sin(N\,s/2)| \le N\,\sin(s/2)$ and in the second integral we apply the inequalities $|\sin(x)| \le 1$ and $\sin(s/2) \ge s/\pi$ for all $s \in [0,\pi)$. Therewith, one obtains

$$\begin{aligned}\int_0^\pi s\,\mathcal{J}_N(s)\,\mathrm{d}s &\le \frac{3}{N\,(2\,N^2+1)}\left\{N^4\int_0^{\pi/N} s\,\mathrm{d}s + \pi^4\int_{\pi/N}^\pi s^{-3}\,\mathrm{d}s\right\}\\ &\le \frac{3}{N\,(2\,N^2+1)}\left\{\frac{\pi^2N^2}{2} + \frac{\pi^2N^2}{2}\right\} = \frac{3\,\pi^2N}{2\,N^2+1}\;.\end{aligned}$$

Using this upper bound in (2.34), one obtains that

$$\left|f(\mathrm{e}^{\mathrm{i}\theta}) - f_N(\mathrm{e}^{\mathrm{i}\theta})\right| \le \left[\frac{3\,\pi}{2} + 1\right]\,\omega\left(\tfrac{1}{N}\right) < 6\,\omega\left(\tfrac{1}{N}\right)$$

for every $\theta \in [-\pi,\pi)$. Since f_N is a trigonometric polynomial of degree $2N-2$, it follows that $B_{2N-2}[f] < 6\,\omega(1/N)$. Assume first that $N = 2\,M$ is an even natural number. Then

$$B_N[f] = B_{2\,M}[f] \le B_{2\,M-2}[f] < 6\,\omega(1/M) = 6\,\omega(2/N) \le 12\,\omega(1/N)\;.$$

Similarly, if $N = 2\,M-1$ is an odd natural number, one obtains

$$B_N[f] = B_{2\,M-1}[f] \le B_{2\,M-2}[f] < 6\,\omega(1/M) = 6\,\omega(2/[N+1]) \le 12\,\omega(1/N)\;.$$

This is what we wanted to show. Moreover, this already proves the theorem for the case $k = 0$.

2) As a consequence of the first part, one obtains for the special case of Lipschitz continuous functions $f \in \mathrm{Lip}_K$ that $B_N[f] < 12\,K/N$.

3) Assume now that the given function f satisfies the condition

$$\textstyle\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta = 0\;. \tag{2.35}$$

Then $f_N \in \mathcal{P}_0(2\,N-2)$ by property (b) of the approximation polynomial f_N. Denote by $b_N[f]$ the best approximation of f by polynomials in $\mathcal{P}_0(N)$, i.e.

$$b_N[f] := \inf_{p_N \in \mathcal{P}_0(N)} \|f - p_N\|_\infty \, .$$

Following the derivation under point 1) of this proof, one obtains that $b_N[f] \leq 12\,\omega(1/N)$ for all $f \in \mathcal{C}(\mathbb{T})$ which satisfy (2.35). Moreover, as under point 2), for all $f \in \mathrm{Lip}_K$ which satisfy (2.35), one obtains

$$b_N[f] \leq 12\,K/N \, . \tag{2.36}$$

4) Assume now that f possesses a bounded derivative f' and denote by $b'_N = b_N[f']$ the best approximation of f' in the class $\mathcal{P}_0(N)$. Then there exists a trigonometric polynomial $u \in \mathcal{P}_0(N)$ such that

$$|f'(\mathrm{e}^{\mathrm{i}\theta}) - u(\mathrm{e}^{\mathrm{i}\theta})| \leq b'_N \, , \qquad \text{for all } \theta \in [-\pi, \pi) \, . \tag{2.37}$$

Integrating u gives a trigonometric polynomial $v \in \mathcal{P}_0(N)$ such that $v'(\mathrm{e}^{\mathrm{i}\theta}) = u(\mathrm{e}^{\mathrm{i}\theta})$. With the definition $\varphi(\mathrm{e}^{\mathrm{i}\theta}) := f(\mathrm{e}^{\mathrm{i}\theta}) - v(\mathrm{e}^{\mathrm{i}\theta})$ relation (2.37) can be written as $|\varphi'(\mathrm{e}^{\mathrm{i}\theta})| \leq b'_N$ for all $\theta \in [-\pi, \pi)$. This shows in particular that $\varphi \in \mathrm{Lip}_{b'_N}$. Using point 2) of this proof, we get $B_N[\varphi] \leq 12\, b'_N[f]/N$. Therefore, by the definition of the $B_N[\varphi]$, there exists a $w \in \mathcal{P}(N)$ such that

$$\left|\varphi(\mathrm{e}^{\mathrm{i}\theta}) - w(\mathrm{e}^{\mathrm{i}\theta})\right| = \left|f(\mathrm{e}^{\mathrm{i}\theta}) - \left[v(\mathrm{e}^{\mathrm{i}\theta}) + w(\mathrm{e}^{\mathrm{i}\theta})\right]\right| \leq \frac{12}{N}\, b'_N$$

for all $\theta \in [-\pi, \pi)$. Setting $u_N = v + w \in \mathcal{P}(N)$, the last inequality shows that

$$B_N[f] \leq \frac{12}{N}\, b'_N \, . \tag{2.38}$$

5) Assume now that f satisfies the conditions of the theorem. Then according to 4), relation (2.38) holds. Moreover, the first derivative f' has again a continuous and bounded derivative f''. Thus $f' \in \mathrm{Lip}_{b''_N}$ where b''_N denotes the best approximation of f'' in $\mathcal{P}_0(N)$. Moreover, since f is continuous on the unit circle $\mathbb{T}$, we have that $\int_{-\pi}^{\pi} f'(\mathrm{e}^{\mathrm{i}\theta})\, \mathrm{d}\theta = f(\pi) - f(-\pi) = 0$. Thus, f' satisfies also the assumption (2.35) under point 3) of this proof. Applying (2.36) to f' one obtains $b'_N = b_N[f''] \leq 12\, b''_N/N$. Now one applies the same arguments to f'', $f^{(3)}$, and so forth, up to $f^{(k-1)}$. This gives the relations

$$b_N^{(n)} \leq \frac{12}{N}\, b_N^{(n+1)} \qquad n = 1, 2, \ldots, k-1 \, . \tag{2.39}$$

Finally, it follows from point 3) of this proof that $b_N^{(k)} = b_N[f^{(k)}] \leq 12\,\omega(1/N)$. Combining this with the inequalities (2.39) and with (2.38), one obtains the statement of the theorem with $C_k = 12^{k+1}$. □

2.2 Hardy Spaces on the Unit Disk

This section gives a short introduction to a class of function spaces which contain analytic functions in the unit disk. Their importance for system theory results from the fact that their elements can be interpreted as causal transfer functions of linear systems which are bounded in a certain L^p-norm. After a short introduction of these so called Hardy spaces, we will present some results which will be needed in later parts of this book.

2.2.1 Basic definitions

Denote by H the set of all functions that are *analytic* (i.e. *holomorphic*) inside the unit disk $\mathbb{D}$. Then, it is clear that every $f(z)$ in H is bounded for all $z \in \mathbb{D}$. However, as $|z|$ approaches 1, the modulus $|f(z)|$ may go to infinity. Hardy spaces are subsets of H whose elements satisfy a certain growth condition toward the boundary of the unit disk.

Definition 2.10 (Hardy spaces). *Let $f \in H$ be an analytic function inside the unit disk $\mathbb{D}$. For $0 < p < \infty$, we set*

$$\|f\|_p := \sup_{0<r<1} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(r\mathrm{e}^{\mathrm{i}\theta})|^p \,\mathrm{d}\theta \right)^{1/p}$$

and for $p = \infty$, we define

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

Then for $0 < p \leq \infty$ the Hardy class H^p *is defined as the set of all functions f analytic in $\mathbb{D}$ for which $\|f\|_p < \infty$.*

It is not hard to verify that $H^\infty \subset H^q \subset H^p$ for all $0 < p < q < \infty$. The Hardy spaces were defined by their behavior inside the unit disk $\mathbb{D}$. The following theorem[2] characterizes the radial limits of functions in H^p.

Theorem 2.11. *Let $f \in H^p$ with $1 \leq p \leq \infty$. Then the radial limit*

$$\widetilde{f}(\mathrm{e}^{\mathrm{i}\theta}) := \lim_{r\nearrow 1} f(r\mathrm{e}^{\mathrm{i}\theta})$$

exists for almost all $\theta \in [-\pi, \pi)$. Moreover $\widetilde{f} \in L^p$ with $\|\widetilde{f}\|_{L^p} = \|f\|_{H^p}$.

Thus, the radial limit of every $f \in H^p$ exists. From now on, this radial limit will also be denoted by f. As a consequence H^p can be considered as a closed subspace of L^p and therefore every H^p with $1 \leq p \leq \infty$ is a Banach space by itself. Moreover, since H^2 is a closed subspace of the Hilbert space L^2, it is also Hilbert space with the inner product

$$\langle f, g\rangle = \lim_{r\nearrow 1} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(r\mathrm{e}^{\mathrm{i}\theta})\, \overline{g(r\mathrm{e}^{\mathrm{i}\theta})} \,\mathrm{d}\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\, \overline{g(\mathrm{e}^{\mathrm{i}\theta})} \,\mathrm{d}\theta \,.$$

[2] For a proof, we refer e.g. to [70, § 17.11], [30, § 3.2]

Conversely, let $f \in H^p$. Then Cauchy's theorem implies that all Fourier coefficients $\hat{f}(n)$ with negative index n vanish, because f is analytic inside the unit disk $\mathbb{D}$ and therefore

$$\hat{f}(n) = \frac{1}{2\pi \mathrm{i}} \oint_{\mathbb{T}} f(\zeta)\, \zeta^{-(n+1)}\, \mathrm{d}\zeta = 0 \quad \text{for all}\, n < 0$$

where the integration over $\mathbb{T}$ has to be done counter-clockwise. Consequently, H^p could be defined as the subspace of those L^p functions for which all negative Fourier coefficients are equal to zero:

$$H^p = \{ f \in L^p : \hat{f}(n) = 0 \text{ for all } n < 0 \} \, .$$

Therefore every $f \in L^p$ with the Fourier series $f(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n=0}^{\infty} \hat{f}(n)\, \mathrm{e}^{\mathrm{i}n\theta}$ can be identified with the H^p-function

$$f(z) = \sum_{n=0}^{\infty} \hat{f}(n)\, z^n$$

which is analytic for every $z \in \mathbb{D}$.

As explained in Example 1.9, there is a natural isometric isomorphism between L^2 and $\ell^2(\mathbb{Z})$, given by associating every $f \in L^2$ with the sequence $\{\hat{f}(n)\}_{n=-\infty}^{\infty}$ of its Fourier coefficients. In the same way, the Hardy space H^2 is isometricly isomorphic to the sequence space $\ell^2(\mathbb{Z}_+)$ because $\ell^2(\mathbb{Z}_+)$ may be considered as the subspace of all sequences $\{\alpha_k\}_{k=-\infty}^{\infty}$ in $\ell^2(\mathbb{Z})$ for which $\alpha_k = 0$ for $k < 0$.

The next theorem gives a useful characterization of the Fourier coefficients of H^1 functions, which will be used frequently in the following. The proof is omitted but may be found in [30] or [41], for example.

Theorem 2.12 (Hardy's inequality). *Let $f \in H^1$ with Fourier expansion $f(z) = \sum_{k=0}^{\infty} \hat{f}(k)\, z^k$. Then its Fourier coefficients satisfy the inequality*

$$\|f\|_1 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(\mathrm{e}^{\mathrm{i}\theta})|\, \mathrm{d}\theta \geq \frac{1}{\pi} \sum_{k=0}^{\infty} \frac{|\hat{f}(k)|}{k+1} \, . \tag{2.40}$$

2.2.2 Zeros of H^p-functions

Let $f \in H^p$ be an arbitrary function in a certain Hardy space H^p. Define by

$$\mathcal{Z}(f) := \{ z \in \mathbb{T} : f(z) = 0 \}$$

the *zero set* of f, i.e. the set of all those points in $\mathbb{D}$ where f is zero. For every function f holomorhic in the unit disk $\mathbb{D}$, it is well known that either $\mathcal{Z}(f) = \mathbb{D}$, or $\mathcal{Z}(f)$ has no limit point in $\mathbb{D}$. In the first case f is identically zero which is of little interest. Thus, the zeros of a non-zero holomorphic function f in $\mathbb{D}$ are isolated points in $\mathbb{T}$, and if the number of zeros is infinite, the limit points of the zeros have to lie outside of $\mathbb{D}$, i.e. on the boundary $\mathbb{T}$ of $\mathbb{D}$. If we only assume that f is holomorphic in $\mathbb{D}$, this is all we can say about the zeros of f, by the Theorem of Weierstrass (see e.g. [70, Chapter 15]). However, if we consider functions in the Hardy spaces H^p which are not only holomorphic in $\mathbb{D}$ but satisfy a certain growth behavior toward the boundary of $\mathbb{D}$, more can be said about the distribution of the zeros in $\mathbb{D}$, namely that the zeros have to converge with a certain rate toward the limit points on $\mathbb{T}$ (if they exist). The basis of deriving these conditions on the zeros of H^p functions is the following *Jensen's formula.*

Theorem 2.13 (Jensen's formula). *Let $f \in H(\mathbb{D})$ be a holomorphic function in $\mathbb{D}$ with $f(0) \neq 0$, let $0 < r < 1$, and let $\alpha_1, \ldots, \alpha_N$ be the zeros of f in the closed disk $\overline{\mathbb{D}}_r(0) = \{z \in \mathbb{C} : |z| \leq r\}$ listed according to their multiplicities. Then*

$$|f(0)| \prod_{n=1}^{N} \frac{r}{|\alpha_n|} = \exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log\left|f(re^{i\theta})\right| \, d\theta\right) . \tag{2.41}$$

Remark 2.14. Since f is considered in the disk $\overline{\mathbb{D}}_r(0) \subset \mathbb{D}$ with $r < 1$ and since f is holomorphic in $\mathbb{D}$, the zeros of f have no limit point in $\overline{\mathbb{D}}_r(0)$. Consequently, the number N of zeros in $\overline{\mathbb{D}}_r(0)$ is finite.

Remark 2.15. The assumption $f(0) \neq 0$ is no real limitation. Because if f has a zero of order m at 0, one can apply Jensen's formula to the function $f(z)/z^m$.

Proof. One orders the zeros $\{\alpha_n\}_{n=1}^N$ of f in $\overline{\mathbb{D}}_r(0)$ according to their location in $\mathbb{D}_r(0)$ and on the boundary of $\mathbb{D}_r(0)$, i.e. such that $|\alpha_1| \leq \cdots \leq |\alpha_M| < r$ and $|\alpha_{M+1}| = \cdots = |\alpha_N| = r$. Define the function

$$g(z) := f(z) \prod_{n=1}^{M} \frac{r^2 - \overline{\alpha}_n z}{r(\alpha_n - z)} \prod_{n=M+1}^{N} \frac{\alpha_n}{\alpha_n - z} , \qquad z \in \mathbb{D} . \tag{2.42}$$

It is clear that g has no zeros in $\overline{\mathbb{D}}_r(0)$ and since $f \in H(\mathbb{D})$, also $g \in H(\mathbb{D})$. Thus, there exists a $\rho > r$ such that g has no zeros in the open disk $\mathbb{D}_\rho(0)$ and is holomorphic in $\mathbb{D}_\rho(0)$. It follows that $\log|g|$ is a harmonic function in $\mathbb{D}_\rho(0)$ (see e.g. [70, § 13.12]). Consequently, $\log|g|$ possesses the mean value property

$$\log|g(0)| = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log\left|g(re^{i\theta})\right| \, d\theta . \tag{2.43}$$

To determine the right hand side of (2.43), one easily verifies that the factors in (2.42) for $1 \leq n \leq M$ have modulus 1. For the remaining factors with $M + 1 \leq n \leq N$ holds

$$\frac{\alpha_n}{\alpha_n - z} = \frac{1}{1 - e^{i(\theta - \tau_n)}}$$

if we write $\alpha_n = r\mathrm{e}^{\mathrm{i}\tau_n}$ and $z = r\mathrm{e}^{\mathrm{i}\theta}$. Therewith (2.43) becomes

$$\log|g(0)| = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|f(r\mathrm{e}^{\mathrm{i}\theta})\right| \,\mathrm{d}\theta - \sum_{n=M+1}^{N} \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|1-\mathrm{e}^{\mathrm{i}(\theta-\tau_n)}\right| \,\mathrm{d}\theta \tag{2.44}$$

It is a consequence of Cauchy's theorem that

$$\tfrac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|1-\mathrm{e}^{\mathrm{i}\theta}\right| \,\mathrm{d}\theta = 0$$

(see e.g. [70, § 15.17]) which implies that the second term on the right hand side of (2.44) is zero. The definition (2.42) of the function g gives at once

$$|g(0)| = |f(0)| \prod_{n=1}^{M} \tfrac{r}{\alpha_n} \,.$$

Taking the exponential function of (2.43) shows finally (2.41). □

The next theorem proves a necessary condition on the zeros of a function f in order that $f \in H^p$ for some $1 \le p \le \infty$.

Theorem 2.16. *Let $f \in H^p$ with $1 \le p \le \infty$ be an analytic function in $\mathbb{D}$ with $f(0) \neq 0$, and let $\{\alpha_n\}_{n=1}^{\infty}$ be the zeros of f, listed according to their multiplicities. Then these zeros satisfy the* Blaschke condition

$$\sum_{n=1}^{\infty} (1-|\alpha_n|) < \infty \,. \tag{2.45}$$

Proof. If f has only finitely many zeros condition (2.45) is satisfied. Therefore, we assume that there are infinitely many zeros. Denote by $N(r)$ the number of zeros of f in the closed disk $\overline{\mathbb{D}}_r(0)$ for a radius $r < 1$. Fix $K \in \mathbb{N}$ and choose $r < 1$ such that $N(r) > K$. Then Jensen's formula gives

$$|f(0)| \prod_{n=1}^{K} \frac{r}{|\alpha_n|} \le |f(0)| \prod_{n=1}^{N(r)} \frac{r}{|\alpha_n|} = \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|f(r\mathrm{e}^{\mathrm{i}\theta})\right| \,\mathrm{d}\theta\right) < \infty$$

where the right hand side is bounded since $f \in H^p \subset H^1$. Thus there exists a constant $C_0 < \infty$ such that $\prod_{n=1}^{K} |\alpha_n| \ge r^K |f(0)|/C_0$. Since this inequality holds for arbitrary K, it is still valid for $K \to \infty$, i.e.

$$\prod_{n=1}^{\infty} |\alpha_n| \ge \frac{|f(0)|}{C_0} > 0 \,.$$

Now we define $u_n := 1 - |\alpha_n|$ for all $n \in \mathbb{N}$ and notice that the power series expansion of the exponential function implies $1 - x \le \exp(-x)$ for all $x \in \mathbb{R}$. Replacing x by u_n and multiplying the resulting inequalities gives

$$0 < \prod_{n=1}^{\infty} |\alpha_n| = \prod_{n=1}^{\infty} (1 - u_n) \leq \prod_{n=1}^{\infty} \exp(-u_n)$$
$$\leq \exp\left(-\sum_{n=1}^{\infty} u_n\right) = \exp\left(-\sum_{n=1}^{\infty} (1 - |\alpha_n|)\right)$$

which implies (2.45). □

Remark 2.17. Let $\{\alpha_n\}$ be a sequence in $\mathbb{D}$ with $|\alpha_n| = (n-1)/n$. This sequence does not satisfy the Blaschke condition (2.45) and therefore there exists no function f in any H^p-space ($1 \leq p \leq \infty$) with zeros at α_n. However, by the Weierstrass factorization theorem there exists a holomorphic function $f \in H(\mathbb{D})$ with zeros at α_n, but with $\|f\|_p = \infty$ for all $1 \leq p \leq \infty$. Conversely, if a function $f \in H^p$ is known to have zeros at $\{\alpha_n\}_{n\in\mathbb{N}}$, then this function has to be identically zero in $\mathbb{D}$.

So the Blaschke condition (2.45) is a necessary condition on the zeros of a holomorphic function f in order that f belongs to a Hardy space H^p. It even turns out that (2.45) is also a sufficient condition for the existence of a function $f \in H^p$, $1 \leq p \leq \infty$ which has zeros only at the prescribed points $\{\alpha_n\}_{n=1}^{\infty}$. The form of such a function is characterized in the following theorem.

Theorem 2.18 (Blaschke product). *Let $\{\alpha_n\}_{n=1}^{\infty}$ be a sequence of nonzero complex numbers in $\mathbb{D}$ such that $\{\alpha_n\}_{n=1}^{\infty}$ satisfies the Blaschke condition* (2.45)*. Let $k \geq 0$ be a nonnegative integer, and define the* Blaschke product

$$B(z) := z^k \prod_{n=1}^{\infty} \frac{|\alpha_n|}{\alpha_n} \frac{\alpha_n - z}{1 - \overline{\alpha}_n z}\,, \qquad z \in \mathbb{D}\,. \tag{2.46}$$

Then $B \in H^\infty \subset H^p$, $1 \leq p < \infty$, and B has zeros only at the points α_n and a zero of order k at 0*. Moreover, $|B(z)| < 1$ for all $z \in \mathbb{D}$ and $|B(\mathrm{e}^{\mathrm{i}\theta})| = 1$ almost everywhere.*

Remark 2.19. The term "Blaschke product" will be used for all functions of the form (2.46) even if the product contains only a finite number of factors and even if it contains no factor, i.e. even if $B(z) = 1$ for all $z \in \mathbb{D}$.

The function $B(z)$ in the above theorem is the product of z^k and of the factors

$$b_n(z) := \frac{|\alpha_n|}{\alpha_n} \frac{\alpha_n - z}{1 - \overline{\alpha}_n z}\,, \quad z \in \mathbb{D}\,. \tag{2.47}$$

Each factor b_n has a zero at $z = \alpha_n$ inside the unit disk $\mathbb{D}$, and a pole at $z = \overline{\alpha}_n^{-1}$ outside the closed unit disk $\overline{\mathbb{D}}$. Thus, each factor $b_n \in H(\mathbb{D})$ is a holomorphic function in $\mathbb{D}$ with precisely one zero at α_n. Moreover, it is easily verified that each factor has the properties that $|b_n(z)| < 1$ for all $z \in \mathbb{D}$ and that $|b_n| = 1$ for all $|z| = 1$. The Blaschke product is given by the infinite product $B(z) = z^k \prod_{n=1}^{\infty} b_n(z)$ of holomorphic functions. To prove Theorem 2.18, we basically have to show that this product converges uniformly to a holomorphic function. Therefore, as a preparation, the following Lemma studies conditions for the uniform convergence of infinite products.

Lemma 2.20. *Let $\{u_n\}_{n=1}^{\infty}$ be a sequence of bounded complex functions on a subset $S \subset \mathbb{D}$ of the unit disk such that the sum $\sum_{n=1}^{\infty} |u_n(z)|$ converges uniformly on S. Then the product*

$$f(z) = \prod_{n=1}^{\infty} [1 + u_n(z)]$$

converges uniformly on S. Moreover $f(z) = 0$ at some $z \in S$ if and only if $u_n(z) = -1$ for some $n \in \mathbb{N}$.

Proof. The assumption on $\{u_n\}$ implies that there exists a constant $C_0 < \infty$ such that $\sum_{n=1}^{\infty} |u_n(z)| \leq C_0$ for all $z \in S$. The power series expansion of the exponential function shows that $1 + x \leq \exp(x)$. Replacing x by $|u_1(z)|, |u_2(z)|, \ldots$ and multiplying the inequalities yields

$$\prod_{n=1}^{\infty} (1 + |u_n(z)|) \leq \exp\left(\sum_{n=1}^{\infty} |u_n(z)|\right) \leq \exp(C_0) =: C_1 < \infty \tag{2.48}$$

for all $z \in S$. Next, we define the partial products

$$p_N(z) := \prod_{n=1}^{N} [1 + u_n(z)] \qquad \text{and} \qquad q_N(z) := \prod_{n=1}^{N} [1 + |u_n(z)|]$$

and show that for every $N \in \mathbb{N}$

$$|p_N(z) - 1| \leq q_N(z) - 1\ . \tag{2.49}$$

For $N = 1$ this inequality is certainly satisfied. For $N > 1$, the statement is proved by induction. For $p_{N+1}(z)$ holds obviously

$$\begin{aligned} p_{N+1}(z) - 1 &= p_N(z)\ [1 + u_{N+1}(z)] - 1 \\ &= [p_N(z) - 1]\ [1 + u_{N+1}(z)] + u_{N+1}(z)\ . \end{aligned}$$

Therewith, it follows for the modulus

$$\begin{aligned} |p_{N+1}(z) - 1| &\leq |p_N(z) - 1|\ |1 + u_{N+1}(z)| + |u_{N+1}(z)| \\ &\leq (q_N(z) - 1)\ (1 + |u_{N+1}(z)|) + |u_{N+1}(z)| \\ &= q_{N+1}(z) - 1 \end{aligned}$$

where for the second line we used that (2.49) holds for p_N.

Now we apply (2.49) to (2.48). This shows that $|p_N(z)-1| \leq q_N(z)-1 \leq C_1-1$ for all $N \in \mathbb{N}$ and all $z \in S$. Taking $C_2 = C_1$, one obtains $|p_N(z)| \leq C_2$ for all $N \in \mathbb{N}$ and $z \in S$. Since $\sum_{n=1}^{\infty} |u_n(z)|$ is assumed to converge uniformly, for every $0 < \epsilon < 1/2$ there exists an $N_0 \in \mathbb{N}$ such that

$$\textstyle\sum_{n=N+1}^{\infty} |u_n(z)| < \epsilon \tag{2.50}$$

for all $N > N_0$ and for all $z \in S$. Now, for $M > N > N_0$ holds

$$\begin{aligned} |p_M(z) - p_N(z)| &= |p_N(z)| \left| \prod_{n=N+1}^{M} [1 + u_n(z)] - 1 \right| \\ &\leq |p_N(z)| \left(\prod_{n=N+1}^{M} [1 + |u_n(z)|] - 1 \right) \\ &\leq |p_N(z)| \, (\mathrm{e}^{\epsilon} - 1) \leq 2\, |p_N(z)| \, \epsilon \leq 2\, C_2\, \epsilon \end{aligned} \tag{2.51}$$

where the first inequality follows from (2.49), whereas the last line from (2.48), (2.50), and from the uniform boundedness of p_N. This last result shows that p_N is a Cauchy sequence in S which converges uniformly to a limit function f on S.

Finally (2.51) implies $|p_M(z)| \geq (1-2\epsilon)|p_{N_0}(z)|$ for all $M > N_0$. It follows for the limit function f that

$$|f(z)| \geq (1 - 2\epsilon)|p_{N_0}(z)|$$

for all $z \in S$. This shows that $f(z) = 0$ if and only if $p_{N_0}(z) = 0$, i.e. if and only if $u_n(z) = -1$ for some n. □

With this we are able to prove Theorem 2.18.

Proof (Theorem 2.18). Without loss of generality, we assume that $k = 0$. Again, we define the individual factors b_n of the Blaschke product by (2.47), and consider the term $1 - b_n$ inside the unit disk. Adding the term $1/|\alpha_n| - 1/|\alpha_n|$ followed by a straight forward rearranging yields

$$\begin{aligned} 1 - b_n(z) &= \frac{1}{|\alpha_n|} \left(1 - \frac{|\alpha_n|^2 - \overline{\alpha}_n z}{1 - \overline{\alpha}_n z} \right) - \frac{1 - |\alpha_n|}{|\alpha_n|} \\ &= (1 - |\alpha_n|) \, \frac{|\alpha_n| + \overline{\alpha}_n z}{(1 - \overline{\alpha}_n z)\, |\alpha_n|} \; . \end{aligned}$$

Remembering that all zeros α_n are inside the unit disk $\mathbb{D}$ gives

$$|1 - b_n(z)| \leq (1 - |\alpha_n|) \, \frac{1 + r}{1 - r} \, , \quad \text{for all } |z| \leq r < 1$$

and consequently

$$\sum_{n=1}^{\infty} |1 - b_n(z)| \le \frac{1+r}{1-r} \sum_{n=1}^{\infty} (1 - |\alpha_n|) \,.$$

This shows that $\sum_{n=1}^{\infty} |1 - b_n(z)|$ converges uniformly on compact subsets of $\mathbb{D}$ since the zeros α_n satisfy the Blaschke condition (2.45).

Setting $u_n(z) = b_n(z) - 1$, Lemma 2.20 implies that $B(z) = \prod_{n=1}^{\infty} b_n(z)$ converges uniformly on compact subsets of $\mathbb{D}$ and that $B(z) = 0$ if and only if $B_n(z) = 0$ for some $n \in \mathbb{N}$. Since every b_n is holomorphic in $\mathbb{D}$ and since B converges uniformly on compact subsets of $\mathbb{D}$, $B(z)$ is also holomorphic in $\mathbb{D}$ (see e.g. [70, Theorem 10.28]). Moreover, since each factor b_n has absolute value less than 1 in $\mathbb{D}$, it follows that $|B(z)| < 1$ for all $z \in \mathbb{D}$, and consequently that $B \in H^\infty$ with $\|B\|_\infty \le 1$.

Since $B \in H^\infty$, the boundary function $B(\mathrm{e}^{\mathrm{i}\theta})$ exists almost everywhere for $\theta \in [-\pi, \pi)$, and since $B(z)$ has absolute value smaller that 1 for all $z \in \mathbb{D}$, the boundary function has to satisfy $|B(\mathrm{e}^{\mathrm{i}\theta})| \le 1$ almost everywhere. Now, let $B_N(z) = \prod_{n=1}^{N} b_n(z)$ the partial product of B. Then $B(z)/B_N(z)$ is again a Blaschke product, and thus holomorphic in $\mathbb{D}$. Consequently, it satisfies the mean value property. Together with the triangle inequality, one has

$$\frac{B(0)}{B_N(0)} \le \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{B(\mathrm{e}^{\mathrm{i}\theta})}{B_N(\mathrm{e}^{\mathrm{i}\theta})} \right| \mathrm{d}\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| B(\mathrm{e}^{\mathrm{i}\theta}) \right| \mathrm{d}\theta$$

where the last equality is a consequence of $|B_N(\mathrm{e}^{\mathrm{i}\theta})| = 1$ for all $\theta \in [-\pi, \pi)$. Now, letting $N \to \infty$, we obtain that $1 \le \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| B(\mathrm{e}^{\mathrm{i}\theta}) \right| \mathrm{d}\theta$. Consequently, since $|B(\mathrm{e}^{\mathrm{i}\theta})| \le 1$, one gets $|B(\mathrm{e}^{\mathrm{i}\theta})| = 1$ almost everywhere. □

Thus, given a function $f \in H^p$ with $1 \le p \le \infty$ with zeros at $\{\alpha_n\}_{n \in \mathbb{N}}$ (which will satisfy the Blaschke condition (2.45)), we can form the Blaschke product $B \in H^\infty$ with the zeros of f. Now we can try to divide out the zeros of f by dividing f by the corresponding Blaschke product B. Of course, the resulting quotient $g := f/B$ is again a holomorphic function in $\mathbb{D}$, and since B has absolute value 1 almost everywhere on the unit circle, we even expect that g may have the same H^p-norm as the original f. That this reasoning is indeed true is shown by the next theorem.

Theorem 2.21. *Let $f \in H^p$ with $1 \le p \le \infty$, let B be the Blaschke product* (2.46) *formed with the zeros of f, and set $g(z) := f(z)/B(z)$, $z \in \mathbb{D}$. Then $g \in H^p$ with $\|g\|_p = \|f\|_p$.*

Proof. Let $\{\alpha_n\}_{n=1}^{\infty}$ be the sequence of zeros of f in $\mathbb{D}$, and let $b_n(z)$ be the factor of the Blaschke product corresponding to the zero α_n as defined in (2.47). Moreover, let

$$B_N(z) = \prod_{n=1}^{N} b_n(z) \,, \qquad z \in \mathbb{D}$$

be the partial Blaschke product formed by the first N zeros of f, and let $g_N = f/B_N$. For every fixed N, $B_N(r\,\mathrm{e}^{\mathrm{i}\theta}) \to 1$ uniformly as $r \to 1$. It follows that $g_N(r\,\mathrm{e}^{\mathrm{i}\theta}) \to f(\mathrm{e}^{\mathrm{i}\theta})$ and consequently that

$$\|g_N\|_p = \|f\|_p \,. \tag{2.52}$$

Since $|b_n(z)| < 1$ for all $z \in \mathbb{D}$ and all n, we have that

$$0 \le |g_1(z)| \le |g_2(z)| \le \cdots \le \infty \quad \text{and} \quad |g_n(z)| \to |g(z)|$$

for every $z \in \mathbb{D}$. Fixing $0 < r < 1$, set $g_r(z) := g(r\,z)$ and $(g_N)_r(z) := g_N(r\,z)$, and applying Lebesgue's monotone convergence theorem, one gets

$$\lim_{N\to\infty} \|(g_N)_r\|_p^p = \lim_{N\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} |g_N(r\,\mathrm{e}^{\mathrm{i}\theta})|^p \,\mathrm{d}\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} |g(r\,\mathrm{e}^{\mathrm{i}\theta})|^p \,\mathrm{d}\theta = \|g_r\|_p^p \,.$$

Since g_N is analytic in $\mathbb{D}$ and because of (2.52), the left hand side is upper bounded by $\|f\|_p^p$ for every $0 < r < 1$. Leting $r \to 1$, we obtain $\|g\|_p \le \|f\|_p$. However, since $|B(z)| \le 1$ for all $z \in \mathbb{D}$, we also have that $|g(z)| \ge |f(z)|$ for all $z \in \mathbb{D}$, which shows that we even have equality, i.e. that $\|g\|_p = \|f\|_p$. □

2.2.3 Inner-Outer factorization

The last theorem showed that every function $f \in H^p$ can be factorized into a Blaschke product and a function without zeros in the unit disk. This section considers a somewhat different factorization of functions in H^p into so called inner and outer functions.

Definition 2.22 (Inner and Outer functions). *An* inner function *is an analytic function $f \in H^\infty$ such that $|f(z)| \le 1$ in the unit disk and such that $|f(\mathrm{e}^{\mathrm{i}\theta})| = 1$ almost everywhere on the unit circle.*
An outer function *is an analytic function O in the unit disk $\mathbb{D}$ of the form*

$$O(z) = c\,\exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\phi(\mathrm{e}^{\mathrm{i}\tau})\,\frac{\mathrm{e}^{\mathrm{i}\tau}+z}{\mathrm{e}^{\mathrm{i}\tau}-z}\,\mathrm{d}\tau\right) \,. \tag{2.53}$$

Here c is a constant with $|c| = 1$, and ϕ is a positive measurable function on $\mathbb{T}$ such that $\log\phi \in L^1$.

Every Blaschke product is an inner function by Theorem 2.18. However, there exist other inner functions. The following theorem characterizes all inner functions as the product of a Blaschke product and a so-called singular function.

Theorem 2.23. *Let $f \in H^p$ be an inner function and let B be the Blaschke product formed with the zeros of f. Then there exists a positive Borel measure μ on $\mathbb{T}$ which is singular with respect to Lebesgue measure and a complex constant c with $|c| = 1$ such that*

$$f(z) = B(z)\,S(z)\,, \qquad z \in \mathbb{D} \tag{2.54}$$

with

$$S(z) = c\,\exp\left(-\int_{-\pi}^{\pi} \frac{\mathrm{e}^{\mathrm{i}\tau}+z}{\mathrm{e}^{\mathrm{i}\tau}-z}\,\mathrm{d}\mu(\tau)\right)\,, \qquad z \in \mathbb{D}\,. \tag{2.55}$$

We will call the function S in the factorization (2.54) a *singular function*. Thus, an inner factor is the product of a Blaschke product B and of a singular function S, in general. Nevertheless, it may happen that B, or S, or even both factors are identically 1.

Example 2.24. Probably the simplest singular function is obtained by taking μ in (2.55) to be the unit mass at $\tau = 0$ and by letting $c = 1$. This yields the singular function

$$S(z) = \exp\left(\frac{z+1}{z-1}\right), \qquad z \in \mathbb{D}\,.$$

It is a holomorphic function in $\mathbb{D}$ with an essential singularity at $z = 1$.

Proof (Theorem 2.23). Let $g := f/B$, then g is a holomorphic function without any zeros in $\mathbb{D}$, from which follows that $\log|g|$ is a harmonic function in $\mathbb{D}$ (see e.g. [70, Theorem 13.12]). By Theorem 2.21 it follows that $|g(z)| \le 1$ for $z \in \mathbb{D}$ and that $|g(\mathrm{e}^{\mathrm{i}\theta})| = 1$ almost everywhere, which implies that $\log|g| \le 0$ in $\mathbb{D}$ and $\log|g(\mathrm{e}^{\mathrm{i}\theta})| = 0$ a.e. on $\mathbb{T}$. It is known that every bounded harmonic function in $\mathbb{D}$ can be represented by the Poisson integral of a unique Borel measure on $\mathbb{T}$ (see e.g. [70, Theorem 11.30]). We conclude for our case that $\log|g|$ is the Poisson integral of $-\mathrm{d}\mu$ with some positive Borel measure μ on $\mathbb{D}$. However, since $\log|g(\mathrm{e}^{\mathrm{i}\theta})| = 0$ a.e. on $\mathbb{T}$ the measure μ has to be singular (with respect to Lebesgue measure). Now $\log|g|$, as the Poisson integral of $-\mathrm{d}\mu$, is the real part of the function

$$G(z) = -\int_{-\pi}^{\pi} \frac{\mathrm{e}^{\mathrm{i}\tau} + z}{\mathrm{e}^{\mathrm{i}\tau} - z}\,\mathrm{d}\mu(\tau)$$

(see e.g. Section 5.1) which implies that S has the form (2.55). □

The previous theorem clarified the general form of an inner function whereas the general form of an outer function is given by (2.53). The next theorem studies basic properties of outer functions needed frequently throughout this book.

Theorem 2.25. *Let O_ϕ be an outer function related to a positive (real valued) measurable function ϕ as in Definition 2.22. Then*

(*a*) $\log|O_\phi|$ *is the Poisson integral of* $\log\phi$.
(*b*) $\lim_{r\to 1}|O_\phi(r\mathrm{e}^{\mathrm{i}\theta})| = \phi(\mathrm{e}^{\mathrm{i}\theta})$ *for almost all* $\theta \in [-\pi, \pi)$.
(*c*) $O_\phi \in H^p$ *if and only if* $\phi \in L^p$ *and* $\|O_\phi\|_p = \|\phi\|_p$.

Proof. Statement (*a*) follows from the definition of the outer function (2.53) since the exponent of O_ϕ can be written as

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\phi(\mathrm{e}^{\mathrm{i}\tau})\,\frac{\mathrm{e}^{\mathrm{i}\tau} + r\mathrm{e}^{\mathrm{i}\theta}}{\mathrm{e}^{\mathrm{i}\tau} - r\mathrm{e}^{\mathrm{i}\theta}}\,\mathrm{d}\tau = (\mathfrak{P}\log\phi)(r\mathrm{e}^{\mathrm{i}\theta}) + \mathrm{i}\,(\mathfrak{Q}\log\phi)(r\mathrm{e}^{\mathrm{i}\theta})$$

with the Poisson integral $\mathfrak{P}\log\phi$ and the conjugate Poisson integral $\mathfrak{Q}\log\phi$ of $\log\phi$ (cf. Section 5.1). It follows that $|O_\phi| = \exp(\mathfrak{P}\log\phi)$ which proves (a).

The Poisson integral $(\mathfrak{P}f)(re^{i\theta})$ of a function $f \in L^1$ converges to f in the L^1-norm as $r \to 1$ (see e.g. Theorem 5.3 for a proof). By this property of the Poisson integral (a) implies (b).

Applying statement (b), we have

$$\begin{aligned}\|\phi\|_p^p &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |\phi(e^{i\theta})|^p \, d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \lim_{r\to 1} |O_\phi(re^{i\theta})|^p \, d\theta \\ &\leq \lim_{r\to 1}\frac{1}{2\pi}\int_{-\pi}^{\pi} |O_\phi(re^{i\theta})|^p \, d\theta = \|O_\phi\|_p^p\end{aligned}$$

where the inequality follows from Fatou's Lemma (see e.g. [70, § 1.28]). Thus $\|\phi\|_p \leq \|O\phi\|_p$. For the converse assume that $\phi \in L^p$. Then Jensen's inequality (cf. [70, §3.3]) gives

$$\begin{aligned}|O_\phi(re^{i\theta})|^p &= \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\phi^p(e^{i\tau})\,\mathcal{P}_r(\theta-\tau)\,d\tau\right) \\ &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi^p(e^{i\tau})\,\mathcal{P}_r(\theta-\tau)\,d\tau\end{aligned}$$

in which the left hand side of the inequality is just the Poisson integral $\mathfrak{P}\log\phi^p$ with the Poisson kernel $\mathcal{P}_r$ (cf. also (5.5) and (5.4) for the definition). Integration of the last inequality with respect to θ and using that the Poisson kernel $\mathcal{P}_r$ satisfies $\int_{-\pi}^{\pi}\mathcal{P}_r(\theta)\,d\theta = 1$ (cf. Section 5.2) gives $\|O_\phi\|_p \leq \|\phi\|_p$. This finishes the proof of (c). □

Finally, the following theorem will give the desired factorization result under point (c). It shows that every function $f \in H^p$ can always be factorized into an inner and an outer function.

Theorem 2.26. *For $1 \leq p \leq \infty$ let $f \in H^p$ be a nonzero function. Then*

(a) $\log|f| \in L^1$.
(b) *the outer function defined by*

$$O_f(z) = \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|f(e^{i\tau})\right| \frac{e^{i\tau}+z}{e^{i\tau}-z}\,d\tau\right), \qquad z \in \mathbb{D}$$

is an element of H^p.
(c) *there exists an inner function I_f such that $f = O_f\, I_f$.*

Proof. We consider first the case $p = 1$. Assume that $f \in H^1$, let B be the Blaschke product (2.46) formed with the zeros of f, and set $g = f/B$. By Theorem 2.21 $g \in H^1$ and $|g(e^{i\theta})| = |f(e^{i\theta})|$ for almost all $\theta \in [-\pi, \pi)$. Therefore, it is sufficient to prove the theorem for g instead of f. Since g is holomorphic without any zero in $\mathbb{D}$ the function $\log|g|$ is harmonic in $\mathbb{D}$ (see e.g. [70, § 13.12]) and therefore it satisfies (2.43). Since $g(0) \neq 0$ we assume, without loss of generality, that $g(0) = 1$ and define the two functions

$$\log^+ x := \begin{cases} 0 & , \quad x < 1 \\ \log x \, , & \quad x \geq 1 \end{cases} \qquad \text{and} \qquad \log^- x := \begin{cases} \log(1/x) \, , & \quad x < 1 \\ 0 & , \quad x \geq 1 \end{cases}$$

on the positive real axis, such that obviously $\log x = \log^+ x - \log^- x$. Therewith, it follows from (2.43) that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log^- |g(\mathrm{e}^{\mathrm{i}\theta})| \,\mathrm{d}\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log^+ |g(\mathrm{e}^{\mathrm{i}\theta})| \,\mathrm{d}\theta \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |g(\mathrm{e}^{\mathrm{i}\theta})| \,\mathrm{d}\theta = \|g\|_1 \, .$$

This shows that $\log^+|g|$ and $\log^-|g|$ are in L^1, such that $\log|g| \in L^1$, which proves (a). It follows that O_f is a well defined outer function, and Theorem 2.25 (c) implies that $O_f \in H^1$, proving (b).

It remains to show (c). To this end we show next that $|g(z)| \leq |O_g(z)|$ for all $z \in \mathbb{D}$. Since we know from Theorem 2.25 that $\log|O_g|$ is equal to the Poisson integral of $\log|g|$, we have to show

$$\log|g(z)| \leq \log|O_g(z)| = (\mathfrak{P} \log|g|)(z) \, . \tag{2.56}$$

For $0 < r < 1$ and $z \in \mathbb{D}$ we define $g_r(z) := g(r\,z)$. Since g is a holomorphic function without zeros in $\mathbb{D}$, $\log|g_r|$ is harmonic in $\mathbb{D}$ (see e.g. [70, Theorem 13.12]) and can be represented as a Poisson integral. We therefore have

$$\log|g_r(z)| = \mathfrak{P}\left[\log|g_r|\right](z) = \mathfrak{P}\left[\log^+|g_r|\right](z) - \mathfrak{P}\left[\log^-|g_r|\right](z) \, . \tag{2.57}$$

We know from Theorem 2.11 that $g_r(\mathrm{e}^{\mathrm{i}\theta}) \to g(\mathrm{e}^{\mathrm{i}\theta})$ as $r \to 1$. It follows that the left hand side of (2.57) converges to $\log|g|$ and that the first term on the right hand side converges to $\mathfrak{P}\left[\log^+|g|\right](z)$ as $r \to 1$. This last statement follows from

$$\begin{aligned} &\left|\mathfrak{P}\left[\log^+|g_r|\right](\rho\,\mathrm{e}^{\mathrm{i}\theta}) - \mathfrak{P}\left[\log^+|g|\right](\rho\,\mathrm{e}^{\mathrm{i}\theta})\right| \\ &\qquad \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\log^+|g_r(\mathrm{e}^{\mathrm{i}\tau})| - \log^+|g(\mathrm{e}^{\mathrm{i}\tau})|\right| \mathcal{P}_\rho(\theta-\tau)\,\mathrm{d}\tau \\ &\qquad\quad \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \left||g_r(\mathrm{e}^{\mathrm{i}\tau})| - |g(\mathrm{e}^{\mathrm{i}\tau})|\right| \mathcal{P}_\rho(\theta-\tau)\,\mathrm{d}\tau \end{aligned}$$

and from the fact that $g_r(\mathrm{e}^{\mathrm{i}\theta}) \to g(\mathrm{e}^{\mathrm{i}\theta})$. Here $\mathcal{P}_\rho$ denotes the Poisson kernel (see (5.5) and (5.4) for the definition), and the last line was obtained using the relation $|\log^+ u - \log^+ v| \leq |u - v|$ for all real numbers u, v, which may easily be verified. Therewith, letting $r \to 1$ in (2.57), one obtains

$$\mathfrak{P}\left[\log^-|g|\right](z) \leq \lim_{r\to 1} \mathfrak{P}\left[\log^-|g_r|\right](z) = \mathfrak{P}\left[\log^+|g_r|\right](z) - \log|g(z)|$$

where the first inequality follows from Fatou's Lemma (see e.g. [70, § 1.28]). Combining $\log^+$ and $\log^-$ to log, one obtains the desired relation (2.56), which proves that $|g(z)| \leq |O_g(z)|$ for all $z \in \mathbb{D}$.

Now we define the function

$$I_g(z) := \frac{g(z)}{O_g(z)} \qquad z \in \mathbb{D} .$$

Obviously, I_g is analytic in $\mathbb{D}$, $|I_g(z)| \leq 1$ for all $z \in \mathbb{D}$ and $|I_g(e^{i\theta})| = 1$ almost everywhere. Thus I_g is an inner function. □

By Theorem 2.23, every inner function can be written as the product of a Blaschke product and an singular function. Consequently, it follows form point (c) of the previous theorem that every $f \in H^p$ may be written as

$$f(z) = O_f(z)\, B_f(z)\, S_f(z) \,, \qquad z \in \mathbb{D}$$

with an outer function O_f, the Blaschke product B_f formed with the zeros of f, and a singular function S_f.

2.3 Vector-valued Hardy Spaces

The previous section introduced the Hardy space of complex valued functions. In general, it is possible to extend the concept of Hardy spaces to functions taking values in arbitrary Banach spaces. To give a completely satisfactory definition of such spaces, one needs some results from the integration theory of functions with values in Banach spaces. Although this is a straight forward generalization of the standard integration of complex valued Lebesgue measurable functions, it would be out of the scope of our intentions here. However, for the case of functions with values in a separable or even a finite dimensional Hilbert spaces, almost the whole theory can be led back to the scalar case of the previous section. Therefore, we shall restrict ourselves to these cases. Later, we will be especially interested in the finite dimensional case, since this is the suitable framework for modeling linear systems with a finite number of inputs and outputs. Nevertheless, the basic definitions are given for the slightly more general case of separable Hilbert spaces.

We start with a formal extension of ℓ^p and L^p spaces to the case of vector valued functions. To emphasize the difference to the scalar case, vector valued functions will be denoted by bold face letters.

Definition 2.27 (Vector-valued ℓ^p spaces). *Let $\mathcal{H}$ be a separable Hilbert space and let $\hat{\boldsymbol{f}} = \{\hat{\boldsymbol{f}}(k)\}_{k=-\infty}^{\infty}$ be a double infinite sequence of elements from $\mathcal{H}$. For $1 \leq p < \infty$ and $p = \infty$ define*

$$\|\hat{\boldsymbol{f}}\|_{\ell^p} := \left(\sum_{k=-\infty}^{\infty} \|\hat{\boldsymbol{f}}(k)\|_{\mathcal{H}}^p \right)^{1/p} \qquad \textit{and} \quad \|\hat{\boldsymbol{f}}\|_{\ell^\infty} := \sup_{k \in \mathbb{Z}} \|\hat{\boldsymbol{f}}(k)\|_{\mathcal{H}}$$

respectively.
Then, for $1 \leq p \leq \infty$ *the set* $\ell^p(\mathcal{H})$ *denotes the set of all double infinite sequences* $\hat{\boldsymbol{f}} = \{\hat{\boldsymbol{f}}(k)\}_{k=-\infty}^{\infty}$ *with values in* $\mathcal{H}$ *for which* $\|\hat{\boldsymbol{f}}\|_{\ell^p} < \infty$ *and* $\ell_+^p(\mathcal{H})$ *denotes the set of all infinite sequences* $\hat{\boldsymbol{f}} = \{\hat{\boldsymbol{f}}(k)\}_{k=0}^{\infty}$ *in* $\mathcal{H}$ *with* $\|\hat{\boldsymbol{f}}\|_{\ell^p} < \infty$*.*

Of course, $\ell_+^p(\mathcal{H})$ my be considered as the subspace of $\ell^p(\mathcal{H})$ in which for all elements $\hat{\boldsymbol{f}}$ holds that $\hat{\boldsymbol{f}}(k) = 0$ for all $k < 0$. Moreover, it is clear that for the special case $\mathcal{H} = \mathbb{C}$, one obtains again the usual ℓ^p spaces. As in the scalar case, $\ell^2(\mathcal{H})$ is a Hilbert space with the inner product

$$\langle \hat{\boldsymbol{f}}, \hat{\boldsymbol{g}} \rangle_{\ell^2(\mathcal{H})} = \sum_{k=-\infty}^{\infty} \langle \hat{\boldsymbol{f}}(k), \hat{\boldsymbol{g}}(k) \rangle_{\mathcal{H}} \, .$$

Definition 2.28 (Vector-valued L^p spaces). *Let* $\mathcal{H}$ *be a separable Hilbert space and let* f *be a measurable function with values in* $\mathcal{H}$*. For* $1 \leq p < \infty$ *define*

$$\|\boldsymbol{f}\|_p := \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \|\boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta})\|_{\mathcal{H}}^p \, \mathrm{d}\theta \right)^{1/p}$$

and for $p = \infty$ *define*

$$\|\boldsymbol{f}\|_\infty := \operatorname*{ess\,sup}_{\zeta \in \mathbb{T}} \|\boldsymbol{f}(\zeta)\|_{\mathcal{H}} \, .$$

Then for $1 \leq p \leq \infty$ *the set* $L^p(\mathcal{H})$ *denotes the set of all measurable functions* $\boldsymbol{f}$ *with values in* $\mathcal{H}$ *for which* $\|\boldsymbol{f}\|_p < \infty$*.*

Of course, if the dimension of the Hilbert space $\mathcal{H}$ is one, the above definition of $L^p(\mathcal{H})$ coincides with the usual L^p-spaces on the unit circle, and $L^2(\mathcal{H})$ is a Hilbert space with the inner product

$$\langle \boldsymbol{f}, \boldsymbol{g} \rangle_{L^2(\mathcal{H})} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \langle \boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta}), \boldsymbol{g}(\mathrm{e}^{\mathrm{i}\theta}) \rangle_{\mathcal{H}} \, \mathrm{d}\theta \, .$$

Since $\mathcal{H}$ is assumed to be separable there exists a complete orthonormal basis $\{\boldsymbol{e}_n\}_{n=1}^{\infty}$ in $\mathcal{H}$ such that every $\boldsymbol{f} \in \mathcal{H}$ can be written as $\boldsymbol{f} = f_1 \, \boldsymbol{e}_1 + f_2 \, \boldsymbol{e}_2 + \cdots + f_n \, \boldsymbol{e}_n + \cdots$ where $f_n := \langle \boldsymbol{f}, \boldsymbol{e}_n \rangle_{\mathcal{H}}$ are the components of $\boldsymbol{f}$ with respect to the basis $\{\boldsymbol{e}_n\}_{n=1}^{\infty}$ and the norm of $\boldsymbol{f}$ in $\mathcal{H}$ is just the ℓ^2-norm of the sequence $(f_1, f_2, \cdots)$ of its components: $\|\boldsymbol{f}\|_{\mathcal{H}}^2 = \sum_{n=1}^{\infty} |f_n|^2$. Moreover, given a sequence $\hat{\boldsymbol{f}} = \{\hat{\boldsymbol{f}}(k)\}_{k=-\infty}^{\infty}$ with values in the Hilbert space $\mathcal{H}$, we can define its *coordinate sequences* $\hat{f}_n = \{\hat{f}_n(k)\}_{k=-\infty}^{\infty}$ with $n = 1, 2, \ldots$ by

$$\hat{f}_n(k) = \langle \hat{\boldsymbol{f}}(k), \boldsymbol{e}_n \rangle_{\mathcal{H}} \, , \qquad k \in \mathbb{Z} \, , \, n \in \mathbb{N} \, .$$

Similarly, given a function $\boldsymbol{f}$ on the unit circle with values in $\mathcal{H}$, its *coordinate functions* f_n, $n = 1, 2, \ldots$ are defined by

$$f_n(\zeta) = \langle \boldsymbol{f}(\zeta), \boldsymbol{e}_n \rangle_{\mathcal{H}}, \qquad \zeta \in \mathbb{T}, n \in \mathbb{N}.$$

The following proposition gives a characterization of the spaces $\ell^p(\mathcal{H})$ and $L^p(\mathcal{H})$ in terms of the individual coordinates. It will be show that $\ell^p(\mathcal{H})$ is equivalent to the set of all sequences $\hat{\boldsymbol{f}} = \{\hat{f}_n\}_{n=1}^{\infty}$ whose individual entries $\hat{f}_n = \{\hat{f}_n(k)\}_{k=-\infty}^{\infty}$ belong to ℓ^p and it will be show that $L^p(\mathcal{H})$ is precisely the set of all sequences $\boldsymbol{f} = \{f_n\}_{n=1}^{\infty}$ whose individual components f_n are elements of L^p.

Proposition 2.29. *Let $\mathcal{H}$ be a separable Hilbert space with an arbitrary orthonormal basis $\{\boldsymbol{e}_n\}_{n=1}^{\infty}$ and let $1 \le p \le \infty$.*
A sequence $\hat{\boldsymbol{f}}$ of elements in $\mathcal{H}$ belongs to $\ell^p(\mathcal{H})$ if and only if all coordinate sequences $\hat{f}_n = \langle \hat{\boldsymbol{f}}, \boldsymbol{e}_n \rangle_{\mathcal{H}}$, $n \in \mathbb{N}$ belong to ℓ^p.
A function $\boldsymbol{f}$ on the unit circle $\mathbb{T}$ and with values in $\mathcal{H}$ belongs to $L^p(\mathcal{H})$ if and only if all coordinate functions $f_n = \langle \boldsymbol{f}, \boldsymbol{e}_n \rangle$, $n \in \mathbb{N}$ belong to L^p.

Proof. We prove the statement for $L^p(\mathcal{H})$. By the identification of $\mathcal{H}$ with ℓ^2 and with the triangle inequality, one has

$$\|\boldsymbol{f}(\zeta)\|_{\mathcal{H}} = \left(\sum_{n=1}^{\infty} |f_n(\zeta)|^2\right)^{1/2} \le \sum_{n=1}^{N} |f_n(\zeta)| \qquad \text{for every } \zeta \in \mathbb{T}$$

and provided that the right hand side exists. Therewith one gets at once $\|\boldsymbol{f}\|_\infty \le \sum_{n=1}^{\infty} \|f_n\|_\infty$. For $p < \infty$ we take both sides to the power p, integrate over the unit circle $\mathbb{T}$, and apply Minkowski's inequality to the right hand side integral. This gives $\|\boldsymbol{f}\|_p \le \sum_{n=1}^{\infty} \|f_n\|_p$ which proves the "if" part of the proposition.
To verify the "only if" part, note that by the identification of $\mathcal{H}$ with ℓ^2 one has that $\|\boldsymbol{f}(\zeta)\|_{\mathcal{H}}^p \ge |f_n(\zeta)|^p$ for all $\zeta \in \mathbb{T}$ and for all $n = \mathbb{N}$. This gives immediately $\|\boldsymbol{f}\|_\infty \ge \|f_n\|_\infty$ and for $p < \infty$, the integration of both sides, gives $\|\boldsymbol{f}\|_p \ge \|f_n\|_p$ for every $1 \le n \le N$.
The analogous proof for $\ell^p(\mathcal{H})$ is left as an exercise. □

As in the scalar case, we want to consider the Fourier series expansion of functions in $L^p(\mathcal{H})$. To avoid the introduction of the integration over functions with values in Hilbert spaces, the Fourier series are introduced in the weak sense, as follows: Let $\boldsymbol{f} \in L^1(\mathcal{H})$, we want to write $\boldsymbol{f}$ in the form

$$\boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{k=-\infty}^{\infty} \hat{\boldsymbol{f}}(k)\, \mathrm{e}^{\mathrm{i}k\theta} \tag{2.58}$$

where $\{\hat{\boldsymbol{f}}(k)\}_{k=-\infty}^{\infty}$ is a sequence of elements from the Hilbert space $\mathcal{H}$. We call this expansion Fourier series of $\boldsymbol{f}$ if for every $\boldsymbol{g} \in \mathcal{H}$

$$\left\langle \boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta}), \boldsymbol{g} \right\rangle_{\mathcal{H}} = \sum_{k=-\infty}^{\infty} \left\langle \hat{\boldsymbol{f}}(k), \boldsymbol{g} \right\rangle_{\mathcal{H}} \mathrm{e}^{\mathrm{i}k\theta}$$

is an ordinary Fourier series of the complex valued function $\langle \boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta}), \boldsymbol{g}\rangle_{\mathcal{H}}$. The coefficients of the scalar Fourier series are of course given by (2.1), thus

$$\langle \hat{\boldsymbol{f}}(k), \boldsymbol{g}\rangle_{\mathcal{H}} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \langle \boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta}), \boldsymbol{g}\rangle_{\mathcal{H}} \mathrm{e}^{-\mathrm{i}k\theta}\,\mathrm{d}\theta\,. \tag{2.59}$$

It may not be immediately clear whether there exists such a sequence $\{\hat{\boldsymbol{f}}(k)\}_{k=-\infty}^{\infty}$ of vectors in $\mathcal{H}$ such that (2.59) is satisfied for all $\boldsymbol{g} \in \mathcal{H}$. However, for a fixed $\boldsymbol{f} \in L^1(\mathcal{H})$ and $k \in \mathbb{Z}$ the right hand side of (2.59) defines a conjugate-linear functional Φ_k on $\mathcal{H}$ which is also continuous, since by the Cauchy-Schwarz inequality

$$\begin{aligned}|\Phi_k(\boldsymbol{g})| &\le \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\langle \boldsymbol{f}(\mathrm{e}^{\mathrm{i}\tau}), \boldsymbol{g}\rangle_{\mathcal{H}}\right|\,\mathrm{d}\tau \\ &\le \frac{1}{2\pi}\int_{-\pi}^{\pi} \|\boldsymbol{f}(\mathrm{e}^{\mathrm{i}\tau})\|_{\mathcal{H}}\,\|\boldsymbol{g}\|_{\mathcal{H}}\,\mathrm{d}\tau = \|\boldsymbol{f}\|_1\,\|\boldsymbol{g}\|_{\mathcal{H}}\,.\end{aligned}$$

But this implies by the Riesz representation theorem (for Hilbert spaces) that there exists a unique $\hat{\boldsymbol{f}}(k) \in \mathcal{H}$ such that (2.59) holds for all $\boldsymbol{g} \in \mathcal{H}$. To determine the coefficient vectors $\hat{\boldsymbol{f}}(k)$ in the Fourier series (2.58), one may choose an orthonormal basis $\{\boldsymbol{e}_n\}$ in $\mathcal{H}$ and write every coefficient vector in this basis as $\hat{\boldsymbol{f}}(k) = \sum_{l\in\mathbb{N}} \hat{f}_l(k)\,\boldsymbol{e}_l$. Plugging this representation into (2.59) together with a $\boldsymbol{g} = \boldsymbol{e}_n$ gives

$$\hat{f}_n(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_n(\mathrm{e}^{\mathrm{i}\tau})\,\mathrm{e}^{-\mathrm{i}k\tau}\,\mathrm{d}\tau$$

where $f_n(\mathrm{e}^{\mathrm{i}\tau}) = \langle \boldsymbol{f}(\mathrm{e}^{\mathrm{i}\tau}), \boldsymbol{e}_n\rangle_{\mathcal{H}}$ is the n-th coordinate of $\boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta})$ with respect to the orthonormal basis $\{\boldsymbol{e}_l\}$. Thus, if an orthonormal basis in $\mathcal{H}$ is fixed, the Fourier coefficients $\hat{\boldsymbol{f}}(k)$ of the series (2.58) can be determined componentwise. Therefore, we will simply write

$$\hat{\boldsymbol{f}}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \boldsymbol{f}(\mathrm{e}^{\mathrm{i}\tau})\,\mathrm{e}^{-\mathrm{i}k\tau}\,\mathrm{d}\tau\,, \qquad k = 0, \pm 1, \pm 2, \ldots \tag{2.60}$$

where the integration on the right hand side means a component wise integration of every coordinate with respect to the chosen orthonormal basis. The above discussion holds in particular for the case $\mathcal{H} = \mathbb{C}^N$ with the usual orthonormal basis $\boldsymbol{e}_1 = \{1, 0, 0, \cdots, 0\}$, $\boldsymbol{e}_2 = \{0, 1, 0, \cdots, 0\}$, ..., $\boldsymbol{e}_N = \{0, 0, 0, \cdots, 1\}$.

Due to the separation of the Fourier series into its components, it is not hard to verify that the *Parseval theorem*

$$\|\boldsymbol{f}\|_2^2 = \sum_{k=-\infty}^{\infty} \|\hat{\boldsymbol{f}}(k)\|_{\mathcal{H}}^2$$

holds for $L^2(\mathcal{H})$ because Parseval's equality holds for every single component f_n.

With these preparations, we can introduce vector valued Hardy spaces.

Definition 2.30 (Vector valued Hardy spaces). *Let $\mathcal{H}$ be a separable Hilbert space, and let $1 \leq p \leq \infty$. Then $H^p(\mathcal{H})$ denotes the subset of $L^p(\mathcal{H})$ of all $\boldsymbol{f} \in L^p(\mathcal{H})$ whose Fourier coefficients* (2.60) *with negative indices vanish, thus*

$$H^p(\mathcal{H}) := \{\boldsymbol{f} \in L^p(\mathcal{H}) : \hat{\boldsymbol{f}}(k) = 0 \text{ for all } k < 0\} .$$

Let $\{\boldsymbol{e}_n\}_{n=1}^{\infty}$ be an arbitrary orthonormal basis in $\mathcal{H}$, then the above discussions on the Fourier series expansion make it clear that the definition of the Hardy spaces $H^p(\mathcal{H})$ is equivalent to the following statement

Proposition 2.31. *A function $\boldsymbol{f} \in L^p(\mathcal{H})$ belongs to $H^p(\mathcal{H})$ if and only if all of its coordinate functions $f_n = \langle \boldsymbol{f}, \boldsymbol{e}_n \rangle_{\mathcal{H}}$ are elements of H^p.*

Similar to the case of scalar functions, every $\boldsymbol{f} \in H^p(\mathcal{H})$ can be associated with a function

$$\boldsymbol{F}(z) := \sum_{k=0}^{\infty} \hat{\boldsymbol{f}}(k)\, z^k \qquad (z \in \mathbb{D}) \tag{2.61}$$

which is analytic for all $z \in \mathbb{D}$, where $\hat{\boldsymbol{f}}(k)$ are the Fourier coefficients (2.60) of $\boldsymbol{f}$. As in the case of scalar functions, we need to show that the function $\boldsymbol{F}(r\mathrm{e}^{\mathrm{i}\theta})$ converges to $\boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta})$ in $L^p(\mathcal{H})$ as $r \to 1$. However, since the general proof can be simply reduced to the scalar case, we just state the result, which is completely analog to the scalar case, but omit the lengthy and technical proof.

Theorem 2.32. *Let $\mathcal{H}$ be a separable Hilbert space, let $1 \leq p \leq \infty$ and let $\boldsymbol{f} \in H^p(\mathcal{H})$. Define $\boldsymbol{F}_r(\mathrm{e}^{\mathrm{i}\theta}) = \boldsymbol{F}(r\mathrm{e}^{\mathrm{i}\theta})$ with $\boldsymbol{F}$ given by* (2.61)*. Then it holds that*

$$\lim_{r \to 1} \|\boldsymbol{F}_r - \boldsymbol{f}\|_p = 0 .$$

2.4 Operator-valued Analytic Functions

Next we consider analytic functions with values in the space of bounded linear operators. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be separable Hilbert spaces and denote by $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$ the set of all bounded linear operators $\mathbf{H}$ from $\mathcal{H}_1$ to $\mathcal{H}_2$. It is known that $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$ is a Banach space with respect to the usual operator norm

$$\|\mathbf{H}\|_{\mathcal{H}_1 \to \mathcal{H}_2} := \sup_{\boldsymbol{f} \in \mathcal{H}_1, \|\boldsymbol{f}\|_{\mathcal{H}_1} \leq 1} \|\mathbf{H}\boldsymbol{f}\|_{\mathcal{H}_2} . \tag{2.62}$$

Therewith, we define operator valued bounded analytic functions.

Definition 2.33 (Bounded analytic functions). *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two separable Hilbert spaces, and let $\{\hat{\mathbf{H}}(k)\}_{k=0}^{\infty}$ be a sequence of elements in $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$. Therewith, we define the function*

$$\mathbf{H}(z) = \sum_{k=0}^{\infty} \hat{\mathbf{H}}(k)\, z^k\,, \qquad z \in \mathbb{D}\,. \tag{2.63}$$

Now $H^\infty(\mathcal{H}_1, \mathcal{H}_2)$ *denotes the set of all* bounded analytic functions with values in $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$, *that is the set of all functions of the form* (2.63) *which are uniformly bounded in* $\mathbb{D}$, *i.e. for which*

$$\|\mathbf{H}\|_\infty := \sup_{z\in\mathbb{D}} \|\mathbf{H}(z)\|_{\mathcal{H}_1\to\mathcal{H}_2} < \infty\,.$$

Equation (2.63) means that the power series on the right hand side is assumed to converge in $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$ for every $z \in \mathbb{D}$. It shows that $\mathbf{H}$ is holomorphic in $\mathbb{D}$. Since $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$ is a Banach space, the usual differentiation is defined on $H^\infty(\mathcal{H}_1, \mathcal{H}_2)$ (cf. Def. 1.16). Then, as in the scalar case, it follows from the power series representation (2.63) that $\mathbf{H}$ is analytic (complex differentiable) for all $z \in \mathbb{D}$. For the particular case $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{C}$, one obtains the usual Hardy space H^∞.

Given a bounded analytic function $\mathbf{H}$ with values in $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$, we will be particularly interested in multiplication operators $\mathbf{O_H} : L^p(\mathcal{H}_1) \to L^p(\mathcal{H}_2)$, with a certain $p \in (1, \infty)$, defined by

$$(\mathbf{O_H}\boldsymbol{f})(\zeta) := \mathbf{H}(\zeta)\, \boldsymbol{f}(\zeta)\,, \qquad \zeta \in \mathbb{T}$$

and $\mathbf{O}^+_\mathbf{H} : H^p(\mathcal{H}_1) \to H^p(\mathcal{H}_2)$ given by

$$(\mathbf{O}^+_\mathbf{H}\boldsymbol{f})(z) := \mathbf{H}(z)\, \boldsymbol{f}(z)\,, \qquad z \in \mathbb{D}\,.$$

The bounded analytic function $\mathbf{H} \in H^\infty(\mathcal{H}_1, \mathcal{H}_2)$ will be called the *symbol* of $\mathbf{O_H}$ and $\mathbf{O}^+_\mathbf{H}$. The norm of these operators is defined as usual by

$$\|\mathbf{O_H}\| = \sup_{\boldsymbol{f}\in L^p(\mathcal{H}_1)} \frac{\|\mathbf{O_H}\boldsymbol{f}\|_{L^p(\mathcal{H}_2)}}{\|\boldsymbol{f}\|_{L^p(\mathcal{H}_1)}} \quad \text{and} \quad \|\mathbf{O}^+_\mathbf{H}\| = \sup_{\boldsymbol{f}\in H^p(\mathcal{H}_1)} \frac{\|\mathbf{O}^+_\mathbf{H}\boldsymbol{f}\|_{H^p(\mathcal{H}_p)}}{\|\boldsymbol{f}\|_{H^p(\mathcal{H}_1)}}\,.$$

The following proposition will formally prove that the norm of these two operators are given by the norm $\|\mathbf{H}\|_\infty$ of the symbol $\mathbf{H}$.

Proposition 2.34. *Let* $\mathbf{H} \in H^\infty(\mathcal{H}_1, \mathcal{H}_2)$ *be a bounded analytic function. Then for the norms of the multiplication operators* $\mathbf{O_H}$ *and* $\mathbf{O}^+_\mathbf{H}$ *with symbol* $\mathbf{H}$ *it holds that*

$$\|\mathbf{O_H}\| = \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} \|\mathbf{H}(\zeta)\|_{\mathcal{H}_1\to\mathcal{H}_2} = \sup_{z\in\mathbb{D}} \|\mathbf{H}(z)\|_{\mathcal{H}_1\to\mathcal{H}_2} = \|\mathbf{O}^+_\mathbf{H}\|\,.$$

Proof. First consider $\mathbf{O_H}$

$$\begin{aligned}
\|\mathbf{O_H}\boldsymbol{f}\|_{L^p(\mathcal{H}_2)} &= \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \|\mathbf{H}(\mathrm{e}^{\mathrm{i}\theta})\boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta})\|^p_{\mathcal{H}_2}\, \mathrm{d}\theta\right)^{1/p} \\
&\le \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \|\mathbf{H}(\mathrm{e}^{\mathrm{i}\theta})\|^p_{\mathcal{H}_1\to\mathcal{H}_2}\, \|\boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta})\|^p_{\mathcal{H}_1}\, \mathrm{d}\theta\right)^{1/p} \\
&\le \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} \|\mathbf{H}(\zeta)\|_{\mathcal{H}_1\to\mathcal{H}_2}\, \|\boldsymbol{f}\|_{L^p(\mathcal{H}_1)}
\end{aligned}$$

which shows that $\|\mathbf{O_H}\| \leq \|\mathbf{H}\|_\infty$. To prove the reverse inequality, we choose an arbitrary $\epsilon > 0$ and define the set

$$M(\epsilon) := \big\{\theta \in [-\pi,\pi) \ : \ \|\mathbf{H}(e^{i\theta})\|_{\mathcal{H}_1\to\mathcal{H}_2} \geq \|\mathbf{H}\|_\infty - \epsilon/2\big\}$$

and denote by $\chi_{M(\epsilon)}$ the indicator function of $M(\epsilon)$. The Lebesgue measure of $M(\epsilon)$ will be denoted by $\mu_\epsilon := \frac{1}{2\pi}\int_{-\pi}^{\pi}\chi_{M(\epsilon)}(e^{i\theta})\,d\theta$. Moreover, to every $\theta \in M(\epsilon)$ there exists a $\boldsymbol{g}(e^{i\theta}) \in \mathcal{H}_1$ such that

$$\|\mathbf{H}(e^{i\theta})\,\boldsymbol{g}(e^{i\theta})\|_{\mathcal{H}_2} \geq (\|\mathbf{H}\|_\infty - \epsilon)\,\|\boldsymbol{g}(e^{i\theta})\|_{\mathcal{H}_1}\ . \tag{2.64}$$

Now, we define the function

$$\boldsymbol{f}_\epsilon(e^{i\theta}) := \tfrac{1}{\mu_\epsilon}\,\chi_{M(\epsilon)}(e^{i\theta})\,\boldsymbol{g}(e^{i\theta})$$

where every $\boldsymbol{g}(e^{i\theta})$ is chosen such that (2.64) holds for all $\theta \in M(\epsilon)$. For this function, one obtains

$$\begin{aligned}\|\mathbf{O_H}\,\boldsymbol{f}_\epsilon\|_{L^p(\mathcal{H}_2)} &= \left(\frac{1}{2\pi}\int_{M(\epsilon)}\tfrac{1}{\mu_\epsilon^p}\,\|\mathbf{H}(e^{i\theta})\,\boldsymbol{g}(e^{i\theta})\|_{\mathcal{H}_2}^p\,d\theta\right)^{1/p} \\ &\geq (\|\mathbf{H}\|_\infty - \epsilon)\,\frac{1}{\mu_\epsilon}\left(\frac{1}{2\pi}\int_{M(\epsilon)}\|\boldsymbol{g}(e^{i\theta})\|_{\mathcal{H}_1}^p\,d\theta\right)^{1/p} \\ &= (\|\mathbf{H}\|_\infty - \epsilon)\,\|\boldsymbol{f}_\epsilon\|_{L^p(\mathcal{H}_1)}\ .\end{aligned}$$

Since ϵ was chosen arbitrary, this shows that $\|\mathbf{O_H}\| \geq \|\mathbf{H}\|_\infty$ and together with the first part of this proof, one has $\|\mathbf{O_H}\| = \|\mathbf{H}\|_\infty$.

It remains to show that $\|\mathbf{O_H^+}\| = \|\mathbf{O_H}\|$. Since $H^p(\mathcal{H}_1) \subset L^p(\mathcal{H}_1)$, it is clear that $\|\mathbf{O_H^+}\| \leq \|\mathbf{O_H}\|$. To prove the reverse inequality, we consider polynomials in $L^p(\mathcal{H}_1)$ of the form

$$\boldsymbol{p}(e^{i\theta}) = \sum_{k=-N_1}^{N_2}\hat{\boldsymbol{p}}(k)\,e^{ik\theta} = e^{-iN_1\theta}\sum_{k=0}^{N_1+N_2}\hat{\boldsymbol{p}}(k-N_1)\,e^{ik\theta} = e^{-iN_1\theta}\,\boldsymbol{p}_c(e^{i\theta})$$

with $\hat{\boldsymbol{p}}(k) \in \mathcal{H}_1$ and $N_1, N_2 \geq 0$. However, the polynomial $\boldsymbol{p}_c(e^{i\theta})$, obtained from $\boldsymbol{p}(e^{i\theta})$ by factoring out $e^{-iN_1\theta}$ belongs to $H^p(\mathcal{H}_1)$, and it is easily verified that $\|\mathbf{O_H^+}\,\boldsymbol{p}_c\|_{H^p(\mathcal{H}_1)} = \|\mathbf{O_H}\,\boldsymbol{p}\|_{L^p(\mathcal{H}_1)}$. Moreover, the polynomials $\mathcal{P}(\mathcal{H}_1)$ of the above form are dense in $L^2(\mathcal{H}_1)$. Therefore, to every $\mathbf{H} \in H^\infty(\mathcal{H}_1,\mathcal{H}_2)$ there exist a polynomials $\boldsymbol{p} \in L^p(\mathcal{H}_1)$ with $\|\boldsymbol{p}\|_{L^p(\mathcal{H}_1)} = 1$ so that

$$\|\mathbf{O_H^+}\,\boldsymbol{p}_c\|_{H^p(\mathcal{H}_1)} = \|\mathbf{O_H}\,\boldsymbol{p}\|_{L^p(\mathcal{H}_2)} \geq \|\mathbf{O_H}\| - \epsilon\ . \tag{2.65}$$

In turn this implies that

$$\|\mathbf{O_H^+}\| = \sup_{\substack{\boldsymbol{f}\in H^p(\mathcal{H}_1)\\ \|\boldsymbol{f}\|_{H^p(\mathcal{H}_1)}\leq 1}}\|\mathbf{O_H^+}\boldsymbol{f}\|_{H^p(\mathcal{H}_2)} \geq \|\mathbf{O_H^+}\,\boldsymbol{p}_c\|_{H^p(\mathcal{H}_1)} \geq \|\mathbf{O_H}\| - \epsilon$$

which shows that $\|\mathbf{O_H^+}\| \geq \|\mathbf{H}\|_\infty$ and altogether that $\|\mathbf{O_H^+}\| = \|\mathbf{H}\|_\infty$. $\square$

Since every symbol $\mathbf{H} \in H^\infty(\mathcal{H}_1, \mathcal{H}_2)$ is analytic in $\mathbb{D}$ it is clear that $(\mathbf{O_H}\boldsymbol{f})(z) = \mathbf{H}(z)\,\boldsymbol{f}(z)$ belongs to $H^p(\mathcal{H}_2)$ provided that $\boldsymbol{f} \in H^p(\mathcal{H}_1)$. For this reason and because $H^p(\mathcal{H}_1)$ and $H^p(\mathcal{H}_2)$ are subspaces of $L^p(\mathcal{H}_1)$ and $L^p(\mathcal{H}_2)$, respectively, the operator $\mathbf{O}_\mathbf{H}^+$ can be considered as the restriction of $\mathbf{O_H}$ to the subspace $H^p(\mathcal{H}_1)$ of $L^p(\mathcal{H}_1)$.

The most important case is $p = 2$. Then the operators $\mathbf{O_H}$ and $\mathbf{O}_\mathbf{H}^+$ are mappings from the Hilbert space $L^2(\mathcal{H}_1)$ into the Hilbert space $L^2(\mathcal{H}_2)$. In this case, it is easily verified that the adjoint of the operator $\mathbf{O_H}$ is given by $\mathbf{O}_\mathbf{H}^* = \mathbf{O}_{\mathbf{H}^*}$ where $\mathbf{H}^*(z) = [\mathbf{H}(z)]^*$ is the adjoint of $\mathbf{H}(z) : \mathcal{H}_1 \to \mathcal{H}_2$ for every $z \in \mathbb{D}$.

Example 2.35. Assume that the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ both have finite dimension. Then, without any loss of generality, we may assume that $\mathcal{H}_1 = \mathbb{C}^N$ and $\mathcal{H}_2 = \mathbb{C}^M$ with the dimensions $N, M \geq 1$ and with the usual Euclidean norm in $\mathbb{C}^N$ and $\mathbb{C}^M$. It is well known that every bounded linear operator $\mathbf{H} \in \mathscr{B}(\mathbb{C}^N, \mathbb{C}^M)$ can be identified with a complex $M \times N$ matrix $\mathbf{H}$ with M rows and N columns. Therefore, $\mathscr{B}(\mathbb{C}^N, \mathbb{C}^M)$ can be identified with the set $\mathbb{C}^{M\times N}$ of all complex $M \times N$ matrices, and the norm of any matrix $\mathbf{H} \in \mathbb{C}^{M\times N}$, induced by the Euclidean vector norm in $\mathbb{C}^N$ and $\mathbb{C}^M$, is known to be

$$\|\mathbf{H}\|_{\mathbb{C}^{M\times N}} = \sup_{\boldsymbol{f}\in\mathbb{C}^N} \frac{\|\mathbf{H}\boldsymbol{f}\|_{\mathbb{C}^M}}{\|\boldsymbol{f}\|_{\mathbb{C}^N}} = \sqrt{\lambda_{\max}\{\mathbf{H}^*\mathbf{H}\}}$$

wherein $\lambda_{\max}(\mathbf{H}^*\mathbf{H})$ is the largest singular value of the matrix $\mathbf{H}$. This norm is also known as the *spectral norm* of $\mathbf{H}$.

Moreover, every $\mathbf{H} \in H^\infty(\mathbb{C}^N, \mathbb{C}^M)$ has the general form (2.63) in which all $\hat{\mathbf{H}}(k) \in \mathbb{C}^{M\times N}$ are complex $M \times N$ matrices. To shorten the notation, the space $H^\infty(\mathbb{C}^N, \mathbb{C}^M)$ of all matrix valued bounded analytic functions will be denoted by $H^\infty(\mathbb{C}^{M\times N})$, and the norm in this space is given by

$$\|\mathbf{H}\|_\infty = \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} \lambda_{\max}\{\mathbf{H}^*(\zeta)\mathbf{H}(\zeta)\} = \sup_{z\in\mathbb{D}} \lambda_{\max}\{\mathbf{H}^*(z)\mathbf{H}(z)\}\ .$$

Notes

Still the classical reference for trigonometric series is the volume of Zygmund [92]. Theorem 2.8 is due to Jackson [51]. Detailed proofs can also be found in [92, Chap. III] or [61, vol. 1, Chap. IV]. There are numerous text books containing the basic theory of Hardy spaces in different detail and various forms. The scalar case can be found for example in [30, 41, 45, 48, 70]. The exposition here and the given proofs are primarily taken from [70] where also most of the omitted proofs and auxiliary results can be found. The vector valued case is considered in detail in [44, 62, 83]. The notion of inner and outer functions was introduced by Beurling in his seminal paper on shift-invariant subspaces [7]. It seems be worthwhile to consider this original approach to the inner-outer factorization since it gives somewhat more descriptive i.e geometrical proofs of the theorems in Section 2.2.3. We also refer the reader to [43, 44, 45, 57, 62, 83].

http://www.springer.com/978-3-642-03638-5

Advanced Topics in System and Signal Theory
A Mathematical Approach
Pohl, V.; Boche, H.
2010, VIII, 241 p. 5 illus., Hardcover
ISBN: 978-3-642-03638-5