Ruprecht-Karls-Universität Heidelberg
Institut für Angewandte Mathematik
Prof. Dr. Jan JOHANNES

Outline of the lecture course

STATISTICS OF INVERSE PROBLEMS

Summer semester 2017

Preliminary version: April 21, 2017

If you find **errors in the outline**, please send a short note
by email to johannes@math.uni-heidelberg.de

MΛTHEMΛTIKON, Im Neuenheimer Feld 205, 69120 Heidelberg
phone: +49 6221 54.14.190 – fax: +49 6221 54.53.31
email: johannes@math.uni-heidelberg.de
webpage: www.razbaer.eu/ag-johannes/vl/SIP-SS17/

Table of contents

1 Introduction **1**

2 Theoretical basics and terminologies **5**

2.1 Hilbert space 5

2.1.1 Abstract smoothness condition 7

2.2 Linear operator between Hilbert spaces 8

2.2.1 Compact, nuclear and Hilbert-Schmidt operator 12

2.2.2 Spectral theory and functional calculus 14

2.2.3 Abstract smoothing condition 18

Chapter 1

Introduction

SHORT SUMMARY

Statistical ill-posed inverse problems are becoming increasingly important in a diverse range of disciplines, including geophysics, astronomy, medicine and economics. Roughly speaking, in all of these applications the observable signal $g = Tf$ is a transformation of the functional parameter of interest f under a linear operator T. Statistical inference on f based on an estimation of g which usually requires an inversion of T is thus called an *inverse problem*. The lecture course focuses on statistical ill-posed inverse problems with noise in the operator where neither the signal g nor the linear operator T are known in advance, although they can be estimated from the data. Our objective in this context is the construction of minimax-optimal fully data-driven estimation procedures of the unknown function f. Special attention is given to four models and their extensions, namely Gaussian inverse regression, density deconvolution, functional linear regression and non-parametric instrumental regression, which lead naturally to statistical ill-posed inverse problems with noise in the operator.

APPLICATIONS

Density deconvolution with unknown error distribution. The biologist who is interested in the density f_X of a gene-expression intensity X, can record in a cDNA microarray the expressed gene intensity X only corrupted by the intensity of a background noise U, that is $Y = X + U$. If the additive measurement error U is independent of X then the density $f_Y = f_X \star f_U$ of Y equals the convolution of f_X and the error density f_U. Consequently, recovering f_X from the estimated density $f_Y = C_{f_U} f_X$ of Y is an inverse problem where C_{f_U} is the convolution operator defined by the error density f_U. In this situation, the density f_X of the random variable X has to be estimated non-parametrically based on an iid. sample from a noisy observation Y of X which is called a density deconvolution problem. There is a vast literature on deconvolution with known error density which leads to a statistical ill-posed inverse problem with known operator. On the other hand, if the error density f_U is estimated from an additional calibration sample of the error U then the deconvolution problem corresponds to a statistical ill-posed inverse problem with noise in the operator.

Functional linear regression. In climatology, prediction of level of ozone pollution based on continuous measurements of pollutant indicators is often modelled by a functional linear model. In this context a scalar response Y (i.e. the ozone concentration) is modelled in dependence of a random function X (i.e. the daily concentration curve of a pollutant indicator). Typically the dependence is assumed to be linear which finds its expression in a linear normal equation $c_{YX} = \Gamma_{XX}\beta$ where c_{YX} is the cross-correlation between Y and X, and Γ_{XX} is the covariance operator associated to the indicator X. Note that both the cross-correlation function c_{YX} and the covariance operator Γ_{XX} need to be estimated in practice. Consequently, the non-parametric estimation of the functional slope parameter β based on an iid. sample from (Y, X) leads to a statistical ill-posed inverse problem with noise in the operator.

Non-parametric instrumental regression. An econometrician who wants to analyse an economic relation between a response Y and an endogenous vector X of explanatory variables, might incorporate a vector of exogenous instruments Z. This situation is usually treated by considering a conditional moment equation $r_{Y|Z} = K_{X|Z}\,\varphi$ where $r_{Y|Z} = \mathbb{E}_{Y|Z}$ is the conditional expectation function of Y given Z and $K_{X|Z}$ is the conditional expectation operator of X given Z. As these are unknown in practice, inference on φ based on an iid. sample from (Y, X, Z) is a statistical ill-posed inverse problem with noise in the operator.

STATISTICAL ILL-POSED INVERSE PROBLEMS

We study non-parametric estimation of the functional parameter of interest f in an inverse problem, that is, its reconstruction based on an estimation of a linear transformation $g = Tf$. It is important to note that in all the applications discussed above both the signal g and the inherent transformation T are unknown in practice, although they can be estimated from the data. The estimated signal $\widehat{g}_\varepsilon$ and operator $\widehat{T}_\sigma$ respectively given by

$$\widehat{g}_\varepsilon = Tf + \sqrt{\varepsilon}\,\dot{W} \quad \text{and} \quad \widehat{T}_\sigma = T + \sqrt{\sigma}\,\dot{B}. \tag{1.1}$$

are noisy versions of g and T contaminated by additive random errors $\dot{W}$ and $\dot{B}$ with respective noise levels ε and σ. Consequently, a statistical inference on the functional parameter of interest f has to take into account that a random noise is present in both the estimated signal $\dot{W}$ and the estimated operator $\dot{B}$.

Gaussian inverse regression with noise in the operator. A particularly interesting situation is given by model (1.1) where the random error $\dot{W}$ and $\dot{B}$ are independent Gaussian white noises. This model is particularly useful to characterise the influence of an *a priori* knowledge of the operator T. To this end we will compare three cases: First, the operator T is *fully known* in advance, i.e., the noise level σ is equal to zero. Second, it is *partially known*, that is, the eigenfunctions of T are known in advance but the "observed" eigenvalues of T are contaminated with an additive Gaussian error. Third, the operator T is *unknown*.

MINIMAX-OPTIMAL ESTIMATION

Typical questions in this context are the non-parametric estimation of the functional parameter f on an interval or in a given point, referred to as global or local estimation, respectively. However, these are special cases in a general framework where the accuracy of an estimator $\widehat{f}$ of f given the estimations (1.1) is measured by a distance $\mathfrak{d}_{\text{ist}}(\widehat{f}, f)$. A suitable choice of the distance covers than the global as well as the local estimation problem. Moreover, denoting by $\mathbb{P}^{\varepsilon;\sigma}_{f,T}|\mathfrak{d}_{\text{ist}}(\widehat{f}, f)|^2$ (or $\mathbb{E}^{\varepsilon;\sigma}_{f,T}|\mathfrak{d}_{\text{ist}}(\widehat{f}, f)|^2$) its expectation w.r.t. the probability measure $\mathbb{P}^{\varepsilon;\sigma}_{f,T}$ associated with the observable quantities (1.1) we call the quantity $\mathbb{P}^{\varepsilon;\sigma}_{f,T}|\mathfrak{d}_{\text{ist}}(\widehat{f}, f)|^2$ risk of the estimator $\widehat{f}$ of f. It is well-known that in terms of its risk the attainable accuracy of an estimation procedure is essentially determined by the conditions imposed on f and the operator T. Typically, these conditions are expressed in the form $f \in \mathcal{F}$ and $T \in \mathcal{T}$ for suitable chosen classes $\mathcal{F}$ and $\mathcal{T}$. The class $\mathcal{F}$ reflects prior information on the solution f, e.g., its level of smoothness, and the class $\mathcal{T}$ imposes among others conditions on the decay of the eigenvalues of the operator T.

Consequently, let us introduce the associated family of probability measures $\mathbb{P}^{\varepsilon,\sigma}_{\mathcal{F},\mathcal{T}}$. The accuracy of $\widehat{f}$ is hence measured by its maximal risk over the classes $\mathcal{F}$ and $\mathcal{T}$, that is,

$$\mathfrak{R}_{\mathfrak{d}}\big[\widehat{f}\,|\,\mathbb{P}^{\varepsilon,\sigma}_{\mathcal{F},\mathcal{T}}\big] := \sup\big\{\mathbb{P}^{\varepsilon,\sigma}_{f,T}|\mathfrak{d}_{\text{ist}}(\widehat{f},f)|^2, \mathbb{P}^{\varepsilon,\sigma}_{f,T} \in \mathbb{P}^{\varepsilon,\sigma}_{\mathcal{F},\mathcal{T}}\big\}.$$

Moreover, $\widehat{f}$ is called minimax-optimal up to a finite positive constant C if $\mathfrak{R}_{\mathfrak{d}}\big[\widehat{f}\,|\,\mathbb{P}^{\varepsilon,\sigma}_{\mathcal{F},\mathcal{T}}\big] \leqslant C\inf_{\widetilde{f}}\mathfrak{R}_{\mathfrak{d}}\big[\widetilde{f}\,|\,\mathbb{P}^{\varepsilon,\sigma}_{\mathcal{F},\mathcal{T}}\big]$ where the infimum is taken over all possible estimators of f. Consequently, minimax-optimality of an estimator $\widehat{f}$ based on observations (1.1) is usually shown by establishing both an upper and a lower bound. More precisely, we search a finite positive quantity $\mathcal{R}^{\varepsilon,n}_{\mathfrak{d}}$ depending only on the noise levels and the classes such that

$$\mathfrak{R}_{\mathfrak{d}}\big[\widehat{f}\,|\,\mathbb{P}^{\varepsilon,\sigma}_{\mathcal{F},\mathcal{T}}\big] \leqslant C_1\mathcal{R}^{\varepsilon,\sigma}_{\mathfrak{d}} \quad \text{and} \quad \mathcal{R}^{\varepsilon,\sigma}_{\mathfrak{d}} \leqslant C_2\inf_{\widetilde{f}}\mathfrak{R}_{\mathfrak{d}}\big[\widetilde{f}\,|\,\mathbb{P}^{\varepsilon,\sigma}_{\mathcal{F},\mathcal{T}}\big]$$

where C_1, C_2 are finite positive constants independent of the noise levels. Moreover, the quantity $\mathcal{R}^{\varepsilon,\sigma}_{\mathfrak{d}}$ is called the minimax-optimal rate of convergence over the family $\mathbb{P}_{\mathcal{F},\mathcal{T}} := \{\mathbb{P}^{\varepsilon,\sigma}_{\mathcal{F},\mathcal{T}}, \varepsilon, \sigma \in (0,1)\}$ if it tends to zero as ε and σ tend to zero.

ADAPTIVE ESTIMATION

In many cases the proposed estimation procedures rely on the choice of at least one tuning parameter, which in turn, crucially influences the attainable accuracy of the constructed estimator. In other words, these estimation procedures can attain the minimax rate $\mathcal{R}^{\varepsilon,n}_{\mathfrak{d}}$ over the family $\mathbb{P}_{\mathcal{F},\mathcal{T}}$ only if the inherent tuning parameters are chosen optimally. This optimal choice, however, follows often from a classical squared-bias-variance compromise and requires a *a priori* knowledge about the classes $\mathcal{F}$ and $\mathcal{T}$, which is usually inaccessible in practice. This motivates its data-driven choice in the context of non-parametric statistics since its very beginning in the fifties of the last century. A demanding challenge is then a fully data driven method to select the tuning parameters in such a way that the resulting data-driven estimator of f still attains the minimax-rate up to a constant over a variety of classes $\mathcal{F}$ and $\mathcal{T}$. The fully data driven estimation procedure is then called *adaptive*.

Chapter 2

Theoretical basics and terminologies

2.1 Hilbert space

For a detailed and extensive survey on functional analysis we refer the reader, for example, to Werner [2011] or the series of textbooks by Dunford and Schwartz [1988a,b,c].

§2.1.1 **Definition**. A normed vector space $(\mathbb{H}, \|\cdot\|_{\mathbb{H}})$ over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ that is complete (in a Cauchy-sense) is called a (real or complex) *Hilbert space* if there exists an inner product $\langle \cdot, \cdot \rangle_{\mathbb{H}}$ on $\mathbb{H} \times \mathbb{H}$ with $|\langle h, h \rangle_{\mathbb{H}}|^{1/2} = \|h\|_{\mathbb{H}}$ for all $h \in \mathbb{H}$. □

§2.1.2 **Property**. *Let* $(\mathbb{H}, \|\cdot\|_1)$ *and* $(\mathbb{H}, \|\cdot\|_2)$ *be complete normed vector spaces. If there exists a constant* $K > 0$ *such that* $\|h\|_1 \leqslant K \|h\|_2$ *for any* $h \in \mathbb{H}$ *then,* $\|\cdot\|_1$ *and* $\|\cdot\|_2$ *are equivalent.* □

§2.1.3 **Property**.
(Cauchy-Schwarz inequality) $|\langle h_1, h_2 \rangle_{\mathbb{H}}| \leqslant \|h_1\|_{\mathbb{H}} \cdot \|h_2\|_{\mathbb{H}}$ *for all* $h_1, h_2 \in \mathbb{H}$. □

§2.1.4 **Examples**.

(i) For $k \in \mathbb{N}$ the *Euclidean space* $\mathbb{K}^k$ endowed with the Euclidean inner product $\langle x, y \rangle := \overline{y}^t x$ and the induced Euclidean norm $\|x\| = (\overline{x}^t x)^{1/2}$ for all $x, y \in \mathbb{K}^k$ is a Hilbert space. More generally, given a strictly positive definite $(k \times k)$-matrix W, $\mathbb{K}^k$ endowed with the weighted inner product $\langle x, y \rangle_W := \overline{y}^t W x$ for all $x, y \in \mathbb{K}^k$ is also a Hilbert space.

(ii) Given $\mathcal{J} \subseteq \mathbb{Z}$, denote by $\mathbb{K}^{\mathcal{J}}$ the vector space of all $\mathbb{K}$-valued sequences over $\mathcal{J}$ where we refer to any sequence $(x_j)_{j \in \mathcal{J}} \in \mathbb{K}^{\mathcal{J}}$ as a whole by omitting its index as for example in «the sequence x» and arithmetic operations on sequences are defined element-wise, i.e., $xy := (x_j y_j)_{j \in \mathcal{J}}$. In the sequel, let $\|x\|_{\ell^p} := (\sum_{j \in \mathcal{J}} |x_j|^p)^{1/p}$, for $p \in [1, \infty)$, and $\|x\|_{\ell^\infty} := \sup_{j \in \mathcal{J}} |x_j|$. Thereby, for $p \in [1, \infty]$, consider $\ell^p(\mathcal{J}) := \left\{(x_j)_{j \in \mathcal{J}} \in \mathbb{K}^{\mathcal{J}}, \|x\|_{\ell^p} < \infty\right\}$, or ℓ^p for short, endowed with the norm $\|\cdot\|_{\ell^p}$. In particular, $\ell^2(\mathcal{J})$ is the usual *Hilbert space of square summable sequences over* $\mathcal{J}$ endowed with the inner product $\langle x, y \rangle_{\ell^2} := \sum_{j \in \mathcal{J}} x_j \overline{y}_j$ for all $x, y \in \ell^2(\mathcal{J})$.

(iii) For a strictly positive sequence $\mathfrak{v}$ consider the *weighted norm* $\|x\|_{\mathfrak{v}}^2 := \sum_{j \in \mathcal{J}} \mathfrak{v}_j^2 |x_j|^2$. We define $\ell_{\mathfrak{v}}^2(\mathcal{J})$, or $\ell_{\mathfrak{v}}^2$ for short, as the completion of $\ell^2(\mathcal{J})$ w.r.t. $\|\cdot\|_{\mathfrak{v}}$ which is a Hilbert space endowed with the inner product $\langle x, y \rangle_{\mathfrak{v}} := \langle \mathfrak{v}x, \mathfrak{v}y \rangle_{\ell^2} = \sum_{j \in \mathcal{J}} \mathfrak{v}_j^2 x_j \overline{y}_j$ for all $x, y \in \ell_{\mathfrak{v}}$.

(iv) Let $\mathscr{B}$ be the Borel-σ-algebra on $\mathbb{K}$. Given a measure space $(\Omega, \mathscr{A}, \mu)$ denote by $\mathbb{K}^\Omega$ the vector space of all $\mathbb{K}$-valued functions $f : \Omega \to \mathbb{K}$. Recall that $\|f\|_{L_\mu^p} = (\mu |f|^p)^{1/p} = \left(\int_\Omega |f(\omega)|^p \mu(d\omega)\right)^{1/p}$, for $p \in [1, \infty)$, and $\|f\|_{L_\mu^\infty} := \inf\{c : \mu(|f| > c) = 0\}$, where for $p \in [1, \infty]$, we write $L^p(\Omega, \mathscr{A}, \mu) := \{f \in \mathbb{K}^\Omega, \mathscr{A}\text{-}\mathscr{B}\text{-measurable}, \|f\|_{L^p} < \infty\}$, $L_\mu^p(\Omega)$ or L_μ^p for short, which is endowed with the norm $\|\cdot\|_{L_\mu^p}$ for short. In case μ is the Lebesgue measure, then we may write $L^p(\Omega, \mathscr{A})$, $L^p(\Omega)$, L^p and $\|\cdot\|_{L^p}$ for short. Moreover,

$L^2(\Omega, \mathscr{A}, \mu)$, $L^2_\mu(\Omega)$ or L^2_μ for short, is the usual *Hilbert space of square μ-integrable, $\mathscr{A}$-$\mathscr{B}$-measurable functions on Ω* endowed with the inner product $\langle f, g\rangle_{L^2_\mu} := \mu(f\overline{g})$ for all $f, g \in L^2_\mu$.

(v) Let X be a random variable (r.v.) on a probability space $(\Omega, \mathscr{A}, \mathbb{P})$ taking its values in a measurable space $(\mathcal{X}, \mathscr{B})$. We denote by $\mathbb{P}^X := \mathbb{P} \circ X^{-1}$ the image probability measure of $\mathbb{P}$ under X on $(\mathcal{X}, \mathscr{B})$. For $p \in [1, \infty]$ we set $L^p_X := L^p(\mathcal{X}, \mathscr{B}, \mathbb{P}^X)$ where L^2_X is a Hilbert space endowed with $\langle f, g\rangle_{L^2_X} = \mathbb{P}^X(f\overline{g})$ for all $f, g \in L^2_X$. □

§2.1.5 **Definition.** A subset $\mathcal{U}$ of a Hilbert space $(\mathbb{H}, \langle\cdot,\cdot\rangle_{\mathbb{H}})$ is called *orthogonal* if

$$\forall u_1, u_2 \in \mathcal{U},\ u_1 \neq u_2 :\ \langle u_1, u_2\rangle_{\mathbb{H}} = 0$$

and *orthonormal system (ONS)* if in addition $\|u\|_{\mathbb{H}} = 1, \forall u \in \mathcal{U}$. We say $\mathcal{U}$ is an *orthonormal basis (ONB)* if $\mathcal{U} \subset \mathcal{U}'$ and $\mathcal{U}'$ is ONS, then $\mathcal{U} = \mathcal{U}'$, i.e., if it is a *complete* ONS.

§2.1.6 **Examples.**

(i) Consider the real Hilbert space $L^2([0,1])$ w.r.t. the Lebesgue measure. The *trigonometric basis* $\{\psi_j, j \in \mathbb{N}\}$ given for $t \in [0,1]$ by

$$\psi_1(t) := 1,\ \psi_{2k}(t) := \sqrt{2}\cos(2\pi k t),\ \psi_{2k+1}(t) := \sqrt{2}\sin(2\pi k t),\ k = 1, 2, \dots,$$

is orthonormal and complete, i.e. an ONB.

(ii) Consider the complex Hilbert space $L^2([0,1))$, then the *exponential basis* $\{e_j, j \in \mathbb{Z}\}$ with

$$e_j(t) := \exp(-\iota 2\pi j t) \text{ for } t \in [0,1) \text{ and } j \in \mathbb{Z},$$

is orthonormal and complete, i.e. an ONB. □

§2.1.7 **Properties.**

(Pythagorean formula) *If $h_1, \dots, h_n \in \mathbb{H}$ are orthogonal, then $\|\sum_{j=1}^n h_j\|^2_{\mathbb{H}} = \sum_{j=1}^n \|h_j\|^2_{\mathbb{H}}$.*

(Bessel's inequality) *If $\mathcal{U} \subset \mathbb{H}$ is an ONS, then $\|h\|^2_{\mathbb{H}} \geqslant \sum_{u \in \mathcal{U}} |\langle h, u\rangle_{\mathbb{H}}|^2$ for all $h \in \mathbb{H}$.*

(Parseval's formula) *An ONS $\mathcal{U} \subset \mathbb{H}$ is complete if and only if $\|h\|^2_{\mathbb{H}} = \sum_{u \in \mathcal{U}} |\langle h, u\rangle_{\mathbb{H}}|^2$ for all $h \in \mathbb{H}$.* □

§2.1.8 **Definition.** Let $\mathcal{U}$ be a subset of a Hilbert space $(\mathbb{H}, \langle\cdot,\cdot\rangle_{\mathbb{H}})$. Denote by $\mathbb{U} := \overline{lin}(\mathcal{U})$ the closure of the linear subspace spanned by the elements of $\mathcal{U}$ and its orthogonal complement in $(\mathbb{H}, \langle\cdot,\cdot\rangle_{\mathbb{H}})$ by $\mathbb{U}^\perp := \big\{h \in \mathbb{H} : \langle h, u\rangle_{\mathbb{H}} = 0, \forall u \in \overline{lin}(\mathcal{U})\big\}$ where $\mathbb{H} = \mathbb{U} \oplus \mathbb{U}^\perp$. □

§2.1.9 **Remark.** If $\mathcal{U} \subset \mathbb{H}$ is an ONS, then there exists an ONS $\mathcal{V} \subset \mathbb{H}$ such that $\mathbb{H} = \overline{lin}(\mathcal{U}) \oplus \overline{lin}(\mathcal{V})$ and for all $h \in \mathbb{H}$ it holds $h = \sum_{u \in \mathcal{U}} \langle h, u\rangle_{\mathbb{H}} u + \sum_{v \in \mathcal{V}} \langle h, v\rangle_{\mathbb{H}} v$ (in a $\mathbb{H}$-sense). In particular, if $\mathcal{U}$ is an ONB then $h = \sum_{u \in \mathcal{U}} \langle h, u\rangle_{\mathbb{H}} u$ for all $h \in \mathbb{H}$. □

§2.1.10 **Definition.** Given $\mathcal{J} \subset \mathbb{Z}$, a sequence $(u_j)_{j \in \mathcal{J}}$ in $\mathbb{H}$ is said to be *orthonormal and complete* (i.e. orthonormal basis) if the subset $\mathcal{U} = \{u_j, j \in \mathcal{J}\}$ is a complete ONS (i.e. ONB). The Hilbert space $\mathbb{H}$ is called *separable*, if there exists a complete orthonormal sequence. □

§2.1.11 **Examples.** The Hilbert space $(\mathbb{R}^k, \langle\cdot,\cdot\rangle_M)$, $(\ell^2_{\mathfrak{v}}, \langle\cdot,\cdot\rangle_{\ell^2_{\mathfrak{v}}})$ and $(L^2_\mu(\Omega), \langle\cdot,\cdot\rangle_{L^2_\mu})$ with σ-finite measure μ are separable. On the contrary, given $\lambda \in \mathbb{R}$ define the function $f_\lambda : \mathbb{R} \to \mathbb{C}$

with $f_\lambda(x) := e^{\iota\lambda x}$ and set $\mathcal{H} = \overline{lin}\{f_\lambda, \lambda \in \mathbb{R}\}$. Observe that $\langle f, g\rangle = \lim_{t\to\infty} \frac{1}{2t}\int_{-t}^{t} f(s)\overline{g(s)}ds$ defines an inner product on $\mathcal{H}$. The completion of $\mathcal{H}$ w.r.t. the induced norm $\|f\| = |\langle f, f\rangle|^{1/2}$ is a Hilbert space which is not separable, since $\|f_\lambda - f_{\lambda'}\| = \sqrt{2}$ for all $\lambda \neq \lambda'$. □

§2.1.12 **Definition**. Given $\mathcal{J} \subseteq \mathbb{Z}$ we call a (possibly finite) sequence $(\mathcal{J}_m)_{m\in\mathcal{M}}$, $\mathcal{M} \subseteq \mathbb{N}$, a *nested sieve in $\mathcal{J}$*, if (i) $\mathcal{J}_k \subset \mathcal{J}_m$, for any $k \leqslant m$, $k, m \in \mathcal{M}$, (ii) $|\mathcal{J}_m| < \infty$, $m \in \mathcal{M}$, and (iii) $\cup_{m\in\mathcal{M}}\mathcal{J}_m = \mathcal{J}$. We write $\mathcal{J}_m^c := \mathcal{J}\backslash\mathcal{J}_m$, $m \in \mathcal{M}$. Denoting $[\![a, b]\!] := [a, b] \cap \mathbb{Z}$ we use typically the nested sieve $([\![1, m]\!])_{m\in\mathbb{N}}$ and $([\![-m, m]\!])_{m\in\mathbb{N}}$ in $\mathcal{J} = \mathbb{N}$ and $\mathcal{J} = \mathbb{Z}$, respectively. Analogously, given an ONS $\mathcal{U} = \{u_j, j \in \mathcal{J}\}$ and setting $\mathbb{U}_m := \overline{lin}\{u_j, j \in \mathcal{J}_m\}$, $m \in \mathcal{M}$, for a nested sieve $(\mathcal{J}_m)_{m\in\mathcal{M}}$ in $\mathcal{J}$ we call the (possibly finite) sequence $(\mathbb{U}_m)_{m\in\mathcal{M}}$ *a nested sieve in* $\mathbb{U} := \overline{lin}\{u_j, j \in \mathcal{J}\}$. We write $\mathbb{U}_m^\perp := \overline{lin}\{u_j, j \in \mathcal{J}_m^c\}$ where $\mathbb{U} = \mathbb{U}_m \oplus \mathbb{U}_m^\perp$. For convenient notations we set further $\mathbb{1}_{\mathcal{J}_m} := (\mathbb{1}_{\mathcal{J}_m}(j))_{j\in\mathcal{J}}$ with $\mathbb{1}_{\mathcal{J}_m}(j) = 1$ if $j \in \mathcal{J}_m$ and $\mathbb{1}_{\mathcal{J}_m}(j) = 0$ otherwise, and analogously $\mathbb{1}_{\mathcal{J}_m^c} := (\mathbb{1}_{\mathcal{J}_m^c}(j))_{j\in\mathcal{J}}$. □

§2.1.13 **Definition**. We call an ONS $\mathcal{U} = \{u_j, j \in \mathcal{J}\}$ in L^2_μ (respectively, in ℓ^2)

(i) *regular w.r.t. a nested sieve $(\mathcal{J}_m)_{m\in\mathcal{M}}$ in $\mathcal{J}$ and a weight sequence* $\mathfrak{v}$ if there is a finite constant $\tau_{u\mathfrak{v}} \geqslant 1$ satisfying $\|\sum_{j\in\mathcal{J}_m} \mathfrak{v}_j^2|u_j|^2\|_{L^\infty_\mu} \leqslant \tau_{u\mathfrak{v}}^2 \sum_{j\in\mathcal{J}_m} \mathfrak{v}_j^2$ for all $m \in \mathcal{M}$;

(ii) *regular w.r.t. a weight sequence* $\mathfrak{a}$ if there exists a finite constant $\tau_{u\mathfrak{a}} \geqslant 1$ such that $\|\sum_{j\in\mathcal{J}} \mathfrak{a}_j^2|u_j|^2\|_{L^\infty_\mu} \leqslant \tau_{u\mathfrak{a}}^2$. □

§2.1.14 **Remark**. According to Lemma 6 of Birgé and Massart [1997] assuming in L^2 a regular ONS $\{u_j, j \in \mathbb{N}\}$ w.r.t. the nested sieve $([\![1, m]\!])_{m\in\mathbb{N}}$ and $\mathfrak{v} \equiv 1$ is exactly equivalent to following property: there exists a finite constant $\tau_u \geqslant 1$ such that for any h belonging to the subspace $\mathbb{U}_m$, spanned by the first m functions $\{u_j\}_{j=1}^m$, holds $\|h\|_{L^\infty} \leqslant \tau_u\sqrt{m}\,\|h\|_{L^2}$. Typical example are bounded basis, such as the trigonometric basis, or basis satisfying the assertion, that there exists a positive constant C_∞ such that for any $(c_1, \dots, c_m) \in \mathbb{R}^m$, $\|\sum_{j=1}^m c_j u_j\|_{L^\infty} \leqslant C_\infty\sqrt{m}|c|_\infty$ where $|c|_\infty = \max_{1\leqslant j\leqslant m} c_j$. Birgé and Massart [1997] have shown that the last property is satisfied for piece-wise polynomials, splines and wavelets. □

§2.1.15 **Example** *(§2.1.6 (i) continued)*. Consider the *trigonometric basis* $\{\psi_j, j \in \mathbb{N}\}$ in the real Hilbert space $L^2([0, 1])$. Since $\sup_{j\in\mathbb{N}} \|\psi_j\|_{L^\infty} \leqslant \sqrt{2}$ setting $\tau_{\psi\mathfrak{v}}^2 := 2$ the trigonometric basis is regular w.r.t. any nested Sieve $(\mathcal{J}_m)_{m\in\mathcal{M}}$ and sequence $\mathfrak{v}$, i.e., §2.1.13 (i) holds with $\|\sum_{j\in\mathcal{J}_m} \mathfrak{v}_j^2|\psi_j|^2\|_{L^\infty} \leqslant \tau_{\psi\mathfrak{v}}^2 \sum_{j\in\mathcal{J}_m} \mathfrak{v}_j^2$. In the particular case of the nested sieve $([\![1, 1+2m]\!])_{m\in\mathbb{N}}$ and $\mathfrak{v} \equiv 1$, we have $\sum_{j=1}^{1+2m} |\psi_j|^2 = \mathbb{1}_{[0,1]} + \sum_{j=1}^m\{2\sin^2(2\pi j\bullet) + 2\cos^2(2\pi j\bullet)\} = 1 + 2m$ and thus, the trigonometric basis is regular with $\tau_\psi^2 := 1$. Moreover, the trigonometric basis is regular w.r.t. any square-summable weight sequence $\mathfrak{a}$, i.e., $\|\mathfrak{a}\|_{\ell^2} < \infty$. Indeed, in this situation we have $\|\sum_{j\in\mathbb{N}} \mathfrak{a}_j^2|\psi_j|^2\|_{\ell_\infty} \leqslant 2\,\|\mathfrak{a}\|_{\ell^2}^2$ and hence §2.1.13 holds with $\tau_{\psi\mathfrak{a}}^2 = 2\,\|\mathfrak{a}\|_{\ell^2}^2$. □

2.1.1 Abstract smoothness condition

§2.1.16 **Notations**. Let $\mathcal{U} = \{u_j, j \in \mathcal{J}\}$ be an ONS with $\mathbb{U} = \overline{lin}\{u_j, j \in \mathcal{J}\} \subseteq \mathbb{H}$. For any $h \in \mathbb{H}$ consider its associated sequence of generalised Fourier coefficients $[h] := ([h]_j)_{j\in\mathcal{J}}$ with generic elements $[h]_j = \langle h, u_j\rangle_{\mathbb{H}}$, $j \in \mathcal{J}$. Given a strictly positive sequence of weights $\mathfrak{v} = (\mathfrak{v}_j)_{j\in\mathcal{J}}$ for $h, g \in \mathbb{H}$ we define $\langle h, g\rangle_{\mathfrak{v}}^2 := \langle \mathfrak{v}[h], \mathfrak{v}[g]\rangle_{\ell^2} = \sum_{j\in\mathcal{J}} \mathfrak{v}_j^2[h]_j\overline{[g]}_j$ and $\|h\|_{\mathfrak{v}}^2 := \sum_{j\in\mathcal{J}} \mathfrak{v}_j^2|[h]_j|^2$. Obviously, $\langle\cdot,\cdot\rangle_{\mathfrak{v}}$ and $\|\cdot\|_{\mathfrak{v}}$ restricted on $\mathbb{U}$ defines on $\mathbb{U}$ a (weighted) *inner product* and its induced (weighted) *norm*, respectively. We denote by $\mathbb{U}_{\mathfrak{v}}$ the completion of $\mathbb{U}$

w.r.t. $\|\cdot\|_{\mathfrak{v}}$. If $(u_j)_{j\in\mathcal{J}}$ is complete in $\mathbb{H}$ then let $\mathbb{H}_{\mathfrak{v}}$ be the completion of $\mathbb{H}$ w.r.t. $\|\cdot\|_{\mathfrak{v}}$. □

§2.1.17 **Example** *(§2.1.15 continued)*. Consider the real Hilbert space $L^2([0,1])$ and the *trigonometric basis* $\{\psi_j, j\in\mathbb{N}\}$. Define further a weighted norm $\|\cdot\|_{\mathfrak{v}}$ w.r.t. the trigonometric basis, that is, $\|h\|_{\mathfrak{v}} := \sum_{j\in\mathbb{N}} \mathfrak{v}_j^2 |\langle h, \psi_j\rangle_{L^2}|^2$. Denote by $L^2_{\mathfrak{v}}([0,1])$ or $L^2_{\mathfrak{v}}$ for short, the completion of $L^2([0,1])$ w.r.t. $\|\cdot\|_{\mathfrak{v}}$.

(P) If we set $\mathfrak{v}_1 = 1$, $\mathfrak{v}_{2k} = \mathfrak{v}_{2k+1} = j^p$, $p\in\mathbb{N}$, $k\in\mathbb{N}$, then $L^2_{\mathfrak{v}}([0,1])$ is a subset of the *Sobolev space* of p-times differentiable periodic functions. Moreover, up to a constant, for any function $h\in L^2_{\mathfrak{v}}([0,1])$, the weighted norm $\|h\|^2_{\mathfrak{v}}$ equals the L^2-norm of its p-th weak derivative $h^{(p)}$ (Tsybakov [2009]).

(E) If, on the contrary, $\mathfrak{v}_j = \exp(-1+j^{2p})$, $p > 1/2$, $j\in\mathbb{N}$, then $L^2_{\mathfrak{v}}([0,1])$ is a *class of analytic functions* (Kawata [1972]).

Note that, the trigonometric basis is regular w.r.t. the weight sequence $1/\mathfrak{v} = \mathfrak{v}^{-1} = (\mathfrak{v}_j^{-1})$ as in §2.1.13 (ii), i.e., $\|1/\mathfrak{v}\|_{\ell^2} < \infty$, in case **(P)** whenever $p > 1/2$ and in case **(E)** if $p > 0$. □

§2.1.18 **Definition** *(Abstract smoothness condition)*. Given a strictly positive sequence of weights $\mathfrak{a} = (\mathfrak{a}_j)_{j\in\mathcal{J}}$ and an ONS $\mathcal{U} = \{u_j, j\in\mathcal{J}\}$ in $\mathbb{H}$ consider the associated weighted norm $\|\cdot\|_{1/\mathfrak{a}}$ and the completion $\mathbb{U}_{1/\mathfrak{a}}$ of $\mathbb{U}$. Let $r > 0$ be a constant. We assume in the following that the function of interest f belongs to the ellipsoid $\mathbb{F}^r_{\mathfrak{a}} := \{h\in\mathbb{U}_{1/\mathfrak{a}} : \|h\|^2_{1/\mathfrak{a}} \leqslant r^2\}$ and hence, $\Pi_{\mathbb{U}^\perp} f = 0$. □

§2.1.19 **Lemma**. *Let $\mathbb{F}^r_{\mathfrak{a}}$ be a class of functions w.r.t. an ONS $\mathcal{U} = \{u_j, j\in\mathcal{J}\}$ in L^2_μ (or analogously in ℓ^2) as given in §2.1.18. If the ONS is regular w.r.t. the weight sequence $\mathfrak{a}$ as in §2.1.13 (ii) for some finite constant $\tau_{u\mathfrak{a}} \geqslant 1$, then for each $f\in\mathbb{F}^r_{\mathfrak{a}}$ holds $\|f\|_{L^\infty_\mu} \leqslant \tau_{u\mathfrak{a}} \|f\|_{1/\mathfrak{a}} \leqslant r\tau_{u\mathfrak{a}}$.*

Proof of Lemma §2.1.19 is given in the lecture. □

§2.1.20 **Examples** *(§2.1.17 continued)*. Consider in $L^2_{\mathfrak{v}}([0,1])$ the *trigonometric basis* $\{\psi_j, j\in\mathbb{N}\}$ and a weight sequence $\mathfrak{v}$ satisfying either §2.1.17 **(P)** with $p > 1/2$ or §2.1.17 **(E)** with $p > 0$. In both cases setting $\tau^2_{\psi\mathfrak{v}} = 2\|1/\mathfrak{v}\|^2_{\ell^2} < \infty$ the trigonometric basis is regular w.r.t. the weight sequence $1/\mathfrak{v}$. Consequently, setting $\mathfrak{a} = 1/\mathfrak{v}$ and $\mathbb{F}^r_{\mathfrak{a}} = \{h\in L^2_{\mathfrak{v}}([0,1]) : \|h\|^2_{\mathfrak{v}} \leqslant r^2\}$, from Lemma §2.1.19 follows $\|f\|^2_{L^\infty} \leqslant 2\|f\|^2_{\mathfrak{v}} \|1/\mathfrak{v}\|^2_{\ell^2}$ for all $f\in\mathbb{F}^r_{\mathfrak{a}}$. □

2.2 Linear operator between Hilbert spaces

§2.2.1 **Definition**. A map $T:\mathbb{H}\to\mathbb{G}$ between Hilbert spaces $\mathbb{H}$ and $\mathbb{G}$ is called *linear operator* if $T(ah_1+bh_2) = aTh_1 + bTh_2$ for all $h_1, h_2\in\mathbb{H}, a, b\in\mathbb{K}$. Its *domain* will be denoted by $\mathcal{D}(T)$, its *range* by $\mathcal{R}(T)$ and its *null space* by $\mathcal{N}(T)$. □

§2.2.2 **Property**. *Let $T:\mathbb{H}\to\mathbb{G}$ be a linear operator, then the following assertions are equivalent:* (i) *T is continuous in zero.* (ii) *T is bounded, i.e., there is $M > 0$ such that $\|Th\|_{\mathbb{G}} \leqslant M\|h\|_{\mathbb{H}}$ for all $h\in\mathbb{H}$.* (iii) *T is uniformly continuous.* □

§2.2.3 **Definition**. The *class of all bounded linear operators* $T:\mathbb{H}\to\mathbb{G}$ is denoted by $\mathscr{L}(\mathbb{H},\mathbb{G})$, or $\mathscr{L}$ and in case of $\mathbb{H}=\mathbb{G}$, $\mathscr{L}(\mathbb{H})$ for short. For $T\in\mathscr{L}(\mathbb{H},\mathbb{G})$ define its *(uniform) norm* as $\|T\|_{\mathscr{L}} := \|T\|_{\mathscr{L}(\mathbb{H},\mathbb{G})} := \sup\{\|Th\|_{\mathbb{G}} : \|h\|_{\mathbb{H}} \leqslant 1, h\in\mathbb{H}\}$. □

§2.2.4 **Examples.**

(i) Let M be a $(m \times k)$ matrix, then $M \in \mathscr{L}(\mathbb{R}^k, \mathbb{R}^m)$. We write $\|M\|_s := \|M\|_{\mathscr{L}(\mathbb{R}^k,\mathbb{R}^m)}$ for short. *(spectral norm)*

(ii) For finite (i.e., $|\mathcal{J}| < \infty$) sequences $(h)_{j\in\mathcal{J}}$ in $\mathbb{H}$ and $(g)_{j\in\mathcal{J}}$ in $\mathbb{G}$ the linear operator $\sum_{j\in\mathcal{J}} h_j \otimes g_j$ defined by $f \mapsto \big[\sum_{j\in\mathcal{J}} h_j \otimes g_j\big] f := \sum_{j\in\mathcal{J}} \langle f, h_j\rangle_{\mathbb{H}}\, g_j$ belongs to $\mathscr{L}(\mathbb{H}, \mathbb{G})$ with $\|\sum_{j\in\mathcal{J}} h_j \otimes g_j\|_{\mathscr{L}} \leqslant \sum_{j\in\mathcal{J}} \|h_j\|_{\mathbb{H}} \|g_j\|_{\mathbb{G}}$. Moreover, it has a finite range contained in $\overline{lin}(\{g_j, j \in \mathcal{J}\})$.

(iii) Let $\mathcal{U} = \{u_j, j \in \mathcal{J}\}$ be an ONS in $\mathbb{H}$ and for any $f \in \mathbb{H}$ consider its *sequence of generalised Fourier coefficients* $[f] := ([f]_j)_{j\in\mathcal{J}}$ given by $[f]_j := \langle f, u_j\rangle_{\mathbb{H}}$, $j \in \mathcal{J}$. The associated *(generalised) Fourier series transform* U defined by $f \mapsto Uf := [f]$ belongs to $\mathscr{L}(\mathbb{H}, \ell^2(\mathcal{J}))$ with $\|U\|_{\mathscr{L}} = 1$.

(iv) For a sequence $\lambda = (\lambda_j)_{j\in\mathcal{J}}$ consider the *multiplication operator* $\mathrm{M}_\lambda : \mathbb{K}^{\mathcal{J}} \to \mathbb{K}^{\mathcal{J}}$ given by $x \mapsto \mathrm{M}_\lambda x := (\lambda_j x_j)_{j\in\mathcal{J}}$. For any bounded sequence λ, i.e, $\|\lambda\|_{\ell^\infty} < \infty$, we have $\|\mathrm{M}_\lambda\|_{\mathscr{L}(\ell^p)} \leqslant \|\lambda\|_{\ell^\infty}$ and hence, $\mathrm{M}_\lambda \in \mathscr{L}(\ell^p)$ for any $p \in [1, \infty]$. Analogously, given a function $\lambda : \Omega \to \mathbb{K}$ the *multiplication operator* $\mathrm{M}_\lambda : \mathbb{K}^\Omega \to \mathbb{K}^\Omega$ is defined as $f \mapsto \mathrm{M}_\lambda f := f\lambda$ where for any bounded (measurable) function λ, i.e, $\|\lambda\|_{L^\infty_\mu} < \infty$, holds $\|\mathrm{M}_\lambda\|_{\mathscr{L}(L^p_\mu)} \leqslant \|\lambda\|_{L^\infty_\mu} < \infty$ and, hence $\mathrm{M}_\lambda \in \mathscr{L}(L^p_\mu)$. On the other hand side, if λ is real-valued (measurable), μ-a.s. finite and non zero, then the subset $\mathcal{D}(\mathrm{M}_\lambda) := \big\{f \in L^2_\mu : \lambda f \in L^2_\mu\big\}$ is dense in L^2_μ. In this situation the *multiplication operator* $\mathrm{M}_\lambda : L^2_\mu \supset \mathcal{D}(\mathrm{M}_\lambda) \to L^2_\mu$ is densely defined (and self-adjoint).

(v) Given a (generalised) Fourier series transform $U \in \mathscr{L}(\mathbb{H}, \ell^2)$ as in (iii) and a multiplication operator $\mathrm{M}_\lambda \in \mathscr{L}(\ell^2)$ for some bounded sequence $\lambda = (\lambda_j)_{j\in\mathcal{J}}$ as in (iv) the linear operator $\nabla_\lambda : \mathbb{H} \to \mathbb{H}$ given by $\mathcal{N}(U) = \mathcal{N}(\nabla_\lambda)$ and $U\nabla_\lambda = \mathrm{M}_\lambda U$, i.e. $U\nabla_\lambda h = \mathrm{M}_\lambda U h = (\lambda_j [h]_j)_{j\in\mathcal{J}}$ belongs to $\mathscr{L}(\mathbb{H})$ with $\|\nabla_\lambda\|_{\mathscr{L}} \leqslant \|\lambda\|_{\ell^\infty} < \infty$. We call ∇_λ *diagonal* w.r.t. U (or $\mathcal{U}$).

(vi) The *integral operator* $T_k : L^2_{\mu_1}(\Omega_1) \to L^2_{\mu_2}(\Omega_2)$ with kernel $k : \Omega_1 \times \Omega_2 \to \mathbb{K}$ defined by

$$[T_k f](\omega_2) := \int_{\Omega_1} h(\omega_1) k(\omega_1, \omega_2) \mu(d\omega_1),\ \omega_2 \in \Omega_2,\ h \in L^2_{\mu_1}(\Omega_1),$$

belongs to $\mathscr{L}(L^2_{\mu_1}(\Omega_1), L^2_{\mu_2}(\Omega_2))$ if $\|k\|^2_{L^2} = \int_{\Omega_1}\int_{\Omega_2} |k|^2 d\mu_1 d\mu_2 < \infty$.

(vii) Let $X \in L^1(\Omega, \mathscr{A}, \mathbb{P})$ and $\mathscr{F} \subset \mathscr{A}$ be a sub-σ-algebra. There exists $Y \in L^1(\Omega, \mathscr{F}, \mathbb{P})$ such that $\mathbb{E}(X\mathbb{1}_F) = \mathbb{E}(Y\mathbb{1}_F)$ for all $F \in \mathscr{F}$, moreover, Y is unique up to equality $\mathbb{P}$-a.s.. Each version Y is called conditional expectation of X given $\mathscr{F}$, symbolically, $\mathbb{E}[X|\mathscr{F}] := Y$. For each $p \in [1, \infty]$ the linear map $\mathbb{E}[\bullet|\mathscr{F}] : L^p(\Omega, \mathscr{A}, \mathbb{P}) \to L^p(\Omega, \mathscr{F}, \mathbb{P}) \subseteq L^p(\Omega, \mathscr{A}, \mathbb{P})$ given by $X \mapsto \mathbb{E}[X|\mathscr{F}]$ is a contraction, that is $\|\mathbb{E}[X|\mathscr{F}]\|_{L^p} \leqslant \|X\|_{L^p}$ and thus $\mathbb{E}[\bullet|\mathscr{F}]$ belongs to $\mathscr{L}(L^p(\Omega, \mathscr{A}, \mathbb{P}))$ with $\|\mathbb{E}[\bullet|\mathscr{F}]\|_{\mathscr{L}} = 1$ (keep in mind that $\mathbb{E}[1|\mathscr{F}] = 1$). Given a r.v. Z on $(\Omega, \mathscr{A}, \mathbb{P})$ and the σ-algebra $\sigma(Z)$ generated by Z we set $\mathbb{E}[X|Z] := \mathbb{E}[X|\sigma(Z)]$. The *conditional expectation operator* of X given Z defined by $K_{X|Z} h := \mathbb{E}[h(X)|Z]$ for $h \in L^1_X$ is then an element of $\mathscr{L}(L^p_X, L^p_Z)$ with $\|K_{X|Z}\|_{\mathscr{L}} = 1$.

(viii) Let $g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then the *convolution operator* $C_g : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ defined by

$$[C_g h](t) := [g * h](t) := \int_{\mathbb{R}} g(t-s) h(s) ds,\ t \in \mathbb{R},\ h \in L^2(\mathbb{R}),$$

belongs to $\mathscr{L}(L^2(\mathbb{R}))$ with $\|C_g\|_{\mathscr{L}} \leqslant \|g\|_{L^1} := \int_{\mathbb{R}} |g(t)| dt$.

(ix) Let $g \in L^2([0,1))$, hence, $g \in L^1([0,1))$, and let $\lfloor\cdot\rfloor$ be the floor function, then the *circular convolution operator* $C_g : L^2([0,1)) \to L^2([0,1))$ defined by

$$[C_g h](t) := [g \circledast h](t) := \int_{[0,1)} g(t-s-\lfloor t-s\rfloor)h(s)ds,\ t \in [0,1),\ h \in L([0,1)),$$

belongs to $\mathscr{L}(L^2([0,1)))$ with $\|C_g\|_{\mathscr{L}} \leqslant \|g\|_{L^1} := \int_0^1 |g(t)|dt$. □

§2.2.5 **Definition**. A (linear) map $\Phi : \mathbb{H} \supset \mathcal{D}(\Phi) \to \mathbb{K}$ is called *(linear) functional* and given an ONS $\{u_j, j \in \mathcal{J}\}$ in $\mathbb{H}$ which belongs to $\mathcal{D}(\Phi)$ we set $[\Phi] = ([\Phi]_j)_{j\in\mathcal{J}}$ with the slight abuse of notations $[\Phi]_j := \Phi(u_j)$. In particular, if $\Phi \in \mathscr{L}(\mathbb{H}, \mathbb{K})$ then $\mathcal{D}(\Phi) = \mathbb{H}$. □

§2.2.6 **Property**. *Let $\Phi \in \mathscr{L}(\mathbb{H}, \mathbb{K})$.*

(Fréchet-Riesz representation) There exists a function $\phi \in \mathbb{H}$ such that $\Phi(h) = \langle \phi, h\rangle_{\mathbb{H}}$ for all $h \in \mathbb{H}$, and hence, given an ONS $\{u_j, j \in \mathcal{J}\}$ in $\mathbb{H}$ we have $[\Phi]_j = [\phi]_j$ for all $j \in \mathcal{J}$. □

§2.2.7 **Example**. Consider an ONB $\mathcal{U} = \{u_j, j \in \mathcal{J}\}$ in $L^2(\Omega)$ (or analogously in $\ell^2(\mathcal{J})$). By *evaluation at a point* $t_o \in \Omega$ we mean the linear functional Φ_{t_o} mapping $h \in L^2(\Omega)$ to $h(t_o) := \Phi_{t_o}(h) = \sum_{j\in\mathcal{J}}[h]_j u_j(t_o)$. Obviously, a point evaluation of h at t_o is well-defined, if $\sum_{j\in\mathcal{J}} |[h]_j u_j(t_o)| < \infty$. Observe that the point evaluation at t_o is generally not bounded on the subset $\{h \in L^2(\Omega) : \sum_{j\in\mathcal{J}} |[h]_j u_j(t_o)| < \infty\}$. □

§2.2.8 **Definition** *(Regular linear functionals)*. Consider an ONS $\mathcal{U} = \{u_j, j \in \mathcal{J}\}$ in $\mathbb{H}$ which belongs to the domain $\mathcal{D}(\Phi)$ of a linear functional Φ. In order to guarantee that $\mathbb{U}_{1/\mathfrak{a}}$ and hence the class $\mathbb{F}^r_{\mathfrak{a}}$ of functions of interest as in §2.1.18 are contained in $\mathcal{D}(\Phi)$ and that $\Phi(f) = \sum_{j\in\mathcal{J}}[\Phi]_j[f]_j$ holds for all $f \in \mathbb{F}^r_{\mathfrak{a}}$, it is sufficient that $\|[\Phi]\|^2_{\ell^2_{\mathfrak{a}}} = \sum_{j\in\mathcal{J}} |[\Phi]_j|^2\mathfrak{a}_j^2 < \infty$. Indeed, $|\Phi(f)|^2 \leqslant \|f\|^2_{1/\mathfrak{a}} \|[\Phi]\|^2_{\ell^2_{\mathfrak{a}}}$ for any $f \in \mathbb{U}_{1/\mathfrak{a}}$ and hence $\Phi \in \mathscr{L}(\mathbb{U}_{1/\mathfrak{a}}, \mathbb{K})$ with $\|\Phi\|_{\mathscr{L}} \leqslant \|[\Phi]\|_{\ell^2_{\mathfrak{a}}}$. We denote by $\mathcal{L}_{\mathfrak{a}}$ the set of all linear functionals with $\|[\Phi]\|^2_{\ell^2_{\mathfrak{a}}} < \infty$. □

§2.2.9 **Remark**. We may emphasise that we neither impose that the sequence $[\Phi] = ([\Phi]_j)_{j\in\mathcal{J}}$ tends to zero nor that it is square summable. The assumption $\Phi \in \mathcal{L}_{\mathfrak{a}}$, however, enables us in specific cases to deal with more demanding functionals, such as in Example §2.2.7 above the evaluation at a given point. □

§2.2.10 **Example** *(§2.2.7 continued)*. Consider an ONB $\mathcal{U} = \{u_j, j \in \mathcal{J}\}$ in $L^2(\Omega)$ and the *evaluation at a point* $t_o \in \Omega$ given by $\Phi_{t_o}(h) = \sum_{j\in\mathcal{J}}[h]_j u_j(t_o)$. Let $L^2_{1/\mathfrak{a}}(\Omega)$ be the completion of $L^2(\Omega)$ w.r.t. a weighted norm $\|\cdot\|_{1/\mathfrak{a}}$ derived from $\mathcal{U}$ and a strictly positive sequence $\mathfrak{a}$. Since $|\Phi_{t_o}(h)|^2 \leqslant \|h\|^2_{1/\mathfrak{a}} \sum_{j\in\mathcal{J}} \mathfrak{a}_j^2|u_j(t_o)|^2$ the point evaluation in t_o is bounded on $L^2_{1/\mathfrak{a}}(\Omega)$ and, thus, belongs to $\mathscr{L}(L^2_{1/\mathfrak{a}}(\Omega), \mathbb{K})$, if $\sum_{j\in\mathcal{J}} \mathfrak{a}_j^2|u_j(t_o)|^2 < \infty$. Consequently, if the ONS $\mathcal{U}$ is regular w.r.t. the weight sequence $\mathfrak{a}$, i.e., §2.1.13 (ii) holds for some finite constant $\tau_{u\mathfrak{a}} \geqslant 1$, then $\|\Phi_{t_o}\|_{\mathscr{L}(L^2_{1/\mathfrak{a}}(\Omega),\mathbb{K})} \leqslant \tau_{u\mathfrak{a}}$ uniformly for any $t_o \in \Omega$. Revisiting the particular situation of Example §2.1.17 and its continuation in §2.1.20, that is, $L^2_{\mathfrak{v}}([0,1])$ w.r.t. the *trigonometric basis* $\{\psi_j, j \in \mathbb{N}\}$ and weight sequence $\mathfrak{v}$ satisfying either §2.1.17 **(P)** with $p > 1/2$ or §2.1.17 **(E)** with $p > 0$, recall that the trigonometric basis is regular w.r.t. $\mathfrak{a} = 1/\mathfrak{v}$ and hence, the point evaluation Φ_{t_o} belongs to $\mathscr{L}(L^2_{\mathfrak{v}}([0,1]), \mathbb{R})$, i.e., $\|\Phi_{t_o}\|_{\mathscr{L}} \leqslant \sqrt{2}\,\|1/\mathfrak{v}\|_{\ell^2}$ for each $t_o \in [0,1]$. □

§2.2.11 **Definition**. If $T \in \mathscr{L}(\mathbb{H}, \mathbb{G})$, then there exists a uniquely determined *adjoint operator* $T^\star \in \mathscr{L}(\mathbb{G}, \mathbb{H})$ satisfying $\langle Th, g\rangle_{\mathbb{G}} = \langle h, T^\star g\rangle_{\mathbb{H}}$ for all $h \in \mathbb{H}, g \in \mathbb{G}$. □

§2.2.12 **Properties**. *Let $S, T \in \mathscr{L}(\mathbb{H}_1, \mathbb{H}_2)$ and $R \in \mathscr{L}(\mathbb{H}_2, \mathbb{H}_3)$. Then we have*

(i) $(S+T)^\star = S^\star + T^\star$, $(RS)^\star = S^\star R^\star$.

(ii) $\|S^\star\|_{\mathscr{L}} = \|S\|_{\mathscr{L}}$, $\|SS^\star\|_{\mathscr{L}} = \|S^\star S\|_{\mathscr{L}} = \|S\|_{\mathscr{L}}^2$.

(iii) $\mathcal{N}(S) = \mathcal{R}(S^\star)^\perp$, $\mathcal{N}(S^\star) = \mathcal{R}(S)^\perp$, $\mathbb{H}_1 = \mathcal{N}(S) \oplus \overline{\mathcal{R}}(S^\star)$ *and* $\mathbb{H}_2 = \mathcal{N}(S^\star) \oplus \overline{\mathcal{R}}(S)$ *where $\overline{\mathcal{R}}(S)$ (respectively, $\overline{\mathcal{R}}(S^\star)$) denotes the closure of the range of S. In particular, S is injective if and only if $\mathcal{R}(S^\star)$ is dense in $\mathbb{H}$.*

(iv) $\mathcal{N}(S^\star S) = \mathcal{N}(S)$ *and* $\mathcal{N}(SS^\star) = \mathcal{N}(S^\star)$. □

§2.2.13 **Examples** *(§2.2.4 continued).*

(i) The adjoint of a $(k \times m)$ matrix M is its $(m \times k)$ transpose matrix M^t.

(ii) The adjoint $U^\star \in \mathscr{L}(\ell^2(\mathcal{J}), \mathbb{H})$ of the *(generalised) Fourier series transform* $U \in \mathscr{L}(\mathbb{H}, \ell^2(\mathcal{J}))$ satisfies $x \mapsto U^\star x := \sum_{j \in \mathcal{J}} x_j u_j$ for $x \in \ell^2(\mathcal{J})$.

(iii) For finite $\mathcal{J}$ the adjoint operator in $\mathscr{L}(\mathbb{G}, \mathbb{H})$ of $\sum_{j \in \mathcal{J}} h_j \otimes g_j \in \mathscr{L}(\mathbb{H}, \mathbb{G})$ satisfies $[\sum_{j \in \mathcal{J}} h_j \otimes g_j]^\star g = \sum_{j \in \mathcal{J}} \langle g, g_j \rangle_{\mathbb{G}} h_j = [\sum_{j \in \mathcal{J}} g_j \otimes h_j] g$.

(iv) Let $\mathrm{M}_\lambda \in \mathscr{L}(L^2_\mu(\Omega))$ (or analogously $\mathrm{M}_\lambda \in \mathscr{L}(\ell^2)$) be a *multiplication operator*, then its adjoint operator $\mathrm{M}_\lambda^\star = \mathrm{M}_{\lambda^\star}$ is a multiplication operator with $\lambda^\star(t) = \overline{\lambda(t)}$, $t \in \Omega$.

(v) Let $T_k \in \mathscr{L}(L^2_{\mu_1}(\Omega_1), L^2_{\mu_2}(\Omega_2))$ be an *integral operator* with kernel k, then its adjoint $T_k^\star = T_{k^\star} \in \mathscr{L}(L^2_{\mu_2}(\Omega_2), L^2_{\mu_1}(\Omega_1))$ is again an integral operator satisfying

$$[T_{k^\star} g](\omega_1) := \int_{\Omega_2} g(\omega_2) k^\star(\omega_2, \omega_1) \mu_2(d\omega_2),\ \omega_1 \in \Omega_1,\ g \in L^2_{\mu_2}(\Omega_2),$$

with kernel $k^\star(\omega_2, \omega_1) := \overline{k(\omega_1, \omega_2)}$, $\omega_1 \in \Omega_1$, $\omega_2 \in \Omega_2$.

(vi) Let $K_{X|Z} \in \mathscr{L}(L^2_X, L^2_Z)$ be the *conditional expectation* of X given Z, then its adjoint operator $K_{X|Z}^\star = K_{Z|X} \in \mathscr{L}(L^2_Z, L^2_X)$ is the conditional expectation of Z given X satisfying $K_{Z|X} g = \mathbb{E}[g(Z)|X]$ for all $g \in L^2_Z$.

(vii) Let $C_g \in \mathscr{L}(L^2(\mathbb{R}))$ be the *convolution operator*, then its adjoint operator $C_g^\star = C_{g^\star}$ is a convolution operator, i.e, $C_{g^\star} h = g^\star * h$, with $g^\star(t) = \overline{g(-t)}$, $t \in \mathbb{R}$. □

§2.2.14 **Definition**.

(i) The *identity* in $\mathscr{L}(\mathbb{H})$ is denoted by $\mathrm{Id}_{\mathbb{H}}$.

(ii) Let $T \in \mathscr{L}(\mathbb{H}, \mathbb{G})$. Obviously, $T : \mathcal{N}(T)^\perp \to \mathcal{R}(T)$ is bijective and continuous whereas its *inverse* $T^{-1} : \mathcal{R}(T) \to \mathcal{N}(T)^\perp$ is continuous (i.e. bounded) if and only if $\mathcal{R}(T)$ is closed. In particular, if $T : \mathbb{H} \to \mathbb{G}$ is bijective (invertible) then its inverse $T^{-1} \in \mathscr{L}(\mathbb{G}, \mathbb{H})$ satisfies $\mathrm{Id}_{\mathbb{G}} = TT^{-1}$ and $\mathrm{Id}_{\mathbb{H}} = T^{-1}T$.

(iii) $U \in \mathscr{L}(\mathbb{H}, \mathbb{G})$ is called *unitary*, if U is invertible with $UU^\star = \mathrm{Id}_{\mathbb{G}}$ and $U^\star U = \mathrm{Id}_{\mathbb{H}}$.

(iv) $V \in \mathscr{L}(\mathbb{H}, \mathbb{G})$ is called *partial isometry*, if $V : \mathcal{N}(V)^\perp \to \mathcal{R}(V)$ is unitary.

(v) $T \in \mathscr{L}(\mathbb{H})$ is called *self-adjoint*, if $T = T^\star$, i.e., $\langle Th, g \rangle_{\mathbb{H}} = \langle h, T^\star g \rangle_{\mathbb{H}}$ for all $h, g \in \mathbb{H}$.

(vi) $T \in \mathscr{L}(\mathbb{H})$ is called *normal*, if $TT^\star = T^\star T$, i.e., $\langle Th, Tg \rangle_{\mathbb{H}} = \langle T^\star h, T^\star g \rangle_{\mathbb{H}}$ for all $h, g \in \mathbb{H}$.

(vii) A self-adjoint $T \in \mathscr{L}(\mathbb{H})$ is called *non-negative* or $T \geqslant 0$ for short, if $\langle Th, h \rangle_{\mathbb{H}} \geqslant 0$ for all $h \in \mathbb{H}$ and *strictly positive* or $T > 0$ for short, if $\langle Th, h \rangle_{\mathbb{H}} > 0$ for all $h \in \mathbb{H} \backslash \{0\}$.

(viii) $\Pi \in \mathscr{L}(\mathbb{H})$ is called *projection* if $\Pi^2 = \Pi$. For $\Pi \neq 0$ are equivalent: (a) Π is an orthogonal projection ($\mathbb{H} = \mathcal{R}(\Pi) \oplus \mathcal{N}(\Pi)$); (b) $\|\Pi\|_{\mathscr{L}} = 1$; (c) Π is non-negative. □

§2.2.15 **Property**. *Let $T \in \mathscr{L}(\mathbb{H})$. If T is invertible, then it is $T^\star$, where $(T^{-1})^\star = (T^\star)^{-1}$. Moreover, if T is normal, then $\|T\|_{\mathscr{L}} = \sup\{|\langle Th, h\rangle_{\mathbb{H}}| : \|h\|_{\mathbb{H}} \leqslant 1, h \in \mathbb{H}\}$.*

(Neumann series) If $\|T\|_{\mathscr{L}} < 1$, then $\|(\mathrm{Id}_{\mathbb{H}} - T)^{-1}\|_{\mathscr{L}} \leqslant (1 - \|T\|_{\mathscr{L}})^{-1}$. □

§2.2.16 **Examples** *(§2.2.4 continued).*

(i) The *(generalised) Fourier series transform* U is a partial isometry with adjoint operator $U^\star x = \sum_{j\in\mathcal{J}} x_j u_j$ for $x \in \ell^2(\mathcal{J})$. Moreover, the orthogonal projection $\Pi_{\mathbb{U}}$ onto $\mathbb{U}$ satisfies $\Pi_{\mathbb{U}} f = U^\star U f = \sum_{j\in\mathcal{J}} [f]_j u_j$ for all $f \in \mathbb{H}$. If $\mathcal{U} = \{u_j, j \in \mathcal{J}\}$ is complete (i.e. ONB), then U is invertible with $UU^\star = \mathrm{Id}_{\ell^2}$ and $U^\star U = \mathrm{Id}_{\mathbb{H}}$ due to Parseval's formula, and hence U is unitary.

(ii) Let $\mathcal{F} \in \mathscr{L}(L^2(\mathbb{R}))$ denote the *Fourier-Plancherel transform* satisfying

$$[\mathcal{F}h](t) = \int_{\mathbb{R}} h(x) e^{-\iota 2\pi x t} dx, \quad \forall h \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}).$$

Then $\mathcal{F}$ is unitary with $[\mathcal{F}^\star h](t) = \int h(x) e^{\iota 2\pi x t} dx$ for all $h \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. We note further for all $h \in L^1$ that $\|\mathcal{F}h\|_{L^\infty} \leqslant \|h\|_{L^1}$, and that $\mathcal{F}h$ is continuous and tends to zero in infinity. Keeping in mind the convolution defined in Examples §2.2.4 (viii) the convolution theorem states $\mathscr{F}(f * g) = \mathscr{F} f \cdot \mathscr{F} g$ for any $f, g \in L^1(\mathbb{R})$.

(iii) A *multiplication operator* $\mathrm{M}_\lambda \in \mathscr{L}(L^2_\mu)$ is normal. If λ is in addition real, it is self-adjoint and if λ is non-negative, then it is non-negative.

(iv) A *diagonal operator* $\nabla_\lambda \in \mathscr{L}(\mathbb{H})$ w.r.t. a partial isometry $U \in \mathscr{L}(\mathbb{H}, \ell^2)$ satisfies $\nabla_\lambda = U^\star \mathrm{M}_\lambda U$ and it shares the properties of the *multiplication operator* $\mathrm{M}_\lambda \in \mathscr{L}(\ell^2)$.

(v) A *conditional expectation operator* $K_{Z|X} \in \mathscr{L}(L^2_X, L^2_Z)$ is an orthogonal projection.

(vi) A *convolution operator* $C_g \in \mathscr{L}(L^2(\mathbb{R}))$ is normal and if g is in addition a real and even ($g(-t) = g(t)$) function, then it is self-adjoint.

(vii) A *circular convolution operator* $C_g \in \mathscr{L}(L^2([0,1)))$ is normal and if g is in addition a real and even ($g(t) = g(1-t)$) function, then it is self-adjoint. □

2.2.1 Compact, nuclear and Hilbert-Schmidt operator

§2.2.17 **Definition**. An operator $K \in \mathscr{L}(\mathbb{H}, \mathbb{G})$ is called *compact*, if $\{Kh : \|h\|_{\mathbb{H}} \leqslant 1, h \in \mathbb{H}\}$ is relatively compact in $\mathbb{G}$. We denote by $\mathscr{K}(\mathbb{H}, \mathbb{G})$ the *subset of all compact operator* in $\mathscr{L}(\mathbb{H}, \mathbb{G})$, and we write $\mathscr{K}(\mathbb{H}) = \mathscr{K}(\mathbb{H}, \mathbb{H})$ for short. □

§2.2.18 **Properties**. *Let $K \in \mathscr{L}(\mathbb{H}, \mathbb{G})$.*

(Schauder's theorem) K is compact, if and only if its adjoint $K^\star \in \mathscr{L}(\mathbb{G}, \mathbb{H})$ is compact.

If there are $K_j \in \mathscr{L}(\mathbb{H}, \mathbb{G})$ with finite dimensional range for each $j \in \mathbb{N}$ such that $\lim_{j\to\infty} \|K_j - K\|_{\mathscr{L}} = 0$, then K is compact. If in addition $\mathbb{G}$ is separable, then the converse holds also true. □

§2.2.19 **Examples** *(§2.2.4 continued).*

(i) For finite $\mathcal{J}$ the operator $\sum_{j\in\mathcal{J}} h_j \otimes g_j \in \mathscr{L}(\mathbb{H}, \mathbb{G})$ is compact.

(ii) A *multiplication operator* $\mathrm{M}_\lambda \in \mathscr{L}(\ell^2)$ is compact, if λ has either only a finite number of entries not equal to zero or zero is the only accumulation point.

(iii) A *diagonal operator* $\nabla_\lambda = U^\star \mathrm{M}_\lambda U \in \mathscr{L}(\mathbb{H})$ w.r.t. a partial isometry $U \in \mathscr{L}(\mathbb{H}, \ell^2)$ is compact if the multiplication operator $\mathrm{M}_\lambda \in \mathscr{L}(\ell^2)$ is compact.

(iv) A *convolution operator* $C_g \in \mathscr{L}(L^2(\mathbb{R}))$ is not compact.

(v) A *circular convolution operator* $C_g \in \mathscr{L}(L^2([0, 1)))$ is compact. □

§2.2.20 **Remark.** Every finite linear combination of compact operators is compact, and hence $\mathscr{K}(\mathbb{H}, \mathbb{G})$ is a vector space. □

§2.2.21 **Definition.** An operator $T \in \mathscr{L}(\mathbb{H}, \mathbb{G})$ is called *nuclear*, if there are sequences $(h_j)_{j\in\mathbb{N}}$ in $\mathbb{H}$ and $(g_j)_{j\in\mathbb{N}}$ in $\mathbb{G}$ with $\sum_{j\in\mathbb{N}} \|h_j\|_{\mathbb{H}} \|g_j\|_{\mathbb{G}} < \infty$ such that $\lim_{n\to\infty} \|\sum_{j=1}^n h_j \otimes g_j - T\|_{\mathscr{L}} = 0$, or $T = \sum_{j\in\mathbb{N}} h_j \otimes g_j$ for short. We denote by $\mathscr{N}(\mathbb{H}, \mathbb{G})$ the *subset of all nuclear operator* in $\mathscr{L}(\mathbb{H}, \mathbb{G})$, and we write $\mathscr{N}(\mathbb{H}) := \mathscr{N}(\mathbb{H}, \mathbb{H})$. Furthermore, let $(f_j)_{j\in\mathbb{N}}$ be any ONB in $\mathbb{H}$ and $T \in \mathscr{N}(\mathbb{H})$, then $\mathrm{tr}(T) := \sum_{j\in\mathbb{N}} \langle T f_j, f_j\rangle_{\mathbb{H}}$ denotes the *trace* of T. □

§2.2.22 **Remark.** We have $\mathscr{N}(\mathbb{H}, \mathbb{G}) \subset \mathscr{K}(\mathbb{H}, \mathbb{G}) \subset \mathscr{L}(\mathbb{H}, \mathbb{G})$. The trace does not depend on the choice of the ONB and is a continuous linear functional on $\mathscr{N}(\mathbb{H})$ with $\|\mathrm{tr}\|_{\mathscr{L}} = 1$. □

§2.2.23 **Properties.** *Let $T \in \mathscr{L}(\mathbb{H}, \mathbb{G})$ and $S \in \mathscr{L}(\mathbb{G}, \mathbb{H})$.*

(i) *T is nuclear, if and only if its adjoint $T^\star \in \mathscr{L}(\mathbb{G}, \mathbb{H})$ is nuclear.*

(ii) *If T is nuclear, then $TS \in \mathscr{N}(\mathbb{H})$, $ST \in \mathscr{N}(\mathbb{G})$ and $\mathrm{tr}(TS) = \mathrm{tr}(ST)$.* □

§2.2.24 **Example.** A *multiplication operator* $\mathrm{M}_\lambda \in \mathscr{L}(\ell^2)$ and, hence an associated *diagonal operator* $U^\star \mathrm{M}_\lambda U \in \mathscr{L}(\mathbb{H})$, is nuclear, if λ is absolute summable, i.e., $\|\lambda\|_{\ell^1} < \infty$, and $\mathrm{tr}(\mathrm{M}_\lambda) = \mathrm{tr}(\nabla_\lambda) = \sum_{j\in\mathcal{J}} \lambda_j$. □

§2.2.25 **Definition.** An operator $T \in \mathscr{L}(\mathbb{H}, \mathbb{G})$ is called *Hilbert-Schmidt*, if there exists an ONB $(h_j)_{j\in\mathbb{N}}$ in $\mathbb{H}$ such that $\|T\|^2_{\mathscr{H}} := \sum_{j\in\mathbb{N}} \|T h_j\|^2_{\mathbb{G}} < \infty$. The number $\|T\|_{\mathscr{H}}$ is called Hilbert-Schmidt norm of T and satisfies $\|T\|_{\mathscr{L}} \leqslant \|T\|_{\mathscr{H}}$. We denote by $\mathscr{H}(\mathbb{H}, \mathbb{G})$ the *subset of all Hilbert-Schmidt operator* in $\mathscr{L}(\mathbb{H}, \mathbb{G})$, and we write $\mathscr{H}(\mathbb{H}) := \mathscr{H}(\mathbb{H}, \mathbb{H})$. □

§2.2.26 **Remark.** We have $\mathscr{N}(\mathbb{H}, \mathbb{G}) \subset \mathscr{H}(\mathbb{H}, \mathbb{G}) \subset \mathscr{K}(\mathbb{H}, \mathbb{G})$. The number $\|T\|_{\mathscr{H}}$ does not depend on the choice of the ONB. The product TS of two Hilbert-Schmidt operator T and S is nuclear. The space $\mathscr{H}(\mathbb{H}, \mathbb{G})$ endowed with the inner product $\langle T, S\rangle_{\mathscr{H}} := \mathrm{tr}(S^\star T)$, $S, T \in \mathscr{H}(\mathbb{H}, \mathbb{G})$ is a Hilbert space and $\|\cdot\|_{\mathscr{H}}$ the induced norm. □

§2.2.27 **Property.** *If $T \in \mathcal{H}(\mathbb{H}, \mathbb{G})$ and $S \in \mathscr{L}(\mathbb{G})$ then $\mathrm{tr}(TST^\star) \leqslant \mathrm{tr}(TT^\star) \|S\|_{\mathscr{L}}$.* □

§2.2.28 **Examples.**

(i) Let $T \in \mathscr{L}(L^2_{\mu_1}(\Omega_1), L^2_{\mu_2}(\Omega_2))$. The operator T is Hilbert-Schmidt if and only if it is an *integral operator* $T = T_k$ with square integrable kernel k and it holds $\|T\|^2_{\mathscr{H}} = \int_{\Omega_1}\int_{\Omega_2} |k(\omega_1, \omega_2)|^2 \mu_1(d\omega_1)\mu_2(d\omega_2)$.

(ii) A *multiplication operator* $\mathrm{M}_\lambda \in \mathscr{L}(\ell(\mathcal{J}))$ and, hence an associated *diagonal operator* $U^\star \mathrm{M}_\lambda U \in \mathscr{L}(\mathbb{H})$, is Hilbert-Schmidt, if $\lambda = (\lambda_j)_{j\in\mathcal{J}}$ is square summable and $\|\mathrm{M}_\lambda\|_{\mathscr{H}} =$

$\|\nabla_\lambda\|_{\mathscr{H}} = \|\lambda\|_{\ell^2} < \infty$.

(iii) Consider the *conditional expectation operator* $K_{X|Z} \in \mathscr{L}(L^2_X, L^2_Z)$ of X given Z. Let in addition $p_{X,Z}$, p_X and p_Z be, respectively, the joint and marginal densities of (X, Z), X and Z w.r.t. a σ-finite measure. In this situation, the operator $K_{X|Z}$ is Hilbert Schmidt if and only if $\mathbb{E}\big[\frac{|p_{XZ}(X,Z)|^2}{|p_X(X)p_Z(Z)|^2}\big] < \infty$. □

2.2.2 Spectral theory and functional calculus

§2.2.29 **Definition**. Consider $T \in \mathscr{L}(\mathbb{H})$. The set $\rho(T) = \{\lambda \in \mathbb{K} : (\lambda \operatorname{Id}_{\mathbb{H}} - T)^{-1} \in \mathscr{L}(\mathbb{H})\}$ and its complement $\sigma(T) = \mathbb{K}\backslash\rho(T)$ is called *resolvent set* and *spectrum* of T, respectively. The subset $\sigma_p(T) = \{\lambda \in \mathbb{K} : \lambda \operatorname{Id}_{\mathbb{H}} - T \text{ is not injective}\}$ of $\sigma(T)$ is called *point spectrum* of T. An element λ of $\sigma_p(T)$ and $h \in \mathbb{H}\backslash\{0\}$ with $Th = \lambda h$ is called *eigenvalue* and *eigenfunction* (eigenvector), respectively. □

§2.2.30 **Properties**. *Consider $T \in \mathscr{K}(\mathbb{H})$.*

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

(ii) *If $\mathbb{H}$ is infinite dimensional, then $0 \in \sigma(T)$.*

(iii) *The (possibly empty) set $\sigma(T)\backslash\{0\}$ is at most countable.*

(iv) *Any $\lambda \in \sigma(T)\backslash\{0\}$ is an eigenvalue of T and its multiplicity is the (finite) dimension of the associated eigenspace $\mathcal{N}(\lambda \operatorname{Id}_{\mathbb{H}} - T)$.*

(v) *In $\sigma(T)$ the only possible accumulation point is zero.* □

§2.2.31 **Example**. The spectrum of a *multiplication operator* $\mathrm{M}_\lambda \in \mathscr{K}(\ell^2)$ and its associated *diagonal operator* $\nabla_\lambda = U^\star \mathrm{M}_\lambda U \in \mathscr{K}(\mathbb{H})$ is given by $\sigma(\mathrm{M}_\lambda) = \sigma(\nabla_\lambda) = \{\lambda_j, j \in \mathcal{J}\} \subset \mathbb{K}$. □

§2.2.32 **Definition**. Let $T \in \mathscr{K}(\mathbb{H})$ be normal ($\mathbb{K} = \mathbb{C}$) or self-adjoint ($\mathbb{K} = \mathbb{R}$). There exist

(i) a sequence $\lambda = (\lambda_j)_{j\in\mathcal{J}}$ in $\mathbb{K}\backslash\{0\}$ with $\|T\|_{\mathscr{L}} = \sup_{j\in\mathcal{J}} |\lambda_j|$ which has either a finite number of entries or zero as accumulation point, and determines a multiplication operator $\mathrm{M}_\lambda \in \mathscr{L}(\ell^2(\mathcal{J}))$,

(ii) an ONS $\{u_j, j \in \mathcal{J}\}$ in $\mathbb{H}$ with $\mathbb{U} := \overline{\mathrm{lin}}\{u_j, j \in \mathcal{J}\}$ and associated generalised Fourier series transform $\mathcal{U} \in \mathscr{L}(\mathbb{H}, \ell^2(\mathcal{J}))$ as defined in §2.2.4,

such that $\mathbb{H} = \mathcal{N}(T) \oplus \mathbb{U}$ and $T = \sum_{j\in\mathcal{J}} \lambda_j\, u_j \otimes u_j = \mathcal{U}^\star \mathrm{M}_\lambda \mathcal{U} = \nabla_\lambda$ (see §2.2.4 (ii), (iv) and (v)). For $j \in \mathcal{J}$, λ_j and u_j are, respectively, a non-zero *eigenvalue* and *associated eigenvector* of T respectively. $\{(\lambda_j, u_j), j \in \mathcal{J}\}$ is called an *eigensystem* of T. □

§2.2.33 **Properties**. *Let $T \in \mathscr{K}(\mathbb{H})$ be self-adjoint with eigensystem $\{(\lambda_j, u_j), j \in \mathcal{J}\}$, i.e., $\sigma(T)\backslash\{0\} = \{\lambda_j, j \in \mathcal{J}\} \subset \mathbb{R}$ denotes the (possibly empty) countable point spectrum of T. The sequence $\lambda = (\lambda_j)_{j\in\mathcal{J}}$ contains each eigenvalue of T repeated according to its multiplicity.*

(i) *If T is nuclear, then λ is absolute summable, i.e. $\|\lambda\|_{\ell_1} < \infty$, and $\operatorname{tr}(T) = \sum_{j\in\mathcal{J}} \lambda_j$.*

(ii) *If T is Hilbert-Schmidt, then λ is square summable and $\|T\|_{\mathscr{H}} = \|\lambda\|_{\ell^2} < \infty$.* □

§2.2.34 **Definition** *(Class of operators with given eigenfunctions)*. Given an ONS $\{u_j, j \in \mathcal{J}\}$ in $\mathbb{H}$ let $\mathcal{E}_u(\mathbb{H})$ or $\mathcal{E}_u$ for short be the subset of $\mathscr{K}(\mathbb{H})$ containing all compact, normal (self-adjoint), linear operators having for some $\mathcal{J}' \subseteq \mathcal{J}$, $\{u_j, j \in \mathcal{J}'\}$ as eigenfunctions, i.e., for each $T \in \mathcal{E}_u(\mathbb{H})$ there exist $\mathcal{J}' \subseteq \mathcal{J}$ and a sequence $(\lambda_j)_{j \in \mathcal{J}'}$ in $\mathbb{K}\setminus\{0\}$ such that T admits $\{(\lambda_j, u_j), j \in \mathcal{J}'\}$ as eigensystem, i.e., $\mathcal{E}_u(\mathbb{H}) \subset \{\nabla_\lambda, \lambda \in \mathbb{K}^{\mathcal{J}}\}$. □

§2.2.35 **Example**. Let $C_g \in \mathscr{K}(L^2([0,1)))$ be a *circular convolution operator*. Consider as in §2.1.6 (ii) the *exponential basis* $\{e_j\}_{j\in\mathbb{Z}}$ in $L^2([0,1))$ and for $f \in L^2([0,1))$ the associated Fourier coefficients $[f]_j = \langle f, e_j\rangle_{L^2}$, $j \in \mathbb{Z}$. Keep in mind that C_g is normal and for all $f \in L^2([0,1))$ the convolution theorem states $[g \circledast f]_j = [g]_j[f]_j$ for all $j \in \mathbb{Z}$. Thereby, $\{([g]_j, e_j),\ j \in \mathbb{Z}\}$ is an eigensystem of the circular convolution operator C_g. In other words, for each $g \in L([0,1))$ we have $C_g \in \mathcal{E}_e(L^2([0,1)))$. □

§2.2.36 **Property**. *Let $T \in \mathscr{K}(\mathbb{H})$ be strictly positive definite and let $(\lambda_j)_{j\in\mathbb{N}}$ be a strictly positive, monotonically non-increasing sequence containing each eigenvalue of T repeated according to its multiplicity. For $m \in \mathbb{N}$ let $\mathcal{H}_m$ be the set of all m-dimensional subspaces $\mathbb{U}_m$ in $\mathbb{H}$, and denote by $\mathbb{U}_m^\perp$ the orthogonal complement of $\mathbb{U}_m$ in $\mathbb{H}$. Furthermore, let $\mathbb{B}_{\mathbb{U}_m} := \{h \in \mathbb{U}_m : \|h\|_{\mathbb{H}} = 1\}$ and $\mathbb{B}_{\mathbb{U}_m^\perp}$ be the unit ball in $\mathbb{U}_m$ and $\mathbb{U}_m^\perp$, respectively.*

(Courant's max-min-principle) $\lambda_m = \max_{\mathbb{U}_m \in \mathcal{H}_m} \min_{h \in \mathbb{B}_{\mathbb{U}_m}} \langle Th, h\rangle_{\mathbb{H}}$,

(Courant's min-max-principle) $\lambda_m = \min_{\mathbb{U}_{m-1} \in \mathcal{H}_{m-1}} \max_{h \in \mathbb{B}_{\mathbb{U}_{m-1}^\perp}} \langle Th, h\rangle_{\mathbb{H}}$. □

§2.2.37 **Definition**. Let $T \in \mathscr{K}(\mathbb{H}, \mathbb{G})$. There exist

(i) a sequence $\mathfrak{s} := (\mathfrak{s}_j)_{j\in\mathcal{J}}$ in $\mathbb{K}\setminus\{0\}$ with $\|T\|_{\mathscr{L}} = \sup_{j\in\mathcal{J}} |\mathfrak{s}_j|$ which has either a finite number of entries or zero as only accumulation point, and determines a multiplication operator $\mathrm{M}_{\mathfrak{s}} \in \mathscr{L}(\ell^2(\mathcal{J}))$,

(ii) an (possibly finite) ONS $\{u_j, j \in \mathcal{J}\}$ in $\mathbb{H}$ with $\mathbb{U} := \overline{lin}\{u_j, j \in \mathcal{J}\}$ and associated generalised Fourier series transform $\mathcal{U} \in \mathscr{L}(\mathbb{H}, \ell^2(\mathcal{J}))$ (a partial isometry),

(iii) an (possibly finite) ONS $\{v_j, j \in \mathcal{J}\}$ in $\mathbb{G}$ with $\mathbb{V} := \overline{lin}\{v_j, j \in \mathcal{J}\}$ and associated generalised Fourier series transform $\mathcal{V} \in \mathscr{L}(\mathbb{G}, \ell^2(\mathcal{J}))$ (a partial isometry),

such that $\mathbb{H} = \mathcal{N}(T) \oplus \mathbb{U}$, $\mathbb{G} = \mathcal{N}(T^\star) \oplus \mathbb{V}$ and $T = \mathcal{V}^\star \mathrm{M}_{\mathfrak{s}} \mathcal{U} = \sum_{j\in\mathcal{J}} \mathfrak{s}_j\, u_j \otimes v_j$. In particular, $\{(|\mathfrak{s}_j|^2, u_j), j \in \mathcal{J}\}$ and $\{(|\mathfrak{s}_j|^2, v_j), j \in \mathcal{J}\}$ are an eigensystem of $T^\star T$ and $TT^\star$ respectively. The numbers $\{\mathfrak{s}_j, j \in \mathcal{J}\}$ and triplets $\{(\mathfrak{s}_j, u_j, v_j), j \in \mathcal{J}\}$ are, respectively, called *singular values* and *singular system* of T. □

§2.2.38 **Properties**. *Let $T \in \mathscr{K}(\mathbb{H}, \mathbb{G})$ with singular system $\{(\mathfrak{s}_j, u_j, v_j), j \in \mathcal{J}\}$ where the (possibly empty) countable point spectrum of $T^\star T$ (respectively, $TT^\star$) is given by $\sigma(T^\star T)\setminus\{0\} = \{|\mathfrak{s}_j|^2, j \in \mathcal{J}\} \subset \mathbb{R}$. The sequence $(|\mathfrak{s}_j|^2)_{j\in\mathcal{J}}$ contains each eigenvalue of $T^\star T$ repeated according to its multiplicity.*

(i) *If T is nuclear, then $\mathfrak{s}$ is absolute summable, i.e. $\|\mathfrak{s}\|_{\ell^1} < \infty$.*

(ii) *If T is Hilbert-Schmidt, then $\mathfrak{s}$ is square summable and $\|T\|_{\mathscr{H}} = \|\mathfrak{s}\|_{\ell^2} < \infty$.* □

§2.2.39 **Definition** *(Class of operators with known eigenfunctions)*. Given an ONS $\{u_j, j \in \mathcal{J}\}$ and $\{v_j, j \in \mathcal{J}\}$ in $\mathbb{H}$ and $\mathbb{G}$, respectively, let $\mathcal{S}_{u,v}(\mathbb{H}, \mathbb{G})$ or $\mathcal{S}_{u,v}$ for short, be the subset of $\mathscr{K}(\mathbb{H}, \mathbb{G})$ containing all compact, linear operators having for some $\mathcal{J}' \subseteq \mathcal{J}$, $\{u_j, j \in \mathcal{J}'\}$ and $\{u_j, j \in \mathcal{J}'\}$ as eigenfunctions, i.e., for each $T \in \mathcal{S}_{u,v}(\mathbb{H}, \mathbb{G})$ there exist $\mathcal{J}' \subseteq \mathcal{J}$ and a sequence $(\mathfrak{s}_j)_{j\in\mathcal{J}'}$ in $\mathbb{K}\setminus\{0\}$ such that T admits $\{(\mathfrak{s}_j, u_j, v_j), j \in \mathcal{J}'\}$ as singular system. □

§2.2.40 **Property** *(Spectral theorem)*. *If $T \in \mathscr{L}(\mathbb{H})$ is self-adjoint, then T is isometrically equivalent to a multiplication operator, i.e., there exist*

(i) *a measurable space (Ω, μ) (σ-finite, if $\mathbb{H}$ is separable),*

(ii) *a bounded (measurable) and μ-a.s. non zero function $\lambda : \Omega \to \mathbb{R}$ with associated multiplication operator $M_\lambda \in \mathscr{L}(L^2_\mu(\Omega))$, and*

(iii) *a partial isometry $\mathcal{U} \in \mathscr{L}(\mathbb{H}, L^2_\mu(\Omega))$,*

such that $T = \mathcal{U}^\star M_\lambda \mathcal{U}$. □

§2.2.41 **Example**. Let $g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ be a real and even function. Consider the associated self-adjoint *convolution operator* $C_g \in \mathscr{L}(L^2(\mathbb{R}))$. Recall that the convolution theorem states $\mathcal{F}(g * f) = \mathcal{F}g \cdot \mathcal{F}f$ for all $f \in L^2(\mathbb{R})$ where $\mathcal{F}$ denotes the *Fourier-Plancherel transform*. Consequently, the operator C_g is unitarily equivalent to the multiplication operator $\mathrm{M}_\lambda \in \mathscr{L}(L^2(\mathbb{R}))$ with $\lambda = [\mathcal{F}g]$, that is $C_g = \mathcal{F}^{-1}\mathrm{M}_\lambda\mathcal{F}$. □

§2.2.42 **Property** *(Spectral theorem Halmos [1963])*. *Let $T : \mathbb{H} \supset \mathcal{D}(T) \to \mathbb{H}$ be a densely-defined self-adjoint operator. There exist*

(i) *a measurable space (Ω, μ) (σ-finite, if $\mathbb{H}$ is separable),*

(ii) *an unitary operator $\mathcal{U} \in \mathscr{L}(\mathbb{H}, L^2_\mu(\Omega))$,*

(iii) *a (measurable) function $\lambda : \Omega \to \mathbb{R}$ (μ-a.s. finite and non zero) and an associated multiplication operator $\mathrm{M}_\lambda : L^2_\mu(\Omega) \supset \mathcal{D}(\mathrm{M}_\lambda) \to L^2_\mu(\Omega)$ with $\mathcal{D}(\mathrm{M}_\lambda) = \big\{f \in L^2_\mu(\Omega) : \lambda f \in L^2_\mu(\Omega)\big\}$*

such that $\mathcal{D}(T) = \{h \in \mathbb{H} : \mathcal{U}h \in \mathcal{D}(M_\lambda)\}$ and

(a) *for all $f \in \mathcal{D}(\mathrm{M}_\lambda)$ we have $\mathrm{M}_\lambda f = \lambda \cdot f = \mathcal{U}T\mathcal{U}^\star f$,*

(b) *for all $h \in \mathcal{D}(T)$ it holds $Th = \mathcal{U}^\star \mathrm{M}_\lambda \mathcal{U}h$,*

i.e., T is unitarily equivalent to the multiplication operator M_λ. □

§2.2.43 **Example**. Let $T \in \mathscr{K}(\mathbb{H})$ be an injective and self-adjoint operator with eigenvalue decomposition $T = \mathcal{U}^\star \mathrm{M}_\lambda \mathcal{U}$ where $\mathcal{U} \in \mathscr{L}(\mathbb{H}, \ell^2)$ is unitary, $\mathrm{M}_\lambda \in \mathscr{L}(\ell^2)$ is a multiplication operator and λ a sequence in $\mathbb{R}\backslash\{0\}$ of eigenvalues repeated according to their multiplicities. If $\mathbb{H}$ is not finite dimensional then the range $\mathcal{R}(T)$ of T is dense in $\mathbb{H}$ but not closed. Therefore, there exists an inverse $T^{-1} : \mathcal{R}(T) \to \mathbb{H}$ of T which is densely-defined and self-adjoint but not continuous. In particular, we have $\mathcal{D}(T^{-1}) = \mathcal{R}(T) = \{h : \lambda^{-1}\mathcal{U}h \in \ell^2\}$ (which is called Picard's condition). Consider the multiplication operator $\mathrm{M}_{1/\lambda} : \ell^2 \supset \mathcal{D}(\mathrm{M}_{1/\lambda}) \to \ell$ with $\mathcal{D}(\mathrm{M}_{1/\lambda}) = \{x \in \ell : x/\lambda \in \ell^2\}$, then $\mathcal{D}(T^{-1}) = \{h \in \mathbb{H} : \mathcal{U}h \in \mathcal{D}(\mathrm{M}_{1/\lambda})\}$ and

(a) for all $x \in \mathcal{D}(\mathrm{M}_{1/\lambda})$ we have $\mathrm{M}_{1/\lambda}x = x/\lambda = \mathcal{U}T^{-1}\mathcal{U}^\star x$,

(b) for all $h \in \mathcal{D}(T^{-1})$ it holds $T^{-1}h = \mathcal{U}^\star \mathrm{M}_{1/\lambda}\mathcal{U}h$,

i.e. T^{-1} is unitarily equivalent to the multiplication operator $\mathrm{M}_{1/\lambda}$. We shall emphasise that $h \in \mathcal{D}(T^{-1}) = \mathcal{R}(T)$ if and only if $\|[h]/\lambda\|^2_{\ell^2} = \sum_{j\in\mathcal{J}} |[h]_j/\lambda_j|^2 < \infty$. On the other hand, for any $k \in \mathbb{N}$ we have $T^k = T\cdots T = \mathcal{U}^\star \mathrm{M}_{\lambda^k}\mathcal{U} = \sum_{j\in\mathcal{J}} \lambda_j^k\, u_j \otimes u_j$ which motivates for a function $g : \mathbb{R} \to \mathbb{R}$ to define the operator

$$g(T)h := \mathcal{U}^\star \mathrm{M}_{g(\lambda)}\mathcal{U}h = \sum_{j\in\mathcal{J}} g(\lambda_j)\, u_j \otimes u_j, \quad \text{for all } h \in \mathbb{H} \text{ with } \|g(\lambda)[h]\|_{\ell^2} < \infty.$$

If g is bounded then $g(T) \in \mathscr{L}(\mathbb{H})$ and $\|g(T)\|_{\mathscr{L}} = \sup\{|g(\lambda_j)|, j \in \mathcal{J}\} \leqslant \|g\|_{L^\infty}$. In particular, it allows to define T^s for all $s \in \mathbb{R}$. □

§2.2.44 **Definition** *(Functional calculus)*. Let $T \in \mathscr{L}(\mathbb{H})$ be self-adjoint and hence isometrically equivalent with multiplication by a bounded function λ in some $L^2_\mu(\Omega)$, that is, $T = U^\star \mathrm{M}_\lambda U$. Given a (measurable) function $g : \mathbb{R} \to \mathbb{R}$ define the multiplication operator

$$\mathrm{M}_{g(\lambda)} : L^2_\mu(\Omega) \supset \mathcal{D}(\mathrm{M}_{g(\lambda)}) \to L^2_\mu(\Omega)$$

with $\mathcal{D}(\mathrm{M}_{g(\lambda)}) = \big\{f \in L^2_\mu(\Omega) : g(\lambda)f \in L^2_\mu(\Omega)\big\}$ and an unitarily equivalent operator

$$g(T)h := \mathcal{U}^\star \mathrm{M}_{g(\lambda)} \mathcal{U} h, \quad \forall h \in \mathcal{D}(g(T)) := \{h \in \mathbb{H} : \mathcal{U}h \in \mathcal{D}(\mathrm{M}_{g(\lambda)})\}$$

where $g(T) : \mathscr{L}(\mathbb{H}) \supset \mathcal{D}(g(T)) \to \mathscr{L}(\mathbb{H})$. Moreover, if g is bounded then $g(T) \in \mathscr{L}(\mathbb{H})$ with $\|g(T)\|_{\mathscr{L}} = \sup\{|g(\lambda)|, \lambda \in \sigma(T)\} \leqslant \|g\|_{L^\infty}$. □

§2.2.45 **Property**. *Let $T \in \mathscr{L}(\mathbb{H}, \mathbb{G})$. Then $\mathcal{R}(T) = \mathcal{R}((T^\star T)^{1/2})$.*

§2.2.46 **Remark**. Considering an ONB $\{u_j, j \in \mathbb{N}\}$ in $\mathbb{H}$, the associated generalised Fourier series transform $\mathcal{U} \in \mathscr{L}(\mathbb{H}, \ell^2)$ and for a sequence $\mathfrak{v}$ the associated multiplication and diagonal operator $\mathrm{M}_\mathfrak{v} : \ell^2 \supset \mathcal{D}(\mathrm{M}_\mathfrak{v}) \to \ell^2$ and $\nabla_\mathfrak{v} = \mathcal{U}^\star \mathrm{M}_\mathfrak{v}\, \mathcal{U} : \mathbb{H} \supset \mathcal{D}(\nabla_\mathfrak{v}) \to \mathbb{H}$ defined as in §2.2.4 (iv) and (v), respectively. If $\mathfrak{v}$ is strictly positive then applying the functional calculus we observe that for any $s \in \mathbb{R}$ we have $\nabla_\mathfrak{v}^s = \mathcal{U}^\star \mathrm{M}_{\mathfrak{v}^s}\, \mathcal{U} = \nabla_{\mathfrak{v}^s}$. Moreover, recall that $\mathbb{H}_{\mathfrak{v}^s}$ denotes the completion of $\mathbb{H}$ w.r.t.. the weighted norm $\|\cdot\|_{\mathfrak{v}^s}$ given by $\|\cdot\|^2_{\mathfrak{v}^s} = \sum_{j\in\mathbb{N}} \mathfrak{v}_j^{2s} |\langle \cdot, u_j\rangle_\mathbb{H}|^2$ where obviously $\|h\|_{\mathfrak{v}^s} = \|\nabla_{\mathfrak{v}^s} h\|_\mathbb{H} = \|\nabla^s_\mathfrak{v} h\|_\mathbb{H}$ for all $h \in \mathcal{D}(\nabla_{\mathfrak{v}^s}) = \mathbb{H}_{\mathfrak{v}^s}$. Introduce further the Hilbert space $(\mathbb{H}_{\mathfrak{v}^s}, \langle\cdot,\cdot\rangle_{\mathfrak{v}^s})$ inner product $\langle\cdot,\cdot\rangle_{\mathfrak{v}^s} = \langle \nabla_{\mathfrak{v}^s}\cdot, \nabla_{\mathfrak{v}^s}\cdot\rangle_\mathbb{H}$. □

§2.2.47 **Definition**. For a monotonically increasing, unbounded sequence $\mathfrak{v}$ with $\mathfrak{v}_1 > 0$ and any $s \in \mathbb{R}$ consider the inner product $\langle\cdot,\cdot\rangle_{\mathfrak{v}^s} = \langle \nabla_{\mathfrak{v}^s}\cdot, \nabla_{\mathfrak{v}^s}\cdot\rangle_\mathbb{H}$ and the norm $\|\cdot\|_{\mathfrak{v}^s} = \|\nabla^s_\mathfrak{v}\cdot\|_\mathbb{H}$. The family $\{(\mathbb{H}_{\mathfrak{v}^s}, \langle\cdot,\cdot\rangle_{\mathfrak{v}^s}), s \in \mathbb{R}\}$ of Hilbert space is called a *Hilbert scale* (see Krein and Petunin [1966] for a rather complete theory). □

§2.2.48 **Properties**. *Let $\{(\mathbb{H}_{\mathfrak{v}^s}, \langle\cdot,\cdot\rangle_{\mathfrak{v}^s}), s \in \mathbb{R}\}$ be a Hilbert scale as introduced in Definition §2.2.47. Then the following assertions hold true:*

(i) *For any $-\infty < s < t < \infty <$ the space $\mathbb{H}_{\mathfrak{v}^t}$ is densely and continuously embedded in $\mathbb{H}_{\mathfrak{v}^s}$.*

(ii) *For $s, t \in \mathbb{R}$ holds $\nabla_\mathfrak{v}^{t-s} = \nabla_\mathfrak{v}^t \nabla_\mathfrak{v}^{-s}$, and in particular, $\nabla_{\mathfrak{v}^s}^{-1} = \nabla_{\mathfrak{v}^{-s}}$.*

(iii) *For $s \geqslant 0$ holds $\mathbb{H}_{\mathfrak{v}^s} = \mathcal{D}(\nabla_{\mathfrak{v}^s})$ and $\mathbb{H}_{\mathfrak{v}^{-s}}$ is the dual space of $\mathbb{H}_{\mathfrak{v}^s}$.*

(iv) *Considering $-\infty < r < s < t < \infty$ for any $h \in \mathbb{H}_{\mathfrak{v}^s}$ the interpolation inequality $\|h\|_{\mathfrak{v}^s} \leqslant \|h\|_{\mathfrak{v}^r}^{(t-s)/(t-r)} \|h\|_{\mathfrak{v}^t}^{(s-r)/(t-r)}$ holds true.* □

§2.2.49 **Example**. Let $T \in \mathscr{K}(\mathbb{H}, \mathbb{G})$ be injective with singular system $\{(\mathfrak{s}_j, u_j, v_j), j \in \mathbb{N}\}$ for some ONB $\{u_j \in \mathbb{N}\}$ in $\mathbb{H}$ and strictly positive, monotonically non-increasing sequence $(\mathfrak{s}_j)_{j\in\mathbb{N}}$ containing each singular value of T repeated according to its multiplicity. Setting $\mathfrak{v} = \mathfrak{s}^{-2}$ the strictly positive definite operator $T^\star T$ admits the spectral representation $T^\star T = \mathcal{U}^\star \mathrm{M}_{\mathfrak{v}^{-1}} \mathcal{U} = \nabla_{\mathfrak{v}^{-1}}$. Obviously, $\mathfrak{v}$ is a monotonically increasing, unbounded sequence with $\mathfrak{v}_1 > 0$. Considering the associated Hilbert scale $\{(\mathbb{H}_{\mathfrak{v}^s}, \langle\cdot,\cdot\rangle_{\mathfrak{v}^s}), s \in \mathbb{R}\}$ it is then an immediate consequence that $\mathbb{H}_{\mathfrak{v}^t} = \mathcal{D}((T^\star T)^t)$ is dense in $\mathbb{H}_{\mathfrak{v}^s} = \mathcal{D}((T^\star T)^s)$ for $0 \leqslant s < t$. We say, a function f satisfies a *source condition*, if $f \in \mathcal{D}((T^\star T)^s)$ for some $s > 0$, i.e., $f = (T^\star T)^s h$ for some $h \in \mathbb{H}$. □

2.2.3 Abstract smoothing condition

§2.2.50 **Definition** *(Link condition)*. Denote by $\mathcal{T}(\mathbb{H})$ or $\mathcal{T}$ for short, the set of all strictly positive operator in $\mathscr{K}(\mathbb{H})$. Given an ONB $\{u_j, j \in \mathcal{J}\}$ in $\mathbb{H}$ and a strictly positive sequence $(\mathfrak{t}_j)_{j\in\mathcal{J}}$ consider the weighted norm $\|\cdot\|_{\mathfrak{t}}^2 = \sum_{j\in\mathcal{J}} \mathfrak{t}_j^2 |\langle\cdot, u_j\rangle_{\mathbb{H}}|^2$. For all $d \geqslant 1$ define the subset $\mathcal{T}_{u,\mathfrak{t}}^d := \mathcal{T}_{u,\mathfrak{t}}^d(\mathbb{H}) := \{T \in \mathcal{T} : d^{-1}\|h\|_{\mathfrak{t}} \leqslant \|Th\|_{\mathbb{H}} \leqslant d\|h\|_{\mathfrak{t}} \text{ for all } h \in \mathbb{H}\}$. We say, T satisfies the *link condition* $\mathcal{T}_{u,\mathfrak{t}}^d$, if $T \in \mathcal{T}_{u,\mathfrak{t}}^d$. Define further subsets of $\mathcal{E}_u = \mathcal{E}_u(\mathbb{H})$ and $\mathcal{S}_{u,v} = \mathcal{S}_{u,v}(\mathbb{H},\mathbb{G})$ containing, respectively, any operator T such that $(T^\star T)^{1/2}$ satisfies the link condition $\mathcal{T}_{u,\mathfrak{t}}^d$, that is, $\mathcal{E}_{u,\mathfrak{t}}^d = \{T \in \mathcal{E}_u : (T^\star T)^{1/2} \in \mathcal{T}_{u,\mathfrak{t}}^d\}$ and $\mathcal{S}_{u,v,\mathfrak{t}}^d = \{T \in \mathcal{S}_{u,v} : (T^\star T)^{1/2} \in \mathcal{T}_{u,\mathfrak{t}}^d\}$, respectively. □

§2.2.51 **Remark**. We shall emphasise that for $T \in \mathscr{K}(\mathbb{H},\mathbb{G})$ the condition $(T^\star T)^{1/2} \in \mathcal{T}_{u,\mathfrak{t}}^d$ is equivalent to $d^{-1}\|h\|_{\mathfrak{t}} \leqslant \|Th\|_{\mathbb{H}} \leqslant d\|h\|_{\mathfrak{t}}$ for all $h \in \mathbb{H}$. Observe further that $T \in \mathcal{S}_{u,v}$ admitting a singular system $\{(\mathfrak{s}_j, u_j, v_j), j \in \mathcal{J}'\}$ with $\mathcal{J}' \subseteq \mathcal{J}$ satisfies the link condition $\mathcal{S}_{u,v,\mathfrak{t}}^d$ if and only if $\mathcal{J}' = \mathcal{J}$ and $d^{-1} \leqslant |\mathfrak{s}_j|/\mathfrak{t}_j \leqslant d$ for all $j \in \mathcal{J}$. Thereby, we have that $T \in \mathcal{S}_{u,v,\mathfrak{t}}^d(\mathbb{H},\mathbb{G})$ if and only if $T^\star \in \mathcal{S}_{v,u,\mathfrak{t}}^d(\mathbb{G},\mathbb{H})$. We shall emphasise, that there are operators satisfying the link condition $\mathcal{T}_{u,\mathfrak{t}}^d$ which do not belong to $\mathcal{E}_u$ (respectively, $\mathcal{S}_{u,v}$), i.e., are not equal to ∇_λ for some sequence λ (not diagonal w.r.t. $\mathcal{U}$), that is admitting eigenfunctions which are not contained in the ONS $\{u_j, j \in \mathcal{J}\}$. Let us briefly give a construction of those. We consider a small perturbation of $\nabla_{\mathfrak{t}}$, that is, $T = \nabla_{\mathfrak{t}} + \nabla_{\mathfrak{t}} A \nabla_{\mathfrak{t}}$ where $A \in \mathscr{L}(\mathbb{H})$ is a non-negative definite operator with spectral norm $c := \|\nabla_{\mathfrak{t}} A\|_{\mathscr{L}}$ strictly smaller than one. Obviously, T is strictly positive definite, and $\|Th\|_{\mathbb{H}} \leqslant \|\mathrm{Id}_{\mathbb{H}} + \nabla_{\mathfrak{t}} A\|_{\mathscr{L}} \|\nabla_{\mathfrak{t}} h\|_{\mathbb{H}} \leqslant (1+c)\|h\|_{\mathfrak{t}}$. On the other hand, we have $\|(\mathrm{Id}_{\mathbb{H}} + \nabla_{\mathfrak{t}} A)^{-1}\|_{\mathscr{L}} = \frac{1}{1-\|\nabla_{\mathfrak{t}} A\|_{\mathscr{L}}} = \frac{1}{1-c}$ by the Neumann series argument §2.2.15, which in turn implies $\|h\|_{\mathfrak{t}} = \|\nabla_{\mathfrak{t}} h\|_{\mathbb{H}} = \|(\mathrm{Id}_{\mathbb{H}} + \nabla_{\mathfrak{t}} A)^{-1}\|_{\mathscr{L}} \|Th\|_{\mathbb{H}} \leqslant \frac{1}{1-c}\|Th\|_{\mathbb{H}}$. Combining both bounds the operator T satisfies the link condition $\mathcal{S}_{u,v,\mathfrak{t}}^d$ for all $d \geqslant \max(1+c, \frac{1}{1-c})$ and is obviously not diagonal w.r.t. $\mathcal{U}$. □

§2.2.52 **Property**. *Let $T \in \mathcal{T}_{u,\mathfrak{t}}^d$.*

(Inequality of Heinz [1951]) For all $|s| \leqslant 1$ holds $\frac{1}{d^{|s|}}\|h\|_{\mathfrak{t}^s} \leqslant \|T^s h\|_{\mathbb{H}} \leqslant d^{|s|}\|h\|_{\mathfrak{t}^s}$. □

§2.2.53 **Example** *(Example §2.2.49 continued)*. Consider the Hilbert scale $\{(\mathbb{H}_{\mathfrak{v}^s}, \langle\cdot,\cdot\rangle_{\mathfrak{v}^s}), s \in \mathbb{R}\}$ associated with the source condition, i.e., $\mathbb{H}_{\mathfrak{v}^s} = \mathcal{D}((T^\star T)^s)$ and $\|\cdot\|_{\mathfrak{v}^s} = \|(T^\star T)^{-s}\cdot\|_{\mathbb{H}}$ for $s > 0$. Suppose further that $(T^\star T)^{1/2} \in \mathcal{T}_{u,\mathfrak{t}}^d$, i.e., T satisfies a link condition for some weighted norm $\|\cdot\|_{\mathfrak{t}}$ defined w.r.t. a certain ONB in $\mathbb{H}$ and a strictly positive sequence $\mathfrak{t}$. Note that in general the two norms $\|\cdot\|_{\mathfrak{t}}$ and $\|\cdot\|_{\mathfrak{v}^s}$ are defined w.r.t. to different orthonormal basis in $\mathbb{H}$. However, rewriting the inequality of Heinz §2.2.52 accordingly it holds $\frac{1}{d^{|s|}}\|\cdot\|_{\mathfrak{t}^s} \leqslant \|(T^\star T)^{s/2}\cdot\|_{\mathbb{H}} \leqslant d^{|s|}\|\cdot\|_{\mathfrak{t}^s}$ or equivalently $\frac{1}{d^{|s|}}\|\cdot\|_{\mathfrak{t}^s} \leqslant \|\cdot\|_{\mathfrak{v}^{-s/2}} \leqslant d^{|s|}\|\cdot\|_{\mathfrak{t}^s}$. In other words the two norms $\|\cdot\|_{\mathfrak{t}^s}$ and $\|\cdot\|_{\mathfrak{v}^{-s/2}}$ are equivalent for any $|s| \leqslant 1$. Recall that $\mathfrak{v}^{-1/2} = \mathfrak{s}$ equals the sequence of singular values of T. We shall emphasise that the equivalence of $\|\cdot\|_{\mathfrak{t}^s}$ and $\|\cdot\|_{\mathfrak{v}^{-s/2}}$ under a link condition holds generally for all $|s| \leqslant 1$ only. However, if the ONB used to construct the norm $\|\cdot\|_{\mathfrak{t}^s}$ for the link condition coincides with the eigenfunctions of $T^\star T$ then the $\|\cdot\|_{\mathfrak{t}^s}$ and $\|\cdot\|_{\mathfrak{v}^{-s/2}}$ are equivalent for all $s \in \mathbb{R}$. □

§2.2.54 **Corollary**. *Let $T \in \mathcal{T}_{u,\mathfrak{t}}^d$ and suppose that $f \in \mathbb{F}_{\mathfrak{a}}^r$ (see Definition §2.1.18) where the two norms $\|\cdot\|_{\mathfrak{t}}$ and $\|\cdot\|_{\mathfrak{a}^{-1}}$ are construct w.r.t. the same ONB in $\mathbb{H}$. Assume in addition that there are constants $a, p > 0$ and a sequence $\mathfrak{v}$ such that $\mathfrak{t} = \mathfrak{v}^a$ and $\mathfrak{a} = \mathfrak{v}^p$. If $p \leqslant 2a$ then for any $f \in \mathbb{F}_{\mathfrak{a}}$ holds $f = (T^\star T)^{p/(2a)} h$ with $\|h\|_{\mathbb{H}} \leqslant d^{p/a}\|h\|_{1/\mathfrak{a}}$, and conversely for any*

$f = (T^{\star}T)^{p/(2a)}h$ *with* $\|h\|_{\mathbb{H}} < \infty$ *we have* $f \in \mathbb{F}_{\mathfrak{a}}$ *with* $\|h\|_{1/\mathfrak{a}} \leqslant d^{p/a}\|h\|_{\mathbb{H}}$.

Proof of Corollary §2.2.54 is given in the lecture. □

§2.2.55 **Lemma**. *Given an ONB* $\{u_j, j \in \mathbb{N}\}$ *in* $\mathbb{H}$ *and a strictly positive non-increasing sequence* $(\mathfrak{t}_j)_{j\in\mathbb{N}}$ *consider the link condition* $\mathcal{T}^d_{u,\mathfrak{t}}$. *Let* $T \in \mathcal{T}(\mathbb{H})$ *admit* $\{(\lambda_j, \psi_j), j \in \mathbb{N}\}$ *as eigensystem where the strictly positive, monotonically non-increasing sequence* $(\lambda_j)_{j\in\mathbb{N}}$ *contains each eigenvalue of* T *repeated according to its multiplicity and the associated eigenbasis* $\{\psi_j, j \in \mathbb{N}\}$ *does eventually not correspond to the ONB* $\{u_j, j \in \mathbb{N}\}$. *If* $T \in \mathcal{T}^d_{u,\mathfrak{t}}$, *then we have* $d^{-1} \leqslant \lambda_j/\mathfrak{t}_j \leqslant d$ *for all* $j \in \mathbb{N}$.

Proof of Lemma §2.2.55 is given in the lecture. □

Bibliography

L. Birgé and P. Massart. From model selection to adaptive estimation. Pollard, David (ed.) et al., Festschrift for Lucien Le Cam: research papers in probability and statistics. New York, NY: Springer. 55-87, 1997.

N. Dunford and J. T. Schwartz. *Linear Operators, Part I: General Theory*. Wiley Classics Library. John Wiley & Sons Ltd, New York, 1988a.

N. Dunford and J. T. Schwartz. *Linear operators. Part II: Spectral theory, self adjoint operators in Hilbert space.* Wiley Classics Library. John Wiley & Sons Ltd, New York, 1988b.

N. Dunford and J. T. Schwartz. *Linear operators. Part III, Spectral Operators.* Wiley Classics Library. John Wiley & Sons Ltd, New York, 1988c.

P. R. Halmos. What does the spectral theorem say? *Amer. Math. Monthly*, 70:241–247, 1963.

E. Heinz. Beiträge zur Störungstheorie der Spektralzerlegung. *Mathematische Annalen*, 123: 415–438, 1951.

T. Kawata. *Fourier analysis in probability theory.* Academic Press, New York, 1972.

S. Krein and Y. I. Petunin. Scales of banach spaces. In *Russian Math. Surveys*, volume 21, pages 85–169, 1966.

A. B. Tsybakov. *Introduction to nonparametric estimation.* Springer Series in Statistics. Springer, New York, 2009.

D. Werner. *Funktionalanalysis*. Springer-Lehrbuch, 2011.

Index

Nested sieve
- $(\mathbb{U}_m)_{m\in\mathcal{M}}$ in $\mathbb{U}$, 7
- $(\mathcal{J}_m)_{m\in\mathcal{M}}$ in $\mathcal{J}$, 7

Norm
- spectral, $\|\cdot\|_s$, 9
- uniform operator, $\|\cdot\|_{\mathscr{L}}$, 8

Operator
- conditional expectation, $K_{X|Z}$, 9, 11, 12, 14
- convolution circular, C_{f_U}, 10
- convolution circular, C_{f_U}, 15
- convolution circular, C_{f_U}, 12, 13
- convolution, C_{f_U}, 9, 11–13, 16
- diagonal, ∇_λ, 9, 12, 13
- Fourier series transform, U, 9, 11, 12
- Fourier-Plancherel transform, $\mathcal{F}$, 12
- linear functional, Φ, 10
- multiplication, M_λ, 14
- multiplication, M_λ, 9, 11–13

Operator classes
- bounded linear, $\mathscr{L}(\mathbb{H},\mathbb{G})$, 8
- compact, $\mathscr{K}(\mathbb{H},\mathbb{G})$, 12, 16
- Hilbert-Schmidt, $\mathscr{H}(\mathbb{H},\mathbb{G})$, 13
- known eigenfunctions, $\mathcal{E}_u(\mathbb{H})$, $\mathcal{S}_{u,v}(\mathbb{H},\mathbb{G})$, 14, 15
- linear functionals, $\mathcal{L}_{1/\mathfrak{a}}$, 10
- link condition, $\mathcal{T}^d_{u,\mathfrak{t}}$, 18
- nuclear, $\mathscr{N}(\mathbb{H},\mathbb{G})$, 13

Orthonormal system (ONS), 6
- regular, 7, 8