## Lecture Notes

# Integral Equations and Boundary Value Problems

*Andreas Neubauer*

and

*Ronny Ramlau*

Industrial Mathematics Institute

Johannes Kepler University Linz

April 30, 2024

# Contents

# Preface

These lecture notes are based on the first five chapters of the book [4] by Heinz. W. Engl. We also recommend the book by R. Kress [6].

# 1. Classification and examples

In general, an equation is called integral equation, when the unknown function appears under an integral sign.

If, for instance, $g : [0,1]^2 \times \mathbb{R} \to \mathbb{R}$ and $f : [0,1] \to \mathbb{R}$ are given, then

$$x(s) - \int_0^1 g(s,t,x(t))\,dt = f(s)\,, \qquad s \in [0,1]\,,$$

is an integral equation for the unknown function $x$. This is a general nonlinear integral equation.

However, we will restrict ourselves to linear integral equations, where the following classification is usual:

If the unknown function only appears under the integral, we speak of an integral equation of the **first kind**, otherwise of an equation of the **second** or **third kind**, depending on whether the unknown function outside the integral is multiplied by a constant or a function.

If the integration limits are fixed, the equation is called **Fredholm** equation. If one limit of integration is a variable, the equation is called a **Volterra** equation.

For instance,

$$\int_0^1 k(s,t)x(t)\,dt = f(s)\,, \qquad s \in [0,1]\,, \tag{1.1}$$

is a (linear) Fredholm integral equation of the first kind,

$$\int_0^s k(s,t)x(t)\,dt = f(s)\,, \qquad s \in [0,1]\,, \tag{1.2}$$

is a (linear) Volterra integral equation of the first kind,

$$a(s)x(s) - \int_0^1 k(s,t)x(t)\,dt = f(s)\,, \qquad s \in [0,1]\,, \tag{1.3}$$

is a (linear) Fredholm integral equation of the third kind, and

$$x(s) - \lambda \int_0^s k(s,t)x(t)\,dt = f(s)\,, \qquad s \in [0,1]\,,$$

is a (linear) Volterra integral equation of the second kind for the unknown function $x$.

This distinction is not always stringent. If for instance $a(s) \neq 0$ in (1.3) for all $s \in [0,1]$, so by dividing through $a$ equation (1.3) kann be rewritten as an integral equation of

the second kind. The Volterra equation (1.2), on the other hand, can be reformulated as a Fredholm equation (1.1): with

$$\tilde{k}(s,t) := \begin{cases} k(s,t)\,, & t \le s\,, \\ 0\,, & t > s\,, \end{cases}$$

equation (1.2) is equivalent to

$$\int_0^1 \tilde{k}(s,t)x(t)\,dt = f(s)\,, \qquad s \in [0,1]\,.$$

Since Volterra equations have special properties, this distinction is still meaningful.

Equations of the first kind differ in theory and numerics completely from those of the second kind. Usually equations of the first kind are **improperly posed** and lead to so called **ill-posed problems**, where we have to expect problems when we want to solve them, since the solution no longer depends continuously on the data. Since these problems are treated in a special course, we only concentrate on equations of the second or third kind.

The function $k$ in the examples above is called **kernel** of the integral equation. If the kernel has a singularity, we speak of singular integral equations. Depending on the quality of the singularity we distinguish between weakly singular and strongly singular integral equations (see Definition 2.12)

$$x(s) - \int_{-1}^1 \frac{x(t)}{s-t}\,dt = f(s)\,, \qquad s \in [-1,1]\,,$$

is for instance strongly singular and

$$x(s) - \int_{-1}^1 \ln|s-t|\,x(t)\,dt = f(s)\,, \qquad s \in [-1,1]\,,$$

weakly singular.

The following examples illustrate, where integral equations usually occur:

**Example 1.1.** We consider the initial value problem

$$x''(t) = g(x(t))\,, \quad t \in [0,1]\,, \qquad x(0) = 1\,,\, x'(0) = 0\,, \tag{1.4}$$

with continuous $g$ given. By integration we obtain that

$$x'(t) = \int_0^t g(x(\tau))\,d\tau + c_1 \quad \text{and} \quad x(s) = \int_0^s \int_0^t g(x(\tau))\,d\tau dt + c_1 s + c_2\,.$$

The initial conditions yield: $c_1 = 0$, $c_2 = 1$. Thus, due to Fubini's theorem and noting that $D := \{(t,\tau) : t \in [0,s], \tau \in [0,t]\} = \{(t,\tau) : \tau \in [0,s], t \in [\tau,s]\}$ (see Figure 1.1), we get

$$\begin{aligned} x(s) &= 1 + \int_0^s \int_0^t g(x(\tau))\,d\tau dt = 1 + \int_D g(x(\tau))\,d(t,\tau) \\ &= 1 + \int_0^s \int_\tau^s g(x(\tau))\,dt d\tau = 1 + \int_0^s (s-\tau) g(x(\tau))\,d\tau\,. \end{aligned}$$

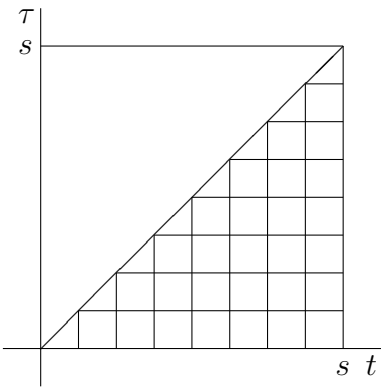


Figure 1.1: Integration area

Hence, $x$ is the solution of the following (nonlinear) Volterra integral equation of the second kind:

$$x(s) = 1 + \int_0^s (s-t)g(x(t))\,dt\,, \qquad s \in [0,1]\,. \tag{1.5}$$

It is rather obvious that initial value problems like the one in (1.4) yield Volterra integral equations, since the solution in the point $s$ may only depend on values $x(t)$ with $t \le s$. This is different for boundary value problems, where $x(s)$ can also depend on future values. This will then lead to Fredholm integral equations as the following example shows.

**Example 1.2.** We consider the boundary value problem

$$x''(t) = g(x(t))\,, \quad t \in [0,1]\,, \qquad x(0) = x(1) = 0\,,$$

with continuous $g$ given. It follows as in Example 1.1 that

$$x(s) = \int_0^s (s-t)g(x(t))\,dt + c_1 s + c_2\,, \qquad s \in [0,1]\,.$$

The boundary conditions now yield: $c_2 = 0$ and $c_1 = \int_0^1 (t-1)g(x(t))\,dt$. Therefore,

$$\begin{aligned} x(s) &= \int_0^s (s-t)g(x(t))\,dt + \int_0^1 s(t-1)g(x(t))\,dt \\ &= \int_0^s \underbrace{[(s-t)+s(t-1)]}_{=t(s-1)} g(x(t))\,dt + \int_s^1 s(t-1)g(x(t))\,dt\,. \end{aligned}$$

Thus, $x$ now satisfies the (nonlinear) Fredholm integral equation of the second kind:

$$x(s) = \int_0^1 k(s,t)g(x(t))\,dt \tag{1.6}$$

with kernel

$$k(s,t) = \begin{cases} t(s-1)\,, & t \le s\,, \\ s(t-1)\,, & t > s \end{cases}\,.$$

The kernel $k$ in (1.6) is continuous in $[0,1]^2$, the first partial derivatives of $k$ are continuous away from the diagonal $\{(t,t) : t \in [0,1]\}$ and have a jump on the diagonal. Such kernels typically occur when treating boundary value probelms for ordinary differential equations via integral equations. We will come back to such kernels in Chapter 5.

**Example 1.3.** Let $f : (0,\infty) \to \mathbb{R}$ be continuous. The **Laplace transform** of $f$ is defined as the function

$$(\mathcal{L}f)(s) := \int_0^\infty e^{-st} f(t)\, dt$$

for all values $s$, where the improper integral exists. We will not discuss existence and the domain of $\mathcal{L}f$ here.

It is an essential property of the Laplace transform that it turns differentiation to an algebraic operation:

$$(\mathcal{L}f')(s) = \int_0^\infty e^{-st} f'(t)\, dt = f(t)e^{-st}\Big|_{t=0}^{t=\infty} + \int_0^\infty s e^{-st} f(t)\, dt = -f(0) + s\cdot(\mathcal{L}f)(s)$$

Note that, due to the existence of $(\mathcal{L}f)$ it holds that $\lim_{t\to\infty} f(t)e^{-st} = 0$. Thus,

$$(\mathcal{L}f')(s) = s\cdot(\mathcal{L}f)(s) - f(0)\,. \tag{1.7}$$

In the same manner, one obtains expressions for the Laplace transform of higher order derivatives. The property of the Laplace transform to turn differentiation into a multiplication with $s$ can be used to transform (systems of) linear differential equations into algebraic equations. If, for instance, we want to solve the initial value probelm ($t \geq 0$)

$$\begin{aligned} f_1'(t) + f_2(t) &= e^t\,, \\ f_2'(t) - f_1(t) &= -e^t\,, \end{aligned} \qquad f_1(0) = 1\,, \quad f_2(0) = 1\,,$$

for the unknown functions $f_1$, $f_2$, using the Laplace transform, we proceed as follows: since $(\mathcal{L}\exp)(s) = (s-1)^{-1}$, for $s > 1$, (1.7) yields

$$\begin{aligned} s\cdot(\mathcal{L}f_1)(s) - 1 + (\mathcal{L}f_2)(s) &= \frac{1}{s-1}\,, \\ s\cdot(\mathcal{L}f_2)(s) - 1 - (\mathcal{L}f_1)(s) &= \frac{-1}{s-1}\,. \end{aligned}$$

This linear system for $\mathcal{L}f_1$ and $\mathcal{L}f_2$ has the solution

$$(\mathcal{L}f_1)(s) = \frac{1}{s-1} - \frac{1}{s^2+1} \quad \text{and} \quad (\mathcal{L}f_2)(s) = \frac{s}{s^2+1}\,.$$

Using a table for Laplace transforms, one can see that

$$(\mathcal{L}\cos)(s) = \frac{s}{s^2+1} \quad \text{and} \quad (\mathcal{L}\sin)(s) = \frac{1}{s^2+1}\,.$$

Thus,

$$f_1(t) = e^t - \sin t \quad \text{and} \quad f_2(t) = \cos t\,.$$

If one would obtain right-hand sides that are not found in tables for Laplace transforms, one would have to solve the following integral equations of the first kind:

$$\int_0^\infty e^{-st} f_i(t)\, dt = (\mathcal{L}f_i)(s)\,, \qquad i = 1,2\,.$$

As mentioned above, this integral equation is improperly posed and one has to be very careful when solving the equation numerically.

Much more important than the Laplace transform is the **Fourier transform**. For appropriate functions $f : \mathbb{R} \to \mathbb{R}$ the Fourier transform $\hat{f}$ is defined by

$$\hat{f}(s) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ist} f(t)\, dt\,. \tag{1.8}$$

If one wants to find $f$ from a given Fourier transform, (1.8) is again an integral equation of the first kind. However, it can be solved explicitly via:

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ist} \hat{f}(s)\, ds\,. \tag{1.9}$$

Thus, it turns out that equation (1.8) is well-posed and can be solved via (1.9) in a stable way.

**Example 1.4.** Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function with compact support. For $s \in \mathbb{R}^+$ and $\omega = (\cos\theta, \sin\theta)$, $\theta \in [0, 2\pi)$ let $\mathcal{L}(s, \omega) := s\omega + t\omega^\perp$ be the line with normal vector $\omega$ and distance $s$ from the origin. The function

$$(Rf)(s, \omega) := \int_{\mathbb{R}} f\left(s\omega + t\omega^\perp\right) dt\,, \qquad \omega \in \mathbb{R}^2\,,\ \|\omega\| = 1\,,\ s > 0\,, \tag{1.10}$$

is called **Radon transform** of $f$. The values of the Radon transform are all possible line integrals of $f$. The inverse problem of determining $f$ from $Rf$ was posed and solved analytically 1917 by the Austrian mathematician Johann Radon.

This problem is of great importance in applications of computerized tomography: let $D \subseteq \mathbb{R}^2$ be a compact domain with a spatially varying density $f$. In medical applications, $D$ symbolizes a cross-section of the human body; in nondestructive testing, $D$ is a cross-section of the material to be tested. The aim is to recover the density $f$ from X-ray measurements in the plane where $D$ lies. These X-rays travel along lines, which are parameterized by $\mathcal{L}(s, \omega)$ (see Figure 1.2).

Physical considerations lead to the following law: the intensity loss $-\Delta I$ of an X-ray beam along a distance $\Delta t$ is proportional to the intensity $I$, the density $f$, and the path length $\Delta t$, i.e.,

$$\Delta I\left(s\omega + t\omega^\perp\right) = -I\left(s\omega + t\omega^\perp\right) f\left(s\omega + t\omega^\perp\right) \Delta t\,.$$

By letting $\Delta t$ tend to 0, one obtains

$$\frac{d}{dt} I\left(sw + tw^\perp\right) = -I\left(sw + tw^\perp\right) f\left(sw + tw^\perp\right). \tag{1.11}$$

We denote by $I_L(s, w)$ and $I_0(s, w)$ the intensity of the X-ray beam measured at the detector and at the emitter, respectively, where detector and emitter, connected by the line $\mathcal{L}(s, \omega)$, are located outside of $D$ so that we can as well assume that they are located *at infinity*. Then, (1.11) has the solution

$$\log I_0(s, w) - \log I_L(s, w) = \int_{\mathbb{R}} f(sw + tw^\perp)\, dt = (Rf)(s, \omega)\,.$$

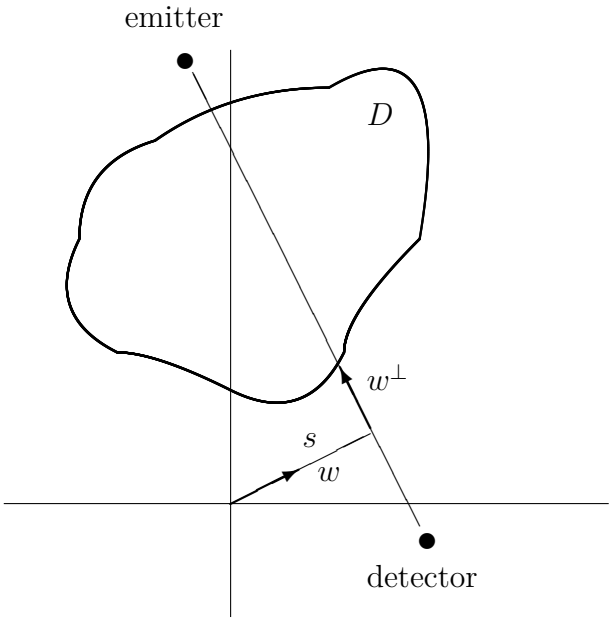


Figure 1.2: Computerized Tomography

Thus, the problem of computerized tomography corresponds to the inversion of the Radon transform.

**Example 1.5.** The following problem was posed and solved by Niels Henrik Abel about 1823: find the equation of a plane curve for which an object sliding without friction in uniform gravity to its lowest point from the height $s$ takes the total time of descent $f(s)$.

Abel's solution begins with the principle of conservation of energy: since the particle is frictionless, and thus loses no energy to heat, its kinetic energy at any point is exactly equal to the difference in potential energy from its starting point. The kinetic energy is $\frac{1}{2}mv^2$, the gravitational potential energy gained in falling from an initial height $s$ to a height $t$ is $mg(s-t)$. Here $m$ denotes the mass, $g$ the gravitational acceleration and $v = v(t)$ the velocity. Thus: $v(t) = \sqrt{2g(s-t)}$

Since the particle is constrained to move along a curve, the velocity satisfies the equation $v(t)\sin\alpha(t) = -\frac{dt}{d\tau}$, where $\tau$ is the time and $\alpha(t)$ is the angle between the $x$-axis and the tangent to the curve (see Figure 1.3).

Under the reasonable assumption that $t$ is strictly monotone and differentiable with respect to $\tau$, the inverse $\tau(t)$ exists and we obtain

$$\frac{d\tau}{dt} = -\frac{1}{\sqrt{2g(s-t)}}\,x(t) \quad \text{with} \quad x(t) := \frac{1}{\sin\alpha(t)}\,.$$

Integrating with respect to $t$ from 0 to $s$ yields:

$$\tau(s) - \tau(0) = -\int_0^s \frac{x(t)}{\sqrt{2g(s-t)}}\,dt$$

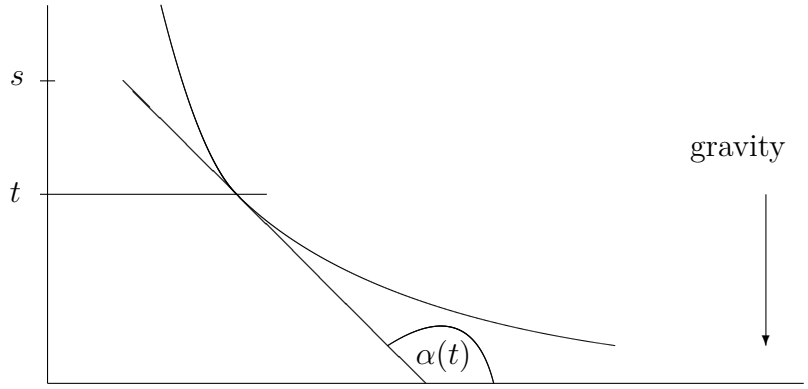


Figure 1.3: Sketch of the curve

Due to the definition of $f$ and the function $\tau$, we have: $\tau(s) - \tau(0) = 0 - f(s)$. Thus, $x$ satisfies the following Volterra integral equation of the first kind:

$$\int_0^s \frac{x(t)}{\sqrt{s-t}}\, dt = \sqrt{2g} f(s)\,. \tag{1.12}$$

Note that knowing $x$, one can uniqely determin $\alpha$ and, hence, also the curve, since only values in $[\frac{\pi}{2}, \pi)$ are reasonable.

This weakly singular integral equation is called **Abel's integral equation**. The same name is used for integral equations, whose kernel involves the expression $(s-t)^\alpha$ with $\alpha \in (0,1)$ in the denominator (see Chapter 4).

# 2. Fredholm integral equations of the second kind

In this chapter we deal with equations of the following type:

$$\lambda x(s) - \int_G k(s,t)x(t)\,dt = f(s)\,, \qquad s \in G\,, \tag{2.1}$$

with $\lambda \neq 0$, where $G \subseteq \mathbb{R}^N$ is compact and Jordan measurable with positive measure. $x$ is an unknown and $f$ a given function. The parameter $\lambda$ will be of interest only for eigenvalue problems. Therefore, we sometimes set $\lambda = 1$.

The kernel $k$ induces an integral operator $K$, defined by

$$(Kx)(s) := \int_G k(s,t)x(t)\,dt\,. \tag{2.2}$$

In the next theorem conditions for the kernel $k$ are given so that the corresponding operator $K$ maps $L^2(G)$ or $C(G)$ to itself, respectively (see also Theorem 2.13 below).

**Theorem 2.1.** If $k \in C(G\times G)$, then $K(L^2(G)) \subseteq C(G)$, especially $K(C(G)) \subseteq C(G)$. If $k \in L^2(G\times G)$, then $K(L^2(G)) \subseteq L^2(G)$. Here, $K$ is the integral operator with kernel $k$.

**Proof:** Let $x \in L^2(G)$, $k \in C(G \times G)$, $s, \sigma \in G$. Then, due to the Cauchy-Schwarz inequality,

$$\begin{aligned} |(Kx)(s) - (Kx)(\sigma)|^2 &\leq \left[\int_G |k(s,t) - k(\sigma,t)|\,|x(t)|\,dt\right]^2 \\ &\leq \int_G |k(s,t) - k(\sigma,t)|^2\,dt\,\|x\|_2^2 \\ &\leq \sup_{t\in G} |k(s,t) - k(\sigma,t)|^2 |G|\,\|x\|_2^2\,. \end{aligned}$$

Since $G \times G$ is compact, $k$ is uniformly continuous. Hence:

$$\lim_{\sigma\to s} \sup_{t\in G} |k(s,t) - k(\sigma,t)| = 0$$

Together with the estimate above this implies the continuity of $Kx$, i.e., $Kx \in C(G)$.

Let us now assume that $k \in L^2(G \times G)$. Then, due to the Cauchy-Schwarz inequality and Fubini's theorem,

$$\begin{aligned} \|Kx\|_2^2 &= \int_G |(Kx)(s)|^2\,ds = \int_G \left|\int_G k(s,t)x(t)\,dt\right|^2 ds \\ &\leq \int_G \left[\int_G |k(s,t)|^2\,dt \int_G |x(t)|^2\,dt\right] ds = \int_G\int_G |k(s,t)|^2\,dtds\,\|x\|_2^2 \\ &= \int_{G\times G} |k(s,t)|^2\,d(s,t)\,\|x\|_2^2 = \|k\|_2^2\,\|x\|_2^2 < \infty \end{aligned}$$

Thus, $Kx \in L^2(G)$. ∎

In the proof above and below, $\|\cdot\|_2$ always denotes the $L^2$-norm on the appropriate space, e.g., in the special case above the norm on $L^2(G \times G)$ and on $L^2(G)$, respectively. Obviously, here and below, all integrals are Lebesgue integrals.

Theorem 2.1 tells us that depending on the type of the kernel the corresponding integral operator can be considered as an operator from $L^2(G)$ to $C(G)$ or $L^2(G)$ or from $C(G)$ to $C(G)$ or also $L^2(G)$. The operator norm of this bounded operator depends on the image and preimage spaces. It is, e.g., an immediate consequence of the proof of Theorem 2.1 that for $K$ considered as an operator from $L^2(G)$ to $L^2(G)$, it holds that

$$\|K\|^2 \leq \int_{G\times G} |k(s,t)|^2 \, d(s,t) \, .$$

## 2.1. Degenerate kernels

Very simple kernels are so called degenerate kernels:

**Definition 2.2.** $k : G \times G \to \mathbb{R}$ is called a **degenerate kernel**, if finitely many functions $\varphi_1, \ldots, \varphi_n, \psi_1, \ldots, \psi_n \in L^2(G)$ exist such that

$$k(s,t) = \sum_{i=1}^{n} \varphi_i(s)\psi_i(t) \qquad \text{a.e.} \tag{2.3}$$

As above an equality of functions in $L^2$ is always meant *almost everywhere* (a.e.), even though it is not always explicitly mentioned.

Degenerate kernels induce integral operators with finite-dimensional ranges and vice versa as the following theorem shows:

**Theorem 2.3.** Let $K : L^2(G) \to L^2(G)$ be an integral operator with $L^2$-kernel $k$ and range $\mathcal{R}(K)$. Then $k$ is degenerate if and only if $\mathcal{R}(K)$ is finite-dimensional.

**Proof:** Let $\langle \cdot, \cdot \rangle$ denote the usual inner product in $L^2(G)$ or $L^2(G \times G)$, respectively.

First we assume that $k$ is degenerate, i.e.,

$$k(s,t) = \sum_{i=1}^{n} \varphi_i(s)\psi_i(t) \quad \text{with} \quad \varphi_1, \ldots, \varphi_n, \psi_1, \ldots, \psi_n \in L^2(G) \, .$$

Then

$$(Kx)(s) \quad = \quad \int_G \sum_{i=1}^{n} \varphi_i(s)\psi_i(t)x(t) \, dt = \sum_{i=1}^{n} \left( \int_G \psi_i(t)x(t) \, dt \right) \varphi_i(s)$$

$$\implies \quad Kx \in \operatorname{span}\{\varphi_1, \ldots, \varphi_n\} \qquad \text{for all} \qquad x \in L^2(G)\,.$$

Therefore, $\mathcal{R}(K) \subseteq \operatorname{span}\{\varphi_1, \ldots, \varphi_n\}$ and, hence, $\mathcal{R}(K)$ is finite-dimensional.

Now we assume that $n := \dim \mathcal{R}(K) < \infty$ and that $\{\varphi_1, \ldots, \varphi_n\}$ is an orthonormal basis of $\mathcal{R}(K)$. This basis can be completed to a complete orthonormal system $\{\varphi_1, \varphi_2, \ldots\}$ of $L^2(G)$. Since, obviously,

$$\int_{G\times G} \varphi_i(s)\overline{\varphi_j(t)}\,\overline{\varphi_k(s)}\varphi_m(t)\,d(s,t) = \delta_{ik}\cdot\delta_{jm}$$

and since, due to Fubini's theorem, it is easy to show that for any $k \in L^2(G\times G)$ with

$$\int_{G\times G} k(s,t)\overline{\varphi_i(s)}\varphi_j(t)\,d(s,t) = 0 \quad \text{for all} \quad i,j \in \mathbb{N}$$

it follows that $k = 0$, it holds that $\{\alpha_{ij}(s,t) := \varphi_i(s)\overline{\varphi_j(t)} : i,j \in \mathbb{N}\}$ is a complete orthonormal system of $L^2(G\times G)$.

Since $k \in L^2(G\times G)$, it can be expanded into a Fourier series as follows: a sequence $(a_{ij})_{(i,j)\in\mathbb{N}\times\mathbb{N}}$ exists with

$$k = \sum_{i,j\in\mathbb{N}} a_{ij}\alpha_{ij} \quad \text{(convergence in } L^2(G\times G)) \quad \text{and} \quad \sum_{i,j\in\mathbb{N}} a_{ij}^2 < \infty\,.$$

For all $x \in L^2(G)$ and $m \in \mathbb{N}$ we now obtain with Fubini's theorem that

$$\langle\, Kx, \varphi_m \,\rangle = \int_G\int_G k(s,t)x(t)\,dt\,\overline{\varphi_m(s)}\,ds = \int_{G\times G} k(s,t)x(t)\overline{\varphi_m(s)}\,d(s,t) = \langle\, k, x_m \,\rangle$$

with $x_m(s,t) := \overline{x(t)}\varphi_m(s)$. Thus, $\langle\, Kx, \varphi_m \,\rangle = \sum\limits_{i,j\in\mathbb{N}} a_{ij}\,\langle\, \alpha_{ij}, x_m \,\rangle$. Since, due to Fubini's theorem,

$$\begin{aligned}\langle\, \alpha_{ij}, x_m \,\rangle &= \int_{G\times G} \varphi_i(s)\overline{\varphi_j(t)}x(t)\overline{\varphi_m(s)}\,d(s,t)\\ &= \int_G x(t)\overline{\varphi_j(t)}\,dt\int_G \varphi_i(s)\overline{\varphi_m(s)}\,ds \;=\; \delta_{im}\,\langle\, x, \varphi_j \,\rangle\end{aligned}$$

it follows that $\langle\, Kx, \varphi_m \,\rangle = \sum\limits_{j\in\mathbb{N}} a_{mj}\,\langle\, x, \varphi_j \,\rangle$.

For $m > n$ we have that $\varphi_m \in \mathcal{R}(K)^\perp$ so that $\langle\, Kx, \varphi_m \,\rangle = 0$.

Choosing especially $x = \varphi_r$ with $r \in \mathbb{N}$, we obtain for $m > n$:

$$0 = \langle\, K\varphi_r, \varphi_m \,\rangle = \sum_{j\in\mathbb{N}} a_{mj}\underbrace{\langle\, \varphi_r, \varphi_j \,\rangle}_{=\delta_{rj}} = a_{mr}\,.$$

Therefore, $k = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{\infty} a_{ij}\alpha_{ij}$, since the Fourier series can be reordered with respect to $\{\alpha_{ij}\}$ without influencing the $L^2$-convergence. Setting $\psi_i(t) := \sum\limits_{j=1}^{\infty} a_{ij}\overline{\varphi_j(t)}$, which is an

element of $L^2(G)$, since $\sum_{j=1}^{\infty} a_{ij}^2 < \infty$, we obtain that

$$k(s,t) = \sum_{i=1}^{n}\sum_{j=1}^{\infty} a_{ij}\alpha_{ij}(s,t) = \sum_{i=1}^{n}\sum_{j=1}^{\infty} a_{ij}\varphi_i(s)\overline{\varphi_j(t)} = \sum_{i=1}^{n}\varphi_i(s)\psi_i(t)\,.$$

Thus, the kernel $k$ is degenerate. ∎

**Remark 2.4.** It is an immediate consequence of the proof above that $\mathcal{R}(K)$ is spanned by the functions $\varphi_1, \ldots, \varphi_n$ and has at most dimension $n$. A Fredholm integral equation of the second kind with a degenerate kernel leads to a linear system:

We consider (2.1) with a kernel of type (2.3), where we assume w.l.og. that the functions $\varphi_1, \ldots, \varphi_n$ are linearly independent. We assume that all functions are real valued. The complex valued case can be treated analogously (see the remark after Theorem 2.6).

Plugging in (2.3) into (2.1), we obtain

$$\lambda x(s) - \sum_{i=1}^{n}\varphi_i(s)\int_G \psi_i(t)x(t)\,dt = f(s)\,, \qquad s \in G\,. \tag{2.4}$$

Building the $L^2$-inner products with $\psi_j$, $1 \le j \le n$, yields

$$\lambda\int_G x(s)\psi_j(s)\,ds - \sum_{i=1}^{n}\int_G \varphi_i(s)\psi_j(s)\,ds\int_G \psi_i(t)x(t)\,dt = \int_G f(s)\psi_j(s)\,ds$$

or eqivalently the linear system

$$\lambda\bar{x} - A\bar{x} = \bar{f} \tag{2.5}$$

with

$$\begin{aligned}
\bar{x}_j &:= \int_G x(s)\psi_j(s)\,ds\,, \quad 1 \le j \le n\,, & \bar{x} &:= (\bar{x}_1, \ldots, \bar{x}_n)^\top\,,\\
\bar{f}_j &:= \int_G f(s)\psi_j(s)\,ds\,, \quad 1 \le j \le n\,, & \bar{f} &:= (\bar{f}_1, \ldots, \bar{f}_n)^\top\,,\\
a_{ij} &:= \int_G \varphi_i(s)\psi_j(s)\,ds\,, \quad 1 \le i,j \le n\,, & A &:= (a_{ji})_{1\le i,j\le n}\,.
\end{aligned} \tag{2.6}$$

If $\lambda$ is no eigenvalue of $A$, system (2.5) has a unique solution. It then follows from (2.4) that

$$x = \frac{1}{\lambda}f + \frac{1}{\lambda}\sum_{i=1}^{n}\bar{x}_i\varphi_i\,. \tag{2.7}$$

I.e., this is the explicit formula of the (unique) solution of (2.4). It is an element of $L^2(G)$ or of $C(G)$ if $f$ and $\varphi_1, \ldots, \varphi_n$ are continuous. Note that one needs the solution of the linear system (2.5).

Solving a Fredholm integral equation of the second kind with a degenerate kernel is, therefore, not an infinite-dimensional problem but a finite-dimensional one. Using well-known results from Linear Algebra one can derive results about existence and uniqueness of solutions of (2.4). We will formulate these results in such a way that we can

later see that they are a special case of the so called **Fredholm alternative** for integral equations of the second kind with a general kernel. We first formulate the Fredholm alternative for real linear systems:

**Theorem 2.5.** Let $B$ be an $n \times n$-matrix with rank $m \leq n$. Then the homogeneous systems

$$Bx = 0 \qquad \text{and} \qquad B^\top y = 0 \tag{2.8}$$

have the same number of $n - m$ linearly independent solutions. The inhomogeneous equation

$$Bx = f \tag{2.9}$$

is solavable if and only if

$$\langle f, y \rangle = 0 \tag{2.10}$$

for all $y$ with $B^\top y = 0$. Here, $\langle \cdot, \cdot \rangle$ denotes the general inner product in $\mathbb{R}^n$.

**Proof:** Since the column rank and row rank of a matrix are equal, also the rank of $B^\top$ is equal to $m$. The number of linearly independent solutions of the homogeneous systems in (2.8), thus satisfies:

$$\dim \mathcal{N}(B) = n - \dim \mathcal{R}(B) = n - m = n - \dim \mathcal{R}(B^\top) = \dim \mathcal{N}(B^\top)$$

This proves the first assertion.

Let us now assume that (2.9) is solvable, i.e., $Bx = f$ for some $x \in \mathbb{R}^n$. Then it holds for every $y$ with $B^\top y = 0$ that

$$\langle f, y \rangle = \langle Bx, y \rangle = \langle x, B^\top y \rangle = 0\,,$$

i.e., (2.10) holds. We have also shown that $\mathcal{R}(B) \subseteq \mathcal{N}(B^\top)^\perp$.

Now we assume that (2.10) holds for all $y$ with $B^\top y = 0$, i.e., for all $y \in \mathcal{N}(B^\top)$. This means that

$$f \in \mathcal{N}(B^\top)^\perp\,. \tag{2.11}$$

Since $\dim \mathcal{R}(B) = m$ and $\dim \mathcal{N}(B^\top)^\perp = n - \dim \mathcal{N}(B^\top) = n - (n - m) = m$, we obtain together with $\mathcal{R}(B) \subseteq \mathcal{N}(B^\top)^\perp$ (see above) that

$$\mathcal{R}(B) = \mathcal{N}(B^\top)^\perp\,.$$

Due to (2.11), this means that $f \in \mathcal{R}(B)$ and, hence, (2.9) is solvable. ∎

Note that this proof includes some steps that are only possible in finite-dimensional spaces.

The equation on the right-hand side of (2.8) is called **adjoint homogeneous equation**.

As an immediate consequence of Theorem 2.5 we obtain the Fredholm alternative for Fredholm integral equations of the second kind with a degenerate kernel:

**Theorem 2.6.** Let $k$ be a degenerate $L^2$-kernel, $\lambda \neq 0$, and $f \in L^2(G)$. Then the equations

$$\lambda x(s) - \int_G k(s,t)x(t)\,dt = 0\,, \qquad s \in G\,, \tag{2.12}$$

and

$$\lambda y(s) - \int_G k(t,s)y(t)\,dt = 0\,, \qquad s \in G\,, \tag{2.13}$$

have the same number of linearly independent solutions.

The inhomogeneous equation (2.1) has a solution $x \in L^2(G)$ if and only if all solutions $y \in L^2(G)$ of (2.13) satisfy

$$\int_G f(s)y(s)\,ds = 0\,.$$

**Proof:** For the special system (2.5) the adjoint homogeneous equation is given by

$$\lambda \bar{y} - A^\top \bar{y} = 0\,. \tag{2.14}$$

As for the derivation of (2.7), one can show that for every solution $\bar{y}$ of (2.14) the function

$$y := \frac{1}{\lambda}\sum_{j=1}^{n} \bar{y}_j \psi_j$$

solves equation (2.13). And conversely for every solution $y$ of (2.13) the vector

$$\left(\int_G \varphi_i(t)y(t)\,dt\right)_{1\le i\le n}$$

solves (2.14). In this sense (2.14) and (2.13) are equivalent. Now Theorem 2.5 implies that the homogeneous equation (2.12) and (2.13) have the same number of linearly independent solutions. This number equals $\dim \mathcal{N}(\lambda I - A)$, i.e., the geometric multiple of the eigenvalue $\lambda$ of $A$.

The inhomogeneous equation (2.1) is solvable if and only if (2.5) is solvable. Due to Theorem 2.5, this is satisfied if and only if

$$\sum_{i=1}^{n} \bar{f}_i \bar{y}_i = 0$$

for all solutions $\bar{y}$ of (2.14). Noting that

$$\int_G f(s)\left(\frac{1}{\lambda}\sum_{i=1}^{n} \bar{y}_i\psi_i(s)\right)ds = \frac{1}{\lambda}\sum_{i=1}^{n}\bar{y}_i\int_G f(s)\psi_i(s)\,ds = \frac{1}{\lambda}\sum_{i=1}^{n}\bar{y}_i\bar{f}_i$$

it follows with the equivalence of (2.14) and (2.13) that (2.1) is solvable if and only if

$$\int_G f(s)y(s)\,ds = 0\,.$$

This proves the assertions. ∎

It is an immediate consequence of Theorem 2.6 that (2.1) is solvable for all right-hand sides $f \in L^2(G)$ if (2.12) (or (2.13), respectively) only have the trivial solution $x = 0$ (or $y = 0$, respectively) or equivalently that $\lambda$ is no eigenvalue of the integral operator induced by the kernel $k$.

The integral operator with kernel $k(t, s)$ instead of $k(s, t)$ is called adjoint integral operator. In case of an $L^2$-kernel its so-called adjoint kernel is only almost everywhere uniquely determined. For equations with complex valued kernels and functions one can either prove Theorem 2.6 in the same way or one can define the adjoint kernel by $\overline{k(t, s)}$. In the latter case one has to replace $\lambda$ by $\bar{\lambda}$ in the adjoint equation (2.13).

Using functional analytic methods, one can show that the Fredholm alternative also holds for general $L^2$-kernels. Thereto, it is necessary to interpret all integral operators as operators between Banach or Hilbert spaces.

A property that is very essential for the theory of linear integral equations is the compactness of the corresponding integral operators.

## 2.2. Compact operators

**Definition 2.7.** Let $X, Y$ be normed spaces and $K : X \to Y$ linear. $K$ is called **compact** if for any bounded set $B \subseteq X$ it holds that $\overline{K(B)}$ is compact.

In the next theorem we collect some basic properties of compact linear operators:

**Theorem 2.8.** Let $X, Y, Z$ be normed spaces and $K : X \to Y$ linear. Then the following assertions hold:

(a) $K$ is compact if and only if $(Kx_n)$ has a convergent subsequence for any bounded sequence $(x_n)$.

(b) If $K$ is compact, then $K$ is bounded, i.e., $K \in L(X, Y)$.

(c) Linear combinations of compact linear operators are compact.

(d) Let $K_1 \in L(X, Y)$ and $K_2 \in L(Y, Z)$. Then $K_1 K_2$ is compact if either $K_1$ or $K_2$ is compact.

(e) Let $Y$ be a Banach space and let $K_n : X \to Y$ be linear and compact for all $n \in \mathbb{N}$. Moreover, $\lim_{n\to\infty} \|K - K_n\| = 0$. Then $K$ is compact.

(f) If $K$ is bounded and $\dim \mathcal{R}(K) < \infty$, then $K$ is compact.

**Proof:**

(a) Let $K$ be compact and let $(x_n)$ be a bounded sequence in $X$. Then $(Kx_n)$ is a sequence in a compact set. By definition of compact sets, it then has a convergent subsequence.

Now assume that $K$ is not compact. Then there exists a bounded set $B \subseteq X$ such that $\overline{K(B)}$ is not compact, i.e., a sequence $(\bar{y}_n)$ exists in $\overline{K(B)}$ that has no convergent subsequence. Note that $y_n \in K(B)$ exists with $\|y_n - \bar{y}_n\| \leq \frac{1}{n}$ for all $n \in \mathbb{N}$. The sequence $(y_n)$ also has no convergent subsequence, because otherwise there would exist $y$ and $(y_{n_k})$ with $y_{n_k} \to y$. But then

$$\|\bar{y}_{n_k} - y\| \leq \|y_{n_k} - y\| + \frac{1}{n_k} \underset{k\to\infty}{\longrightarrow} 0\,,$$

which is a contradiction to the choice of $(\bar{y}_n)$. Obviously for all $n \in \mathbb{N}$ an $x_n \in B$ exists with $Kx_n = y_n$. Thus, we found a bounded sequence $(x_n)$, where $(Kx_n)$ has no convergent subsequence.

(b) Assume that $K$ is not bounded. Then a sequence $(x_n)$ exists with $\|x_n\| = 1$ and $\|Kx_n\| \to \infty$ as $n \to \infty$. But then $(Kx_n)$ cannot have a convergent subsequence. This is a contradiction to (a). Thus, $K$ is bounded and, hence, as a linear operator also continuous.

(c) Let $K_1$ and $K_2$ be compact operators and let $\lambda_1, \lambda_2 \in \mathbb{R}$. Moreover, let $(x_n)$ be a bounded sequence. Since $K_1$ is compact, there exists a $y_1$ and $(n_k)$ such that $K_1 x_{n_k} \to y_1$ as $k \to \infty$. Since $x_{n_k}$ is also bounded, it follows with the compactness of $K_2$ that $y_2$ and $(n_{k_l})$ exist with $K_2 x_{n_{k_l}} \to y_2$ as $l \to \infty$. Since also $K_1 x_{n_{k_l}} \to y_1$ as $l \to \infty$, we finally get that

$$\lambda_1 K_1 x_{n_{k_l}} + \lambda_2 K_2 x_{n_{k_l}} \underset{l\to\infty}{\longrightarrow} \lambda_1 y_1 + \lambda_2 y_2\,.$$

Due to (a), this shows that $\lambda_1 K_1 + \lambda_2 K_2$ is compact.

(d) Assume that $K_1$ is compact and that $K_2$ is bounded. For a bounded sequence $(x_n)$ also $(K_2 x_n)$ is bounded. Therefore, due to (a), $(K_1 K_2 x_n)$ has a convergent subsequence. Thus, $K_1 K_2$ is compact.

Now let $K_1$ be bounded and $K_2$ be compact. For a bounded sequence $(x_n)$, due to (a), there exists a convergent subsequence $(K_2 x_{n_k})$. But, due to the continuity of $K_1$, then also $(K_1 K_2 x_{n_k})$ will converge. Thus, again $K_1 K_2$ is compact.

(e) Let $(x_n)$ be bounded in $X$ with $\|x_n\| \leq C$ for all $n \in \mathbb{N}$. Since $K_1$ is compact, the sequence $(K_1 x_n)_{n\in\mathbb{N}}$ has a convergent subsequence $(K_1 x_{n_1(k)})_{k\in\mathbb{N}}$. Since $K_2$ is compact, the sequence $(K_2 x_{n_1(k)})_{k\in\mathbb{N}}$ has a convergent subsequence $(K_2 x_{n_2(k)})_{k\in\mathbb{N}}$. By induction we may construct subsequences $(x_{n_i(k)})_{k\in\mathbb{N}}$ of $(x_{n_{i-1}(k)})_{k\in\mathbb{N}}$ such that $(K_i x_{n_i(k)})_{k\in\mathbb{N}}$ converges for all $i \in \mathbb{N}$. Note that then also $(K_i x_{n_j(k)})_{k\in\mathbb{N}}$ converges for all $i, j \in \mathbb{N}$ with $j \geq i$.

Setting $\bar{x}_i := x_{n_i(i)}$ for all $i \in \mathbb{N}$, we have that $(\bar{x}_i)_{i\in\mathbb{N}}$ is a subsequence of $(x_n)_{n\in\mathbb{N}}$ and also that $(\bar{x}_i)_{i\geq j}$ is a subsequence of $(x_{n_j(k)})_{k\in\mathbb{N}}$. We show that

$$(K\bar{x}_i) \qquad \text{is convergent.} \tag{2.15}$$

Let $\varepsilon > 0$ be arbitray, but fixed, and let $n \in \mathbb{N}$ be such that

$$\|K_n - K\| < \frac{\varepsilon}{3C}\,.$$

This is possible, since $\lim_{n\to\infty} \|K - K_n\| = 0$.

It follows from the definition of $\bar{x}_i$ that $(K_n\bar{x}_i)_{i\in\mathbb{N}}$ converges. Let $i_0 \in \mathbb{N}$ be such that $\|K_n\bar{x}_i - K_n\bar{x}_j\| < \frac{\varepsilon}{3}$ for all $i, j \geq i_0$. Then

$$\begin{aligned}\|K\bar{x}_i - K\bar{x}_j\| &\leq \|K\bar{x}_i - K_n\bar{x}_i\| + \|K_n\bar{x}_i - K_n\bar{x}_j\| + \|K_n\bar{x}_j - K\bar{x}_j\| \\ &\leq \|K - K_n\|\,(\|\bar{x}_i\| + \|\bar{x}_j\|) + \frac{\varepsilon}{3} < \frac{\varepsilon}{3C}2C + \frac{\varepsilon}{3} = \varepsilon\end{aligned}$$

for all $i, j \geq i_0$. Thus, $(K\bar{x}_i)$ is a Cauchy sequence. Since $Y$ is a Banach space and, hence, complete, this implies (2.15). Since $(x_n)$ was an arbitrary bounded sequence, the compactness of $K$ follows with (a).

(f) This assertion is an immediate consequence of the Heine-Borel theorem which tells us that the closure of a bounded subset of the finite-dimensional space $\mathcal{R}(K)$ is compact.

∎

According to Theorem 2.8 (e) and (f), the norm limit of finite-dimensional bounded linear operators is compact. This fact is used in several compactness proofs for integral operators. Under certain conditions one can prove a converse result, namely that compact operators are limits of finite-dimensional operators:

**Lemma 2.9.** Let $X$ and $Y$ be Banach spaces and let $T_n, T : X \to Y$, $n \in \mathbb{N}$, be bounded linear operators satisfying that $T_n \to T$ pointwise, i.e., $T_n x \to Tx$ for all $x \in X$. Then

$$\lim_{n\in\mathbb{N}} \sup_{x\in A} \|(T_n - T)y\| = 0$$

for every compact set $A \subset X$.

Especially, if $P_n \in L(Y)$, $n \in \mathbb{N}$, is such that $\dim \mathcal{R}(P_n) < \infty$ and that $P_n \to I$ pointwise, then $\|P_n K - K\| \to 0$ for any compact linear operator $K : Z \to Y$, where $Z$ is a normed space.

**Proof:** Due to the Banach-Steinhaus theorem $\sup_{m\in\mathbb{N}} \|T_m - T\| < \infty$. Let $A$ be a compact subset of $X$ and let $\varepsilon > 0$ be arbitrary, but fixed. Then there exist $k \in \mathbb{N}$ and $x_1, \ldots, x_k \in A$ so that $\inf\{\|x - x_i\| : i \in \{1, \ldots, k\}\} < \varepsilon$ for all $x \in A$. Therefore,

$$\begin{aligned}\sup_{x\in A} \|(T_n - T)x\| &\leq \sup\{\|(T_n - T)x_i\| : i \in \{1, \ldots, k\}\} \\ &\quad + \|(T_n - T)\| \inf\{\|x - x_i\| : i \in \{1, \ldots, k\}\} \\ &\leq \sup\{\|T_n x_i - Tx_i\| : i \in \{1, \ldots, k\}\} + \varepsilon \sup_{m\in\mathbb{N}} \|T_m - T\|\end{aligned}$$

for all $n \in \mathbb{N}$. Since the first term in the estimate above converges to 0 as $n \to \infty$, it holds that

$$\limsup_{n\to\infty} \sup_{x\in A} \|(T_n - T)x\| \leq \varepsilon \sup_{m\in\mathbb{N}} \|T_m - T\|\,.$$

Since this holds for all $\varepsilon > 0$, it follows that $\lim_{n\to\infty} \sup_{x\in A} \|(T_n - T)x\| = 0$.

Let now $P_n \in L(Y)$ with $P_n \to I$ pointwise and let $K$ be compact. Then $\overline{K(U)}$ is compact, where $U$ denotes the unit sphere of $Z$. Using the result from above with $X := Y$, $T_n := P_n$, $T := I$, and $A := \overline{K(U)}$, we obtain that

$$\begin{aligned} 0 &\leq \limsup_{n\to\infty} \|P_n K - K\| = \limsup_{n\to\infty} \sup_{z\in U} \|P_n K z - K z\| \\ &\leq \limsup_{n\to\infty} \sup_{y\in\overline{K(U)}} \|(P_n - I)y\| = 0 \end{aligned}$$

and, hence, $\lim_{n\to\infty} \|P_n K - K\| = 0$. ∎

Obviously, the requirement that $\dim \mathcal{R}(P_n) < \infty$ is not needed.

A practical example of a sequence of operators $(P_n)$ as in Lemma 2.9 is the sequence of orthogonal projectors of a separable Hilbert space $Y$ onto a sequence of finite-dimensional subspaces $(Y_n)$ with the property $\overline{\bigcup_{n\in\mathbb{N}} Y_n} = Y$. Thus, a compact linear operator with values in a separable Hilbert space is the norm limit of finite-dimensional bounded linear operators.

We will prove in Lemma 2.15 that the identity operator $I : X \to X$ is compact if and only if $X$ is finite-dimensional. Since $I$ is always bounded, the converse result of Theorem 2.8 (b) is not true in infinite-dimensional spaces, i.e., there exist non-compact bounded linear operators.

Next we deal with the question, under what conditions an integral operator is compact. As a first result we show that an integral operator with a continuous kernel is compact on $L^2(G)$ and $C(G)$, respectively:

**Theorem 2.10.** Let $k \in C(G \times G)$ and let $K$ be the corresponding integral operator according to (2.2). Then $K$ is compact as an operator from each of the spaces $L^2(G)$ and $C(G)$ to each of these spaces.

**Proof:** Actually, this theorem consists of four results depending on the chosen image and preimage. Since $\|x\|_2 \leq \sqrt{|G|}\,\|x\|_\infty$ for all $x \in C(G)$, the embedding operator $E : C(G) \to L^2(G)$ is bounded. Since the operators

$$K : C(G) \to C(G)\,, \quad K : C(G) \to L^2(G)\,, \quad \text{and} \quad K : L^2(G) \to L^2(G)$$

each can be written as composition of the operator $K : L^2(G) \to C(G)$ with the embedding operator $E$, it follows from Theorem 2.8 (d) that we only have to prove the compactness of $K : L^2(G) \to C(G)$. The well-definedness of this operator is already guaranteed by Theorem 2.1.

We will do this by approximating $K$ in operator norm sense by a sequence of finite-dimensional operators and will then apply Theorem 2.8 (e).

Let $\Delta_{1,n}, \ldots, \Delta_{n,n} \subseteq \mathbb{R}^N$, $n \in \mathbb{N}$, be pairwise disjoint open Jordan measurable sets (e.g., Cartesian products of open intervals) with $\bigcup_{i=1}^n \overline{(\Delta_{i,n} \cap G)} = G$. Since $G$ is bounded, a

partition $\{\Delta_{i,n}\}$ can be chosen such that

$$\lim_{n\to\infty} \max_{1\le i\le n} \sup_{t,\tau\in\Delta_{i,n}} \|t-\tau\| = 0\,. \tag{2.16}$$

Now we choose a system of intermediate points $t_{i,n} \in \Delta_{i,n} \cap G$ and define

$$\begin{aligned} k_n : G\times G &\to \mathbb{R} \\ (s,t) &\mapsto \begin{cases} k(s,t_{i,n})\,, & t\in\Delta_{i,n}\cap G\,, \\ 0\,, & t\in G\setminus\bigcup_{i=1}^n \Delta_{i,n}\,. \end{cases} \end{aligned}$$

with corresponding integral operator $K_n$ according to (2.2). The integral exists in the Lebesgue sense, since $G\setminus\bigcup_{i=1}^n\Delta_{i,n}$ can only contain boundary points of the sets $\Delta_{i,n}$ and boundaries are Jordan measurable null sets. Thus, $|G\setminus\bigcup_{i=1}^n\Delta_{i,n}| = 0$ and

$$(K_n x)(s) = \sum_{i=1}^n k(s,t_{i,n}) \int_{\Delta_{i,n}\cap G} x(t)\,dt$$

for all $x\in L^2(G)$, $s\in G$, and $n\in\mathbb{N}$.

This implies that $K_n$ is a linear operator from $L^2(G)$ to $C(G)$ with finite-dimensional range $\mathcal{R}(K_n)\subseteq \operatorname{span}\{k(.,t_{i,n}) : 1\le i\le n\}\subseteq C(G)$. Since, due to the Cauchy-Schwarz inequality,

$$\|K_n x\|_\infty \le \sum_{i=1}^n \|k(.,t_{i,n})\|_\infty \int_{\Delta_{i,n}\cap G} |x(t)|\,dt \le \|k\|_\infty \int_G |x(t)|\,dt \le \|k\|_\infty \|x\|_2 \sqrt{|G|}\,,$$

$K_n : L^2(G)\to C(G)$ is bounded. Together with Theorem 2.8 (f) we get the compactness of $K_n : L^2(G)\to C(G)$.

Let now $\varepsilon>0$ be arbitrary, but fixed. Since $k$ is uniformly continuous on the compact set $G\times G$, a $\delta>0$ exists such that

$$\|t-\tau\|<\delta \Longrightarrow |k(s,t)-k(s,\tau)| < \frac{\varepsilon}{\sqrt{|G|}} \quad \text{for all } s\in G\,.$$

Together with (2.16) this implies that $n_0\in\mathbb{N}$ exists such that

$$|k(s,t)-k(s,t_{i,n})| < \frac{\varepsilon}{\sqrt{|G|}} \quad \text{for all} \quad s\in G\,,\ n\ge n_0\,,\ i\in\{1,\dots,n\}\,,\ t\in\Delta_{i,n}\,.$$

This, together with the Cauchy-Schwarz inequality, yields that

$$\begin{aligned} |(Kx)(s)-(K_n x)(s)| &= \left|\sum_{i=1}^n \int_{\Delta_{i,n}\cap G} k(s,t)x(t)\,dt - \sum_{i=1}^n k(s,t_{i,n}) \int_{\Delta_{i,n}\cap G} x(t)\,dt\right| \\ &\le \sum_{i=1}^n \left|\int_{\Delta_{i,n}\cap G} (k(s,t)-k(s,t_{i,n}))x(t)\,dt\right| \\ &\le \sum_{i=1}^n \sqrt{\int_{\Delta_{i,n}\cap G} |k(s,t)-k(s,t_{i,n})|^2\,dt}\sqrt{\int_{\Delta_{i,n}\cap G} |x(t)|^2\,dt} \\ &\le \sqrt{\sum_{i=1}^n \int_{\Delta_{i,n}\cap G} |k(s,t)-k(s,t_{i,n})|^2\,dt \sum_{i=1}^n \int_{\Delta_{i,n}\cap G} |x(t)|^2\,dt} \\ &\le \sqrt{\sum_{i=1}^n \frac{\varepsilon^2}{|G|}|\Delta_{i,n}\cap G|\,\|x\|_2^2} \;=\; \varepsilon\,\|x\|_2 \end{aligned}$$

for all $n \geq n_0$, $x \in L^2(G)$, and $s \in G$. This implies that

$$\|K - K_n\| = \sup_{\|x\|_2 \leq 1} \sup_{s \in G} |(Kx)(s) - (K_n x)(s)| \leq \varepsilon \,,$$

where $\|K - K_n\|$ denotes the operator norm for linear operators from $L^2(G)$ to $C(G)$. Since $\varepsilon > 0$ was arbitrary, this implies that $\lim_{n\to\infty} \|K - K_n\| = 0$, which together with Theorem 2.8 (e) yields the compactness of $K : L^2(G) \to C(G)$. ∎

Approximating integral operators with $L^2$-kernels by operators with continuous kernels, one can show that such operators are compact from $L^2(G)$ to $L^2(G)$:

**Theorem 2.11.** Let $k \in L^2(G \times G)$ and let $K$ be the corresponding integral operator according to (2.2). Then $K$ is compact from $L^2(G)$ to $L^2(G)$.

**Proof:** Since $C(G \times G)$ is dense in $L^2(G \times G)$, a sequence $k_n$ in $C(G \times G)$ exists with $\lim_{n\to\infty} \|k_n - k\|_2 \to 0$. Due to Theorem 2.10, the integral operators $K_n$ induced by the kernels $k_n$ are compact from $L^2(G)$ to $L^2(G)$. One can see as in the proof of Theorem 2.1 that $\|K_n - K\|_{L(L^2(G))} \leq \|k_n - k\|_2$. Therefore, $\lim_{n\to\infty} \|K_n - K\|_{L(L^2(G))} = 0$. Thus, Theorem 2.8 (e) implies the compactness of $K : L^2(G) \to L^2(G)$. ∎

When dealing with boundary value problems for elliptic differential equations, integral operators with kernels having singularities play an important role. It turns out that the quality of the singularity decides about the compactness of the corresponding integral operator. We distinguish between weakly singular and strongly singular integral operators. The distinction depends on the dimension $N$ and is done such that weakly singular in contrast to strongly singular operators are still compact.

We only consider the very important case where the singularities lie on the diagonal $\{(s, s) : s \in G\}$:

**Definition 2.12.** Let $G \subseteq \mathbb{R}^N$ be compact and Jordan measurable with positive measure. The kernel $k : (G \times G) \setminus \{(s, s) : s \in G\} \to \mathbb{R}$ is called **weakly singular** if it is continuous and if $M > 0$ and $\alpha > 0$ exist such that

$$|k(s, t)| \leq M|s - t|^{\alpha - N} \,, \qquad s \neq t \in G \,. \tag{2.17}$$

Here $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^N$.

**Theorem 2.13.** Let $k$ be a weakly singular kernel and let $K$ be the corresponding integral operator according to (2.2). Then $K$ is well-defined in $C(G)$ (considering the integral as an improper one) and compact from $C(G)$ to $C(G)$.

**Proof:** Let $x \in C(G)$ and $s \in G$. Then it follows with (2.17) that

$$\begin{aligned}\int_G |k(s,t)x(t)|\,dt &\leq M\,\|x\|_\infty \int_G |s-t|^{\alpha-N}\,dt \\ &\leq M\,\|x\|_\infty \int_{\{\tau:|\tau|\leq \operatorname{diam}(G)\}} |\tau|^{\alpha-N} d\tau \\ &= CM\,\|x\|_\infty \int_0^{\operatorname{diam}(G)} r^{\alpha-N} r^{N-1}\,dr \;=\; CM\,\|x\|_\infty \frac{\operatorname{diam}(G)^\alpha}{\alpha}\,.\end{aligned}$$

In the last integral we switched to $N$-dimensional spherical coordinates, i.e.,

$$\begin{aligned} g : \mathbb{R}_0^+ \times [0,\pi]^{N-2} \times [0,2\pi] &\to \mathbb{R}^N \\ g_1(r,\varphi_1,\ldots,\varphi_{N-1}) &:= r\cos(\varphi_1)\,, \\ g_i(r,\varphi_1,\ldots,\varphi_{N-1}) &:= r\left(\prod_{j=1}^{i-1}\sin(\varphi_j)\right)\cos(\varphi_i)\,, \qquad i=2,\ldots,N-1\,, \\ g_N(r,\varphi_1,\ldots,\varphi_{N-1}) &:= r\prod_{j=1}^{N-1}\sin(\varphi_j) \end{aligned}$$

The determinant of the transformation matrix is given by

$$|g'(r,\varphi_1,\ldots,\varphi_{N-1})| = r^{N-1}\prod_{i=1}^{N-2}\sin^{N-1-i}(\varphi_i)\,.$$

The integration with respect to $\varphi_i$, $i=1,\ldots,N-1$ yields the constant

$$C := \frac{2\pi^{\frac{N}{2}}}{\Gamma\left(\frac{N}{2}\right)}\,.$$

Thus, the improper integral $\int_G k(s,t)x(t)\,dt$ is absolutely convergent for all $s \in G$ and, hence, the integral operator $K$ is well-defined in $L^\infty(G)$.

We will now approximate $K$ by a sequence of integral operators $(K_n)$ with continuous kernels: for $n \in \mathbb{N}$ let

$$k_n(s,t) := \begin{cases} \rho_n(|s-t|)k(s,t)\,, & s \neq t\,, \\ 0\,, & s = t\,, \end{cases} \qquad \text{with} \quad \rho_n(r) := \begin{cases} 0\,, & r \in \left[0, \frac{1}{2n}\right]\,, \\ 2nr-1 & r \in \left[\frac{1}{2n}, \frac{1}{n}\right]\,, \\ 1\,, & r \geq \frac{1}{n}\,, \end{cases}$$

and let $K_n$ be the integral operator according to (2.2) with the continuous kernel $k_n$. Due to Theorem 2.10, $K_n$ is a compact linear operator from $C(G)$ to $C(G)$ for all $n \in \mathbb{N}$.

For all $x \in C(G)$, $n \in \mathbb{N}$, and $s \in G$ we have the estimate:

$$\begin{aligned}|(Kx)(s) - (K_n x)(s)| &\leq \|x\|_\infty \int_G |k(s,t) - \rho_n(|s-t|)k(s,t)|\,dt \\ &= \|x\|_\infty \int_{\{t\in G:|t-s|\leq\frac{1}{n}\}} |k(s,t)|(1-\rho_n(|s-t|))\,dt \\ &\leq \|x\|_\infty \int_{\{t\in G:|t-s|\leq\frac{1}{n}\}} |k(s,t)|\,dt\end{aligned}$$

$$\begin{aligned} &\leq \|x\|_\infty M \int_{\{\tau:|\tau|\leq\frac{1}{n}\}} |\tau|^{\alpha-N} d\tau \\ &= CM\|x\|_\infty \int_0^{\frac{1}{n}} r^{\alpha-N} r^{N-1}\, dr \;=\; CM\|x\|_\infty \frac{1}{\alpha n^\alpha}\,, \end{aligned}$$

where we again switched to $N$-dimensional spherical coordinates in the last integral.

This implies that $(K_n x)$ converges uniformly (in $s$) towards $Kx$. Thus, $Kx \in C(G)$, since for any $s, \bar{s} \in G$

$$|Kx(s) - Kx(\bar{s})| \leq 2CM\|x\|_\infty \frac{1}{\alpha n^\alpha} + |K_n x(s) - K_n x(\bar{s})|\,.$$

The convergence of $(K_n x)$ is also uniform with respect to $x$ in bounded subsets of $C(G)$ so that $\|K - K_n\| \to 0$. This, due to Theorem 2.8 (e), yields the compactness of $K$. ∎

The integral operator in Abel's integral equation (1.12) can be written in the form (2.2) with kernel

$$k(s,t) = \begin{cases} (s-t)^{-\frac{1}{2}}\,, & t \in [0,s)\,, \\ 0\,, & t \geq s\,. \end{cases}$$

Since $N = 1$ this kernel satisfies (2.17) with $\alpha = \frac{1}{2}$ and $M = 1$. Due to Definition 2.12, the kernel is weakly singular. Other typical examples of weakly singular integral operators are

$$(Kx)(s) := \int_0^1 \ln|s-t| x(t)\, dt \qquad \text{or} \qquad (Kx)(s) := \int_G \frac{x(t)}{|s-t|^{N-1}}\, dt\,, \quad G \subseteq \mathbb{R}^N\,.$$

We will now derive properties of compact operators that are important for the solution of Fredholm integral equations of the second kind. We start with results that do not need the adjoint equation (compare Theorem 2.6).

## 2.3. Riesz theory

For the time being, we assume that $K : X \to X$ is compact, $X$ is a normed space, and

$$L := I - K\,. \tag{2.18}$$

The Riesz theory includes statements about such operators $L$, i.e., compact perturbations of the identity. For the proof of the three Riesz theorems we need some preparatory lemmata.

**Lemma 2.14 (Riesz).** Let $X$ be a normed space, $Y \subset X$ a closed subspace, and $\varepsilon \in (0,1)$. Then there exists an $x \in X$ with $\|x\| = 1$ so that $\|y - x\| \geq \varepsilon$ for all $y \in Y$.

**Proof:** Let $z \in X \setminus Y$. Since $Y$ is closed, $\alpha := \inf\{\|z-y\| : y \in Y\} > 0$. Since $\frac{\alpha}{\varepsilon} > \alpha$, we may choose a $w \in Y$ with $\alpha \le \|z-w\| \le \frac{\alpha}{\varepsilon}$. Let $x := \frac{z-w}{\|z-w\|}$. Then $\|x\| = 1$ and

$$\begin{aligned} \|y-x\| &= \frac{1}{\|z-w\|}\,\|z-w-\|z-w\|\,y\| \\ &= \frac{1}{\|z-w\|}\,\|z-(w+\|z-w\|\,y)\| \\ &\ge \frac{\alpha}{\|z-w\|} \ge \frac{\alpha}{(\frac{\alpha}{\varepsilon})} = \varepsilon \end{aligned}$$

for all $y \in Y$. Note that $w + \|z-w\|\,y \in Y$. ∎

This lemma immediately implies the following interesting result:

**Lemma 2.15.** The identity operator $I : X \to X$ is compact if and only if $X$ is a finite-dimensional normed space.

**Proof:** If $\dim X < \infty$, then the compactness of $I$ follows from Theorem 2.8 (f).

To prove the converse we assume that $\dim X = \infty$. We inductively define a sequence $(x_n)$ in $X$ and a sequence $(Y_n)$ of closed subspaces as follows: choose $x_1 \in X$ with $\|x_1\| = 1$ and $Y_1 := \operatorname{span}\{x_1\}$.

Let us assume that $Y_n := \operatorname{span}\{x_1, \dots, x_n\}$ is already defined. Since $Y_n$ is a proper closed subspace of $X$, Lemma 2.14 (with $\varepsilon = \frac{1}{2}$) implies that there exists an $x_{n+1} \in X$ with $\|x_{n+1}\| = 1$ and $\|x_{n+1} - x_m\| \ge \frac{1}{2}$ for all $m \le n$. We then define the finite-dimensional subspace $Y_{n+1} := \operatorname{span}\{x_1, \dots, x_n, x_{n+1}\}$.

The sequence $(x_n)$ lies in the bounded unit ball $\{x \in X : \|x\| \le 1\}$ and has, due to its construction, no convergent subsequence. Therefore, the unit ball is not compact and, hence, $I$ is not compact. ∎

**Theorem 2.16.** Let $X$ be a normed space and let $K : X \to X$ be compact and $L$ as in (2.18). Then the following assertions hold:

(a) **(1. Riesz theorem)** $\dim \mathcal{N}(L) < \infty$.

(b) **(2. Riesz theorem)** $\mathcal{R}(L)$ is closed.

**Proof:**

(a) Since $L$ is continuous, $\mathcal{N}(L) = L^{-1}(\{0\})$ is closed. Let $\widetilde{K} := K|_{\mathcal{N}(L)}$. It follows from the definition of compact operators that the restriction of a compact operator to a subspace is also compact. Thus, $\widetilde{K}$ is compact. If $x \in \mathcal{N}(L)$, then $\widetilde{K}x = x - Lx = x$. Therefore, $\widetilde{K}$ is the identity operator on $\mathcal{N}(L)$. Now Lemma 2.15 implies that $\dim \mathcal{N}(L) < \infty$.

(b) Let $y \in \overline{\mathcal{R}(L)}$ be arbitrary but fixed. Then there exists a sequence $(\bar{x}_n)$ in $X$ with

$$L\bar{x}_n \to y. \tag{2.19}$$

For all $n \in \mathbb{N}$ let $z_n \in \mathcal{N}(L)$ be such that

$$\|\bar{x}_n - z_n\| \leq d(\bar{x}_n, \mathcal{N}(L)) + \frac{1}{n}\,, \tag{2.20}$$

where $d(\bar{x}_n, \mathcal{N}(L))$ denotes the minimal distance of $\bar{x}_n$ to $\mathcal{N}(L)$, and let

$$x_n := \bar{x}_n - z_n\,. \tag{2.21}$$

We show that

$$(x_n) \quad \text{is bounded} \tag{2.22}$$

by assuming the contrary. Then a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ exists with

$$\|x_{n_k}\| \geq k\,, \qquad k \in \mathbb{N}\,. \tag{2.23}$$

Setting $v_k := \frac{x_{n_k}}{\|x_{n_k}\|}$, we have that $(v_k)$ is bounded. Since $K$ is compact, $(Kv_k)$ has a convergent subsequence

$$Kv_{k_j} \to z \in X. \tag{2.24}$$

Since, due to (2.19), $(L\bar{x}_n)$ is bounded and since $Lx_n = L\bar{x}_n$, (2.23) implies that

$$\|Lv_k\| = \frac{1}{\|x_{n_k}\|}\,\|L\bar{x}_{n_k}\| \to 0 \quad \text{as} \quad k \to \infty$$

Together with (2.24) this implies that

$$v_{k_j} = Lv_{k_j} + Kv_{k_j} \to z\,. \tag{2.25}$$

Since $L$ is continuous, this yields that

$$Lz = \lim_{j\to\infty} Lv_{k_j} = 0$$

and, hence, that $z \in \mathcal{N}(L)$. But then also

$$z_{n_k} + \|x_{n_k}\|\, z \in \mathcal{N}(L)\,, \qquad k \in \mathbb{N}\,.$$

This together with (2.20) and (2.21) yields that

$$\begin{aligned}
\|v_k - z\| &= \frac{1}{\|x_{n_k}\|}\,\|x_{n_k} - \|x_{n_k}\|\, z\| \\
&= \frac{1}{\|x_{n_k}\|}\,\|\bar{x}_{n_k} - (z_{n_k} + \|x_{n_k}\|\, z)\| \;\geq\; \frac{1}{\|x_{n_k}\|} d(\bar{x}_{n_k}, \mathcal{N}(L)) \\
&\geq \frac{1}{\|x_{n_k}\|}\left(\|\bar{x}_{n_k} - z_{n_k}\| - \frac{1}{n_k}\right) = 1 - \frac{1}{n_k\,\|x_{n_k}\|} \underset{k\to\infty}{\longrightarrow} 1\,,
\end{aligned}$$

which is a contradiction to (2.25). Therefore, a subsequence with (2.23) cannot exist, i.e., (2.22) holds.

Since $K$ is compact, the sequence $(Kx_n)$ has a convergent subsequence $(Kx_{n_k})$. Due to $x_{n_k} = Lx_{n_k} + Kx_{n_k} = L\bar{x}_{n_k} + Kx_{n_k}$, it follows together with (2.19) that $(x_{n_k})$ converges to some $x \in X$. Since $L$ is continuous, we now obtain that

$$Lx = \lim_{k\to\infty} Lx_{n_k} = \lim_{k\to\infty} L\bar{x}_{n_k} = y\,.$$

Thus, $y \in \mathcal{R}(L)$. Since $y \in \overline{\mathcal{R}(L)}$ was arbitrary, this implies that $\mathcal{R}(L)$ is closed.

∎

Theorem 2.16 implies that the solution set of Fredholm integral equations of the second kind is finite-dimensional. Moreover, if one approximates the right-hand side $y$ by a convergent sequence $(y_n)$ so that the equations with $y_n$ are solvable for all $n \in \mathbb{N}$, then the original problem is also solvable.

For integral equations of the first kind these properties do not hold as the following theorem suggests:

**Theorem 2.17.** Let $X, Y$ be Banach spaces and let $K : X \to Y$ be compact and such that a closed subspace $Z \subseteq X$ exists with $X = \mathcal{N}(K) \oplus Z$. Then $\mathcal{R}(K)$ is closed if and only if $\mathcal{R}(K)$ is finite-dimensional. In that case also $Z$ is finite-dimensional.

**Proof:** If $\mathcal{R}(K)$ is finite-dimensional, then $\mathcal{R}(K)$ is closed.

Let us now assume that $\mathcal{R}(K)$ is closed and let $\widetilde{K} := K|_Z : Z \to \mathcal{R}(K)$. Then $\mathcal{R}(K)$ is a Banach space and $\widetilde{K}$ is a bijective linear operator. Due to the open mapping theorem $\widetilde{K}^{-1}$ is bounded. Therefore, Theorem 2.8 (d) implies that $\widetilde{K}\widetilde{K}^{-1} = I_{\mathcal{R}(K)}$ is compact. Now Lemma 2.15 implies that $\mathcal{R}(K)$ is finite-dimensional. Since $\widetilde{K}^{-1}\widetilde{K} = I_Z$, this also holds for $Z$. ∎

**Remark 2.18.** The condition on the existence of $Z$, i.e., the existence of a topological complement of $\mathcal{N}(K)$, is always satisfied if $X$ is a Hilbert space or if $\mathcal{N}(K)$ is finite-dimensional. Therefore, Theorem 2.17 is applicable to integral equations of the first kind in $L^2$ and yields the following result:

If the kernel is non-degenerate so that $\mathcal{R}(K)$ is infinite-dimensional (see Theorem 2.3), then it is not closed. Therefore, the equation is not solvable for all right-hand sides. Even if one has a convergent sequence of right-hand sides where the equation is always solvable, the limit problem need not be solvable.

If $\mathcal{N}(K) = \{0\}$ so that $K^{-1}$ exists on $\mathcal{R}(K)$, the operator $K^{-1}$ is unbounded. Therefore, the solutions of the integral equation do not depend continuously on the right-hand sides, which obviously leads to problems if one wants to solve such equations numerically.

For the solution theory, it would be advantageous if it were possible to write $X$ as a direct sum of $\mathcal{N}(L)$ and $\mathcal{R}(L)$. But this need not be possible even in finite-dimensional spaces: if, for instance,

$$L := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

then $\mathcal{N}(L) = \{0\} \times \mathbb{R}$ and $\mathcal{R}(L) = \{0\} \times \mathbb{R} = \mathcal{N}(L)$. Thus $\mathcal{N}(L) \oplus \mathcal{R}(L) \neq \mathbb{R}^2$. However,

$$L^2 := \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

so that $\mathcal{N}(L^2) = \mathbb{R}^2$, $\mathcal{R}(L^2) = \{0\}$, and $\mathcal{N}(L^2) \oplus \mathcal{R}(L^2) = \mathbb{R}^2$.

We will see that a similar decomposition is always possible for operators $L = I - K$ with compact $K$. We will show that $\nu \in \mathbb{N}_0$ exists with $X = \mathcal{N}(L^\nu) \oplus \mathcal{R}(L^\nu)$. This fact is

part of the so-called **Ries-Schauder theory** which is essential for existence, uniqueness, and stability results for integral equations of the second kind.

**Theorem 2.19 (3. Riesz theorem).** Let $X$ be a normed space and let $K : X \to X$ be compact and $L$ as in (2.18). Then there exists a number $\nu \in \mathbb{N}_0$, called **Riesz index** of $L$ (or $K$), so that for all $\mu \in \mathbb{N}_0$

$$\mathcal{N}(L^\mu) \subsetneq \mathcal{N}(L^{\mu+1}) \quad \text{if} \quad \mu < \nu\,, \quad \text{and} \quad \mathcal{N}(L^\mu) = \mathcal{N}(L^{\mu+1}) \quad \text{if} \quad \mu \geq \nu\,, \tag{2.26}$$

$$\mathcal{R}(L^\mu) \supsetneq \mathcal{R}(L^{\mu+1}) \quad \text{if} \quad \mu < \nu\,, \quad \text{and} \quad \mathcal{R}(L^\mu) = \mathcal{R}(L^{\mu+1}) \quad \text{if} \quad \mu \geq \nu\,. \tag{2.27}$$

Note that $\nu = 0$ is possible.

Moreover, it holds that $\dim \mathcal{N}(L^\nu) < \infty$ and that

$$X = \mathcal{N}(L^\nu) \oplus \mathcal{R}(L^\nu)\,. \tag{2.28}$$

The projectors onto $\mathcal{N}(L^\nu)$ and $\mathcal{R}(L^\nu)$, induced by the decomposition (2.28), are continuous. $L$ maps the closed subspace $\mathcal{R}(L^\nu)$ bijectively onto itself and has a continuous inverse as restriction to $\mathcal{R}(L^\nu)$.

**Proof:** Since the proof is rather long, we split it into several steps.

(i) First, we show the existence of $\nu_1 \in \mathbb{N}_0$ so that (2.26) holds with $\nu = \nu_1$.

Obviously, $\mathcal{N}(L^k) \subseteq \mathcal{N}(L^{k+1})$ for all $k \in \mathbb{N}$. If $\mathcal{N}(L^k) = \mathcal{N}(L^{k+1})$, then also $\mathcal{N}(L^{p+1}) = \mathcal{N}(L^p)$ for all $p > k$. This can be seen as follows:

$$\begin{aligned} x \in \mathcal{N}(L^{p+1}) &\implies 0 = L^{p+1}x = L^{k+1}(L^{p-k}x) \implies L^{p-k}x \in \mathcal{N}(L^{k+1}) = \mathcal{N}(L^k) \\ &\implies 0 = L^k(L^{p-k}x) = L^p x \implies x \in \mathcal{N}(L^p) \end{aligned}$$

Thus, $\mathcal{N}(L^{p+1}) \subseteq \mathcal{N}(L^p)$. Since the other inclusion is always correct, we have shown that $\mathcal{N}(L^p) = \mathcal{N}(L^{p+1})$.

Therefore, we only have to show that a $\nu_1 \in \mathbb{N}_0$ exists with $\mathcal{N}(L^{\nu_1}) = \mathcal{N}(L^{\nu_1+1})$. Thereto, we assume the contrary, i.e., $\mathcal{N}(L^k) \subsetneq \mathcal{N}(L^{k+1})$ for all $k \in \mathbb{N}_0$. But then Lemma 2.14 implies that for all $k \in \mathbb{N}_0$ an $x_k \in \mathcal{N}(L^{k+1})$ exists with $\|x_k\| = 1$ and $\|x_k - x\| \geq \frac{1}{2}$ for all $x \in \mathcal{N}(L^k)$. Note that Lemma 2.14 is applicable, since $\mathcal{N}(L^k)$ is closed. Let us consider this sequence $(x_k)$: for $n > m$ we have that

$$L^n(x_m + Lx_n - Lx_m) = L^{n-m-1}L^{m+1}x_m + L^{n+1}x_n - L^{n-m}L^{m+1}x_m = 0$$

and, therefore, $x_m + Lx_n - Lx_m \in \mathcal{N}(L^n)$ and $\|x_n - (x_m + Lx_n - Lx_m)\| \geq \frac{1}{2}$, since this inequality holds for all elements in $\mathcal{N}(L^n)$. Now

$$x_n - (x_m + Lx_n - Lx_m) = Kx_n - Kx_m$$

implies that $\|Kx_n - Kx_m\| \geq \frac{1}{2}$ for all $n > m$ so that $(Kx_n)$ cannot have a convergent subsequence. Since $(x_n)$ is bounded, this is a contradiction to the compactness of $K$. Thus, our assumption was wrong and, therefore, $\nu_1 \in \mathbb{N}_0$ exists.

(ii) Now we show the existence of $\nu_2 \in \mathbb{N}_0$ so that (2.27) holds with $\nu = \nu_2$.

Obviously, $\mathcal{R}(L^k) \supseteq \mathcal{R}(L^{k+1})$ for all $k \in \mathbb{N}$. If $\mathcal{R}(L^k) = \mathcal{R}(L^{k+1})$, then also $\mathcal{R}(L^{p+1}) = \mathcal{R}(L^p)$ for all $p > k$. This can be seen as follows:

$$\begin{aligned} x \in \mathcal{R}(L^p) &\implies x = L^p z = L^{p-k}(L^k z) \ \wedge \ L^k z \in \mathcal{R}(L^k) = \mathcal{R}(L^{k+1}) \\ &\implies L^k z = L^{k+1} y \ \wedge \ x = L^{p-k}(L^{k+1} y) = L^{p+1} y \in \mathcal{R}(L^{p+1}) \end{aligned}$$

Thus, $\mathcal{R}(L^p) \subseteq \mathcal{R}(L^{p+1})$. Since the other conclusion is always correct, we have shown that $\mathcal{R}(L^p) = \mathcal{R}(L^{p+1})$.

Therefore, we only have to show that a $\nu_2 \in \mathbb{N}_0$ exists with $\mathcal{R}(L^{\nu_2}) = \mathcal{R}(L^{\nu_2+1})$. Thereto, we assume the conntrary, i.e., $\mathcal{R}(L^k) \supsetneq \mathcal{R}(L^{k+1})$ for all $k \in \mathbb{N}_0$. Since, due to the binomial theorem,

$$L^n = (I - K)^n = I - K_n \qquad \text{with} \qquad K_n = \sum_{i=1}^{n} (-1)^{i-1} \binom{n}{i} K^i \tag{2.29}$$

for all $n \in \mathbb{N}$ and since products and sums of compact operators are again compact (see Theorem 2.8 (c), (d)), $K_n$ is compact. Thus, Theorem 2.16 is also applicable to $L^n$, especially $\mathcal{R}(L^n)$ is closed. Now Lemma 2.14 implies that for all $k \in \mathbb{N}_0$ a $y_k \in \mathcal{R}(L^k)$ exists with $\|y_k\| = 1$ and $\|y_k - y\| \geq \frac{1}{2}$ for all $y \in \mathcal{R}(L^{k+1})$. Let us consider this sequence $(y_k)$: since $y_k \in \mathcal{R}(L^k)$, $x_k$ exists with $L^k x_k = y_k$. But then

$$y_m + L y_n - L y_m = L^{n+1}(L^{m-n-1} x_m + x_n - L^{m-n} x_m) \in \mathcal{R}(L^{n+1}) \quad \text{for} \quad m > n \in \mathbb{N}$$

so that $\frac{1}{2} \leq \|y_n - (y_m + L y_n - L y_m)\| = \|K y_n - K y_m\|$ for $m > n$, which is a contradiction to the compactness of $K$. Thus, our assumption was wrong and, therefore, $\nu_2 \in \mathbb{N}_0$ exists.

(iii) Let $\nu_1$ and $\nu_2$ be minimal with the properties

$$\mathcal{N}(L^{\nu_1}) = \mathcal{N}(L^{\nu_1+1}) \qquad \text{and} \qquad \mathcal{R}(L^{\nu_2}) = \mathcal{R}(L^{\nu_2+1}) \,.$$

We still have to show that

$$\nu_1 = \nu_2 \,. \tag{2.30}$$

First we assume that $\nu_1 > \nu_2$. Then, due to the definition of $\nu_2$,

$$\mathcal{R}(L^{\nu_1}) = \mathcal{R}(L^{\nu_1-1}) = \ldots = \mathcal{R}(L^{\nu_2}) \,.$$

Thus, $L^{\nu_1-1} x \in \mathcal{R}(L^{\nu_1-1}) = \mathcal{R}(L^{\nu_1})$ and, hence, $\bar{x}$ exists with $L^{\nu_1-1} x = L^{\nu_1} \bar{x}$. Now, due to the definition of $\nu_1$,

$$\begin{aligned} x \in \mathcal{N}(L^{\nu_1}) &\implies 0 = L^{\nu_1} x = L^{\nu_1+1} \bar{x} \implies \bar{x} \in \mathcal{N}(L^{\nu_1+1}) = \mathcal{N}(L^{\nu_1}) \\ &\implies 0 = L^{\nu_1} \bar{x} = L^{\nu_1-1} x \implies x \in \mathcal{N}(L^{\nu_1-1}) \,. \end{aligned}$$

Thus, $\mathcal{N}(L^{\nu_1}) \subseteq \mathcal{N}(L^{\nu_1-1})$ holds which is a contradiction to the minimality of $\nu_1$. Therefore, $\nu_1 \leq \nu_2$.

Now we assume that $\nu_1 < \nu_2$. Then, due to the definition of $\nu_1$,

$$\mathcal{N}(L^{\nu_1}) = \mathcal{N}(L^{\nu_1+1}) = \ldots = \mathcal{N}(L^{\nu_2-1}) = \mathcal{N}(L^{\nu_2}) \,.$$

Let $y = L^{\nu_2-1}x \in \mathcal{R}(L^{\nu_2-1})$. Then $Ly = L^{\nu_2}x \in \mathcal{R}(L^{\nu_2}) = \mathcal{R}(L^{\nu_2+1})$ and, hence, $\bar{x}$ exists with $Ly = L^{\nu_2+1}\bar{x}$. Thus,

$$\begin{aligned} L^{\nu_2}(x - L\bar{x}) = Ly - Ly = 0 &\implies x - L\bar{x} \in \mathcal{N}(L^{\nu_2}) = \mathcal{N}(L^{\nu_2-1}) \\ &\implies 0 = L^{\nu_2-1}(x - L\bar{x}) = L^{\nu_2-1}x - L^{\nu_2}\bar{x} \\ &\implies y = L^{\nu_2-1}x = L^{\nu_2}\bar{x} \in \mathcal{R}(L^{\nu_2}) \end{aligned}$$

showing that $\mathcal{R}(L^{\nu_2-1}) \subseteq \mathcal{R}(L^{\nu_2})$ holds which is a contradiction to the minimality of $\nu_2$. This proves (2.30), i.e., the existence of the Ries index $\nu$.

(iv) Since, due to (2.29), $L^\nu = I - K_\nu$, Theorem 2.16 (a) implies that $\dim \mathcal{N}(L^\nu) < \infty$.

(v) In a next step, we prove (2.28): let $x \in \mathcal{N}(L^\nu) \cap \mathcal{R}(L^\nu)$. Then $x = L^\nu \bar{x}$ for some $\bar{x}$ and $L^\nu x = 0$. But then $L^{2\nu}\bar{x} = 0$, i.e., $\bar{x} \in \mathcal{N}(L^{2\nu}) = \mathcal{N}(L^\nu)$ and, hence, $x = L^\nu \bar{x} = 0$. Therefore, we have shown that

$$\mathcal{N}(L^\nu) \cap \mathcal{R}(L^\nu) = \{0\}\,. \tag{2.31}$$

Now we show that

$$X = \mathcal{N}(L^\nu) + \mathcal{R}(L^\nu), \tag{2.32}$$

which together with (2.31) implies (2.28). Thereto, let $x \in X$ be arbitrary, but fixed. We have to construct elements in $\mathcal{N}(L^\nu)$ and $\mathcal{R}(L^\nu)$ so that their sum equals $x$: Since $L^\nu x \in \mathcal{R}(L^\nu) = \mathcal{R}(L^{2\nu})$, $\bar{x}$ exists with $L^\nu x = L^{2\nu}\bar{x}$. Let $n := x - L^\nu \bar{x}$. Then

$$L^\nu n = L^\nu x - L^{2\nu}\bar{x} = 0 \implies n \in \mathcal{N}(L^\nu)\,.$$

Since $x = n + L^\nu \bar{x}$, it holds that $x \in \mathcal{N}(L^\nu) + \mathcal{R}(L^\nu)$. This proves (2.32).

(vi) Now we prove the continuity of the projection operators induced by the decomposition (2.28): for $n \in \mathcal{N}(L^\nu)$ let

$$|n| := \inf\{\|n + r\| : r \in \mathcal{R}(L^\nu)\} = \inf\{\|n - r'\| : r' \in \mathcal{R}(L^\nu)\} = d(n, \mathcal{R}(L^\nu))\,.$$

We show that $|\cdot|$ is a norm on $\mathcal{N}(L^\nu)$: since $\mathcal{R}(L^\nu)$ is closed, $|n| > 0$ if $n \notin \mathcal{R}(L^\nu)$. Therefore, $|n| = 0$ implies that $n \in \mathcal{R}(L^\nu)$. But then (2.31) implies that $n = 0$.

If $\lambda \in \mathbb{R} \setminus \{0\}$, then

$$\begin{aligned} |\lambda n| &= \inf\{\|\lambda n + r\| : r \in \mathcal{R}(L^\nu)\} \\ &= \inf\{|\lambda| \left\|n + \frac{r}{\lambda}\right\| : r \in \mathcal{R}(L^\nu)\} \\ &= |\lambda| \inf\{\|n + r'\| : r' \in \mathcal{R}(L^\nu)\} = |\lambda|\,|n|\,. \end{aligned}$$

Obviously, $\|n + m + r + s\| \le \|n + r\| + \|m + s\|$ for all $r, s \in \mathcal{R}(L^\nu)$. Taking the infimum of this inequality with respect to $r$ and $s$ yields $|n + m| \le |n| + |m|$ for all $n, m \in \mathcal{N}(L^\nu)$. Altogether we have shown that $|\cdot|$ is a norm on $\mathcal{N}(L^\nu)$.

Another norm on $\mathcal{N}(L^\nu)$ is the restriction of the original norm. Since $\mathcal{N}(L^\nu)$ is finite-dimensional and since all norms on finite-dimensional spaces are equivalent, there must exist a constant $C > 0$ such that $\|n\| \le C|n|$ for all $n \in \mathcal{N}(L^\nu)$.

This implies that

$$\|Px\| \le C|Px| = C \inf\{\|Px + r\| : r \in \mathcal{R}(L^\nu)\} \le C\,\|Px + (x - Px)\| = C\,\|x\|$$

for all $x \in X$, where $P : X \to \mathcal{N}(L^\nu)$ denotes the projection onto $\mathcal{N}(L^\nu)$. Note that $x - Px \in \mathcal{R}(L^\nu)$. Thus, $P$ is bounded and, hence, continuous. The same is obviously true for $I - P$, the projector onto $\mathcal{R}(L^\nu)$.

(vii) In a final step, we prove that $\bar{L} := L|_{\mathcal{R}(L^\nu)}$ is a continuously invertible bijection onto $\mathcal{R}(L^\nu)$. Since, due to (2.31),

$$\begin{aligned} \mathcal{R}(\bar{L}) &= L(\mathcal{R}(L^\nu)) = \mathcal{R}(L^{\nu+1}) = \mathcal{R}(L^\nu) \\ \mathcal{N}(\bar{L}) &= \mathcal{N}(L) \cap \mathcal{R}(L^\nu) \subseteq \mathcal{N}(L^\nu) \cap \mathcal{R}(L^\nu) = \{0\} \end{aligned}$$

$\bar{L} : \mathcal{R}(L^\nu) \to \mathcal{R}(L^\nu)$ is bijective. Hence, the inverse $\bar{L}^{-1}$ exists. We assume that $\bar{L}^{-1}$ is unbounded. Then a sequence $(x_n)$ exists in $\mathcal{R}(L^\nu)$ with $\|x_n\| = 1$ for all $n \in \mathbb{N}$ and $\left\|\bar{L}^{-1}x_n\right\| \to \infty$ as $n \to \infty$. Let $\bar{x}_n := \frac{x_n}{\left\|\bar{L}^{-1}x_n\right\|}$. Then $\bar{x}_n \to 0$. Since $\left\|\bar{L}^{-1}\bar{x}_n\right\| = 1$ and since $K|_{\mathcal{R}(L^\nu)}$ as a restriction of a compact operator is also compact, $(K(\bar{L}^{-1}\bar{x}_n))$ has a convergent subsequence $(K(\bar{L}^{-1}\bar{x}_{n_k}))$, whose limit is denoted by $z$. Since $\mathcal{R}(L^\nu)$ is closed, $z \in \mathcal{R}(L^\nu)$. Now

$$\bar{L}^{-1}\bar{x}_n - K(\bar{L}^{-1}\bar{x}_n) = \bar{L}\bar{L}^{-1}\bar{x}_n = \bar{x}_n$$

implies that

$$\lim_{k\to\infty} \bar{L}^{-1}\bar{x}_{n_k} = \lim_{k\to\infty} \left(\bar{x}_{n_k} + K(\bar{L}^{-1}\bar{x}_{n_k})\right) = z\,.$$

Thus, $\bar{L}z = \lim_{k\to\infty} \bar{L}(\bar{L}^{-1}\bar{x}_{n_k}) = \lim_{k\to\infty} \bar{x}_{n_k} = 0$ and, hence, $z = 0$, which is a contradiction to $\|z\| = \lim_{k\to\infty} \left\|\bar{L}^{-1}\bar{x}_{n_k}\right\| = 1$. Therefore, the assumption of $\bar{L}^{-1}$ being unbounded was wrong, i.e., $\bar{L}^{-1}$ is bounded.

∎

**Remark 2.20.** In the last part of the proof above, the following principle was used: construct a bounded sequence $z_n$ (here $\bar{L}^{-1}\bar{x}_n$). Using the compactness a convergent subsequence $(Kz_{n_k})$ of $(Kz_n)$ is chosen. If $w_n := z_n - Kz_n$ is convergent, then

$$z_{n_k} = w_{n_k} + Kz_{n_k}$$

is also convergent. This conclusion from the convergence of the image sequence $(Kz_{n_k})$ to the preimage sequence $(z_{n_k})$ is only possible for equations of the second kind since the unknown function also appears outside the compact operator $K$.

By the way, the last step of the proof would have been simpler if $X$ had been complete: the boundedness of $\bar{L}^{-1}$ would have then been a consequence of the open mapping theorem by Banach.

All results from Theorem 2.19 remain valid if $L = S - K$ with compact $K$ and a continuously invertible $S$, since

$$Sx - Kx = f \quad \Longleftrightarrow \quad x - S^{-1}Kx = S^{-1}f$$

and since $S^{-1}K$ is compact, due to Theorem 2.8 (d).

We can draw the following conclusions from the theorems of Riesz for the equation

$$x - Kx = f \tag{2.33}$$

and its homogeneous version

$$x - Kx = 0\,, \tag{2.34}$$

when $K : X \to X$ is compact, including Fredholm integral equations of the second kind:

If (2.34) only has the trivial solution $x = 0$, then (2.33) is uniquely solvable for all $f \in X$ and this solution depends continuously on $f$. One then says that problem (2.33) is correctly posed. This can be seen from Theorem 2.19 as follows: since $\mathcal{N}(I-K) = \{0\}$, the Riesz index $\nu = 0$, i.e., $\mathcal{R}(I-K) = \mathcal{R}((I-K)^0) = X$. Therefore, $I-K$ is surjective. Moreover, $L|_{\mathcal{R}(L^\nu)} = L$ and, hence, $L^{-1}$ is continuous.

Equation (2.34) has nontrivial solutions if and only if (2.33) is not solvable for all $f \in X$. If one only allows right-hand sides $f \in \mathcal{R}((I-K)^\nu)$ and solutions $x \in \mathcal{R}((I-K)^\nu)$, then (2.33) is uniquely solvable and the solution depends continuously on $f$. This is especially important for the case $\nu = 1$ (see Theorem 2.29 below), since the set of admissible right-hand sides then coincides with the set of right-hand sides, where a solution exists at all. The considered solutions are then the ones without a contribution from the null space.

The solution sets of (2.33) and (2.34) are always finite-dimensional.

## 2.4. Fredholm theory

In cases $\nu \geq 1$, the Riesz theory gives no information for which concrete right-hand sides $f$ equation (2.33) is solvable. Thereto, one needs an adjoint equation as in Theorem 2.6. In Hilbert spaces this is quite obvious. In Banach spaces, one could use the adjoint operator defined on the dual space. However, such dual spaces can be quite ugly, e.g., the dual space of continuous functions is the space of functions with bounded variation. Therefore, it is more convenient to derive the Fredholm theory for so-called dual systems:

**Definition 2.21.** Let $X, Y$ be normed spaces. A mapping $\langle \cdot, \cdot \rangle : X \times Y \to \mathbb{C}$ (or $\mathbb{R}$) is called **bilinear form** if

$$\begin{aligned} \langle \alpha_1 x_1 + \alpha_2 x_2, y \rangle &= \alpha_1 \langle x_1, y \rangle + \alpha_2 \langle x_2, y \rangle \\ \langle x, \beta_1 y_1 + \beta_2 y_2 \rangle &= \beta_1 \langle x, y_1 \rangle + \beta_2 \langle x, y_2 \rangle \end{aligned}$$

for all $x, x_1, x_2 \in X$, $y, y_1, y_2 \in Y$, and $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{C}$ (or $\mathbb{R}$). The bilinear form is called **nondegenerate** if for all $x \in X$, $x \neq 0$, there exists a $y \in Y$ with $\langle x, y \rangle \neq 0$, and if for all $y \in Y$, $y \neq 0$, there exists an $x \in X$ with $\langle x, y \rangle \neq 0$.

The normed spaces $X, Y$ equipped with a nondegenerate bilinear form are called a **dual system**, denoted by $\langle X, Y \rangle$.

The bilinear form is called **bounded** if a $\gamma > 0$ exists such that $|\langle x, y \rangle| \leq \gamma \|x\| \, \|y\|$ for all $x \in X$ and $y \in Y$.

If one replaces the second requirement for a bilinear form by

$$\langle x, \beta_1 y_1 + \beta_2 y_2 \rangle = \bar{\beta}_1 \langle x, y_1 \rangle + \bar{\beta}_2 \langle x, y_2 \rangle$$

one gets a so-called **sesquilinear form**, where an analogous theory can be derived.

**Example 2.22.**

(a) Let $X$ be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$. Then $\langle X, X \rangle$ is a dual system (over $\mathbb{R}$) with a bounded bilinear form. The boundedness ($\gamma = 1$) follows from the Cauchy-Schwarz inequality.

This is the reason why we use the same notation for inner products and nondegenerate bilinear forms.

(b) Let $X$ be a Banach space with dual space $X^*$. $\langle \cdot, \cdot \rangle : X \times X^* \to \mathbb{C}$ (or $\mathbb{R}$) is defined by $\langle x, f \rangle := f(x)$. Then $\langle X, X^* \rangle$ is a dual system with bounded bilinear form ($\gamma = 1$). The nondegeneracy follows from the Hahn-Banach theorem.

(c) Let $X$ be a Banach space that is continuously and densely embedded in the real Hilbert space $H$ with inner product $\langle \cdot, \cdot \rangle_H$, i.e., $I : X \to H$ is continuous and $I(X)$ is dense in $H$. If we define $\langle x, y \rangle := \langle Ix, Iy \rangle_H$ for all $x, y \in X$, then $\langle X, X \rangle$ is a dual system with bounded bilinear form. The nondegeneracy is shown as follows: let $x \in X$ be such that $\langle x, y \rangle = 0$ for all $y \in X$, i.e., $\langle Ix, Iy \rangle_H = 0$ for all $y \in X$. Since $I(X)$ is dense in $H$, it follows that $\langle Ix, h \rangle_H = 0$ for all $h \in H$ and, hence, $Ix = x = 0$. In an analogous way one shows the second part. The boundedness follows from $|\langle x, y \rangle| = |\langle Ix, Iy \rangle_H| \le \|Ix\|_H \|Iy\|_H \le \|I\|^2_{X,H} \|x\|_X \|y\|_X$.

An important special case is the following: $X = C(G)$, $H = L^2(G)$ with

$$\langle f, g \rangle_H := \int_G f(t) g(t) r(t)\, dt\,,$$

where $r : G \to \mathbb{R}^+$ is a continuous weight function. Then

$$\|x\|^2_H = \int_G x^2(t) r(t)\, dt \le \|x\|^2_\infty \int_G r(t)\, dt \quad \Longrightarrow \quad \gamma = \|I\|^2_{X,H} \le \int_G r(t)\, dt\,.$$

Most of the time $r \equiv 1$ so that $\gamma = \|I\|^2_{X,H} \le |G|$.

This example allows a comfortable formulation of the Fredholm theory in $C(G)$ without using its dual space.

(d) Finally, we give an example of an unbounded nondegenerate bilinear form, due to Kress [5]: let

$$X := \left\{ x \in C(0,1] : \exists M, \alpha > 0 \, \forall t \in (0,1] : |x(t)| \le M t^{\alpha - \frac{1}{2}} \right\}$$

with $\|x\| := \sup_{t \in (0,1]} \left( \sqrt{t} |x(t)| \right)$ and $\langle x, y \rangle := \int_0^1 x(t) y(t)\, dt$. The unboundedness can be easily seen using the functions $x_n(t) := t^{\frac{1}{n} - \frac{1}{2}}$.

**Definition 2.23.** Let $\langle X, Y \rangle$ be a dual system. $A : X \to X$ and $B : Y \to Y$ are called **adjoint** to each other if for all $x \in X$ and $y \in Y$: $\langle Ax, y \rangle = \langle x, By \rangle$.

**Remark 2.24.** In case of Example 2.22 (a), (b), a linear bounded operator $A : X \to X$ always has an adjoint, namely the Hilbert space adjoint $A^* : X \to X$ in case (a) and the Banach space adjoint $A^* : X^* \to X^*$ in case (b).

For general dual systems the existence of an adjoint operator is not automatically guaranteed: let, e.g., $X = C[0,1]$, $H = L^2[0,1]$, and $\langle \cdot, \cdot \rangle$ as in Example 2.22 (c) with $r \equiv 1$, i.e., $\langle x, y \rangle := \int_0^1 x(t)y(t)\,dt$. Moreover, $A : X \to X$ is defined by $Ax(t) := x(0)$.

Let us assume that $A$ has an adjoint operator $B$ in the dual system $\langle X, X \rangle$. Setting $v := By$ with $y \equiv 1$, we get that

$$x(0) = \langle Ax, y \rangle = \langle x, By \rangle = \langle x, v \rangle = \int_0^1 x(t)v(t)\,dt \qquad \text{for all} \qquad x \in X\,.$$

But then $\langle x, v \rangle = 0$ for all $x \in D := \{x \in X : x(0) = 0\}$. Since $D$ is dense in $H$, $\langle x, v \rangle = 0$ for all $x \in H$. But then $v = 0$, so that $\langle x, v \rangle \neq x(0)$ for all functions $x$ with $x(0) \neq 0$. This is a contradiction. This argument shows that such a $v$ can even not exist in $H = L^2[0,1]$. Thus, $A$ has no adjoint operator in this dual system.

It is an immediate consequence of the properties of a dual system that, whenever $A$ has an adjoint operator $B$, then $B$ is uniquely determined and linear.

For integral operators the existence (and compactness) of an adjoint operator is always guaranteed in the dual system $\langle C(G), C(G) \rangle$. One can even give an explicit formula for it:

**Theorem 2.25.**

(a) Let $k \in L^2(G \times G)$ and let $K$ be the induced integral operator, i.e.,

$$(Kx)(s) := \int_G k(s,t)x(t)\,dt$$

for all $x \in L^2(G)$ and $s \in G$. Moreover,

$$(K'y)(s) := \int_G k(t,s)y(t)\,dt \tag{2.35}$$

for all $y \in L^2(G)$ and $s \in G$. Then $K$ and $K'$ are adjoint to each other in the dual system $\langle L^2(G), L^2(G) \rangle$.

(b) Let $k \in C(G \times G)$ or let $k$ be a weakly singular kernel with induced integral operator $K$ on $C(G)$. Moreover, let $K'$ be defined as in (2.35) for all $y \in C(G)$ and $s \in G$. Then $K$ and $K'$ are adjoint to each other in the dual systems $\langle C(G), C(G) \rangle$ with $\langle x, y \rangle := \int_G x(t)y(t)\,dt$.

In both cases $K'$ is again compact.

**Proof:** According to Theorems 2.1, 2.11, and 2.13, in all cases $K'$ exists and is compact. The fact that $K$ and $K'$ are adjoint to each other follows from the equation

$$\begin{aligned}\langle Kx, y\rangle &= \int_G \int_G k(s,t)x(t)\,dt\,y(s)\,ds = \int_G x(t)\int_G k(s,t)y(s)\,ds\,dt \\ &= \int_G x(t)(K'y)(t)\,dt = \langle x, K'y\rangle .\end{aligned}$$

Interchanging the order of integration is allowed due to Fubini's theorem, whose applicability in case of a continuous or $L^2$-kernel follows from the integrability of the function $(s,t)\mapsto |k(s,t)x(t)y(s)|$. If $k$ is weakly singular, then $(Kx)(s) = \lim_{n\to\infty}(K_n x)(s)$ (uniformly in $s$) where $K_n$ are integral operators with continuous kernels $k_n(s,t)$ (see the proof of Theorem 2.13). For the approximation operator $K_n$ it follows as above that $K_n'$ (with kernel $k_n(t,s)$) is adjoint to $K_n$. Due to the uniformity of the convergence (in $s$), we obtain

$$\begin{aligned}\langle Kx, y\rangle &= \int_G \lim_{n\to\infty}(K_n x)(s)y(s)ds = \lim_{n\to\infty}\langle K_n x, y\rangle = \lim_{n\to\infty}\langle x, K_n' y\rangle \\ &= \int_G x(s)\lim_{n\to\infty}(K_n' y)(s)ds = \langle x, K'y\rangle\end{aligned}$$

for all $x, y \in C(G)$. ∎

For the proof of the main theorem of the Fredholm theory we need a lemma about the existence of a biorthogonal system to finitely many linear independent vectors:

**Lemma 2.26.** Let $\langle X, Y\rangle$ be a dual system and let $n\in\mathbb{N}$. Then to every set of linearly independent elements $x_1,\ldots,x_n \in X$ there exist elements $y_1,\ldots,y_n\in Y$ such that

$$\langle x_i, y_j\rangle = \delta_{ij}, \quad i,j = 1,\ldots,n.$$

The same statement holds with the roles of $X$ and $Y$ interchanged.

**Proof:** The proof is done by induction: if $n=1$, the assertion is true, since $\langle\cdot,\cdot\rangle$ is nondegenerate, i.e., $y\in Y$ exists with $\langle x_1, y\rangle =: c \neq 0$. But then, setting $y_1 := \frac{y}{c}$ yields that $\langle x_1, y_1\rangle = 1$. Assume that the assertion holds for $n\in\mathbb{N}$ linearly independent elements.

Induction step: let $x_1,\ldots,x_{n+1}\in X$ be $n+1$ linearly independent elements. Then, due to the induction hypothesis, applied to $x_1,\ldots,x_{m-1},x_{m+1},\ldots,x_{n+1}$, there exist elements $y_1^{(m)},\ldots,y_{m-1}^{(m)},y_{m+1}^{(m)},\ldots,y_{n+1}^{(m)}\in Y$ for all $m=1,\ldots,n+1$ such that

$$\left\langle x_i, y_j^{(m)}\right\rangle = \delta_{ij}, \qquad i,j\in\{1,\ldots,n+1\}\setminus\{m\}. \tag{2.36}$$

Due to the linear independence of $x_1,\ldots,x_{n+1}$,

$$x_m - \sum_{\substack{j=1\\ j\neq m}}^{n+1}\left\langle x_m, y_j^{(m)}\right\rangle x_j \neq 0.$$

Since $\langle \cdot, \cdot \rangle$ is nondegenerate, this implies the existence of $w_m \in Y$ with

$$0 \neq \left\langle x_m - \sum_{\substack{j=1 \\ j \neq m}}^{n+1} \left\langle x_m, y_j^{(m)} \right\rangle x_j, w_m \right\rangle = \left\langle x_m, w_m - \sum_{\substack{j=1 \\ j \neq m}}^{n+1} \langle x_j, w_m \rangle y_j^{(m)} \right\rangle =: \alpha_m \,.$$

Setting

$$y_m := \frac{1}{\alpha_m} \left( w_m - \sum_{\substack{j=1 \\ j \neq m}}^{n+1} \langle x_j, w_m \rangle y_j^{(m)} \right),$$

it obviously holds that $\langle x_m, y_m \rangle = 1$ and that, due to (2.36),

$$\langle x_i, y_m \rangle = \frac{1}{\alpha_m} \left( \langle x_i, w_m \rangle - \sum_{\substack{j=1 \\ j \neq m}}^{n+1} \langle x_j, w_m \rangle \left\langle x_i, y_j^{(m)} \right\rangle \right) = 0$$

for all $i \neq m$. Thus, we constructed elements $y_1, \ldots, y_{n+1} \in Y$ such that $\langle x_i, y_j \rangle = \delta_{ij}$ for all $i, j = 1, \ldots, n+1$. The proof for the interchanged roles is analogous. ∎

Now we are in the position to prove the Fredholm alternative:

**Theorem 2.27 (Fredholm alternative).** Let $\langle X, Y \rangle$ be a dual system, $K : X \to X$ and $K' : Y \to Y$ are compact and adjoint to each other. Then the following assertions hold:

The homogeneous equations

$$x - Kx = 0 \tag{2.37}$$

and

$$y - K'y = 0 \tag{2.38}$$

have the same finite number of linearly independent solutions, i.e.,

$$\dim(\mathcal{N}(I - K)) = \dim(\mathcal{N}(I - K')) < \infty \,.$$

Moreover, the inhomogeneous equation

$$x - Kx = f \tag{2.39}$$

with $f \in X$ is solvable if and only if $\langle f, y \rangle = 0$ for all solutions $y$ of (2.38), and the inhomogeneous equation

$$y - K'y = g$$

with $g \in Y$ is solvable if and only if $\langle x, g \rangle = 0$ for all solutions $x$ of (2.37).

**Proof:** It was already shown in Theorem 2.16 that

$$m := \dim(\mathcal{N}(I - K)) \qquad \text{and} \qquad n := \dim(\mathcal{N}(I - K'))$$

are finite. It remains to be shown that $n = m$. Let us assume that $m < n$.

Let $x_1, \ldots, x_m$ be a basis for $\mathcal{N}(I - K)$ (only possible for $m > 0$) and let $y_1, \ldots, y_n$ be a basis for $\mathcal{N}(I - K')$. Due to Lemma 2.26, there exist elements $a_1, \ldots, a_m \in Y$ and $b_1, \ldots, b_n \in X$ with

$$\langle x_i, a_j \rangle = \delta_{ij}\,, \quad i, j \in \{1, \ldots, m\}\,, \qquad \langle b_i, y_j \rangle = \delta_{ij}\,, \quad i, j \in \{1, \ldots, n\}\,. \tag{2.40}$$

Let $T : X \to X$ be defined by

$$Tx := \sum_{i=1}^{m} \langle x, a_i \rangle\, b_i\,. \tag{2.41}$$

Note that $T \equiv 0$ for $m = 0$. We show that

$$\mathcal{N}(I - K + T) = \{0\}\,. \tag{2.42}$$

This is trivial for $m = 0$, but has to be shown for $m > 0$. Let $x \in \mathcal{N}(I - K + T)$. Using that $y_j - K'y_j = 0$, it follows with (2.40) that

$$\begin{aligned} \langle x, a_j \rangle &= \langle x, a_j \rangle + \langle x, y_j - K'y_j \rangle) = \left\langle x, \sum_{i=1}^{m} \langle b_i, y_j \rangle\, a_i \right\rangle + \langle x - Kx, y_j \rangle \\ &= \left\langle \sum_{i=1}^{m} \langle x, a_i \rangle\, b_i, y_j \right\rangle + \langle x - Kx, y_j \rangle = \langle x - Kx + Tx, y_j \rangle = 0 \end{aligned} \tag{2.43}$$

for all $j \in \{1, \ldots, m\}$. Therefore, $Tx = 0$ and, hence, $x \in \mathcal{N}(I - K)$. Since $x_1, \ldots, x_m$ is a basis for $\mathcal{N}(I - K)$, there exist coefficients $\alpha_1, \ldots, \alpha_m$ with $x = \sum_{i=1}^{m} \alpha_i x_i$. Now (2.40) and (2.43) imply that

$$\alpha_j = \langle x, a_j \rangle = 0\,, \quad j \in \{1, \ldots, m\} \quad \Longrightarrow \quad x = 0\,.$$

This proves (2.42). Now we show that

$$\mathcal{R}(I - K + T) = X\,. \tag{2.44}$$

Let $\nu$ be the Riesz index of $(I - K)$ and $P : X \to \mathcal{N}((I - K)^\nu)$ the projector induced by the decomposition (2.28). First, we show that

$$(I - K - P) : X \to X \quad \text{is bijective and continuously invertible}\,. \tag{2.45}$$

Since, due to Theorem 2.19, $P$ is continuous and since $\dim(\mathcal{R}(P)) = \dim(\mathcal{N}((I - K)^\nu)$ is finite, Theorem 2.8 (f) implies that $P$ is compact. Thus, also $K + P$ is compact. Let now $x \in \mathcal{N}(I - K - P)$. Then $(I - K)x = Px$. Together with $Px \in \mathcal{N}((I - K)^\nu)$ this implies that $(I - K)^{\nu+1}x = 0$, and hence, due to the definition of $\nu$ that $(I - K)^\nu x = 0$, i.e., $x \in \mathcal{N}((I - K)^\nu)$. But then $x = Px = (I - K)x$. This implies that

$$x = (I - K)x = (I - K)^2 x = \ldots = (I - K)^\nu x = 0\,.$$

Since $x \in \mathcal{N}(I - K - P)$ was arbitrary, this shows that $\mathcal{N}(I - K - P) = \{0\}$. Now Theorem 2.19, applied to the compact operator $K + P$ yields (2.45).

To prove (2.44), we have to show that

$$(I - K + T)x = z \tag{2.46}$$

is solvable for any $z \in X$. We will construct a solution using the operator

$$A : \mathcal{R}(P+T) \to X\,, \quad \text{defined by} \quad A := (I - K + T)(I - K - P)^{-1}\,.$$

Since

$$A = (I - K - P + P + T)(I - K - P)^{-1} = I + (P + T)(I - K - P)^{-1}\,,$$

$\mathcal{R}(A) \subseteq \mathcal{R}(P+T)$, i.e., $A$ maps the finite-dimensional space $\mathcal{R}(P+T)$ into itself.

If $w \in \mathcal{N}(A)$, then, due to (2.42), $(I - K - P)^{-1}w = 0$ and, hence, $w = 0$. Therefore, $A$ is injective and, due to $\dim(\mathcal{R}(P+T)) < \infty$, also surjective as an operator to $\mathcal{R}(P+T)$. Thus, equation

$$v + (P + T)(I - K - P)^{-1}v = Av = (P + T)(I - K - P)^{-1}z \tag{2.47}$$

has a unique solution $v \in \mathcal{R}(P+T)$. Setting $x := (I - K - P)^{-1}(z - v)$, we obtain that

$$\begin{aligned}(I - K + T)x &= (I - K - P)x + (P + T)x \\ &= z - v + Av - (P + T)(I - K - P)^{-1}v \\ &= z - v + Av + (v - Av) \;=\; z\,.\end{aligned}$$

Thus, $x$ solves (2.46). Since $z \in X$ was arbitrary, this proves (2.44) and together with (2.42) that $(I - K + T)$ is bijective.

As a consequence equation

$$(I - K + T)x = b_{m+1}$$

has a unique solution $x_{m+1}$. Now (2.40), (2.41), and $y_{m+1} \in \mathcal{N}(I - K')$ imply that

$$\begin{aligned}1 &= \langle\, b_{m+1}, y_{m+1} \,\rangle \;=\; \langle\, x_{m+1} - Kx_{m+1} + Tx_{m+1}, y_{m+1} \,\rangle \\ &= \langle\, x_{m+1} - Kx_{m+1}, y_{m+1} \,\rangle + \sum_{i=1}^{m} \langle\, x_{m+1}, a_i \,\rangle \,\langle\, b_i, y_{m+1} \,\rangle \\ &= \langle\, x_{m+1}, y_{m+1} - K'y_{m+1} \,\rangle \;=\; 0\,.\end{aligned}$$

This is a contradiction. Thus, the assumption $m < n$ was wrong. In an analogous way one shows that the assumption $m > n$ leads to a contradiction. This then proves that $m = n$.

Finally, we have to prove the assertions about the inhomogeneous equations: let $f \in X$, $x$ a solution of (2.39), and $y$ an arbitrary solution of (2.38). Then

$$\langle\, f, y \,\rangle = \langle\, (I - K)x, y \,\rangle = \langle\, x, (I - K')y \,\rangle = 0\,.$$

Let us now assume that $\langle\, f, y \,\rangle = 0$ for all solutions $y$ of (2.38). We have to prove the solvability of (2.39) and distinguish two cases:

We have already shown that $m := \dim(\mathcal{N}(I - K)) = \dim(\mathcal{N}(I - K')) < \infty$. If $m = 0$, then (2.39) is solvable, due to Theorem 2.19: since $\nu = 0$, $\mathcal{R}(I-K) = \mathcal{R}((I-K)^0) = X$. Note that $y = 0$ is then the only solution of (2.38).

If $m > 0$, then we have already shown that there exists a unique $x$ with $(I-K+T)x = f$. Together with (2.40), (2.41), and $y_j \in \mathcal{N}(I - K')$ we get that

$$\begin{aligned} \langle x, a_j \rangle &= \langle x, y_j - K'y_j \rangle + \langle x, a_j \rangle \\ &= \langle x - Kx, y_j \rangle + \sum_{i=1}^{m} \langle x, a_i \rangle \langle b_i, y_j \rangle \\ &= \langle x - Kx + Tx, y_j \rangle = \langle f, y_j \rangle = 0 . \end{aligned}$$

for all $j \in \{1, \ldots, m\}$. Thus, $Tx = 0$, i.e., $x$ also solves (2.39).

The assertion about the other inhomogeneous equation follows analogoulsy. ∎

**Remark 2.28.** This version of the Fredholm alternative includes the Hilbert space version and the Banach space version, where one uses the dual space (sometimes called Schauder theory) (see Example 2.22 (a) and (b)). However, it is also applicable to integral equations in $C(G)$, where the adjoint is considered on the same space (see Example 2.22 (c) and Theorem 2.25).

As can be easily checked, all theorems above remain valid if in the dual form the bilinear form is replaced by a sesquilinear form. This then includes complex Hilbert spaces, e.g., $L^2(G)$ with

$$\langle x, y \rangle := \int_G x(t)\overline{y(t)}\, dt .$$

Note that then in (2.35) $k(t, s)$ has to be replaced by $\overline{k(t, s)}$.

We now present a characterization for the case of Riesz index $\nu = 1$ in a dual system:

**Theorem 2.29.** Let $\langle X, Y \rangle$, $K$, and $K'$ be as in Theorem 2.27, and assume that $\dim(\mathcal{N}(I - K)) > 0$. Then the following two assertions are equivalent:

(a) $K$ and $K'$ have Riesz index 1.

(b) The matrix $(\langle x_i, y_j \rangle)_{i,j\in\{1,\ldots,m\}}$ is regular for some bases $x_1, \ldots, x_m$ of $\mathcal{N}(I - K)$ and $y_1, \ldots, y_m$ of $\mathcal{N}(I - K')$.

**Proof:** We prove that $\neg(a) \Leftrightarrow \neg(b)$.

Noting that, due to Theorem 2.27, $K$ and $K'$ have the same Riesz index $\nu$, it holds that $\nu > 1$ is equivalent to $\mathcal{N}(I - K) \subsetneq \mathcal{N}((I - K)^2)$. But this is equivalent to the existence of an $x \in X$ with

$$(I - K)^2 x = 0 \qquad \text{and} \qquad (I - K)x =: f \neq 0$$

According to Theorem 2.27, this is equivalent to $\langle f, y \rangle = 0$ for some $f \in \mathcal{N}(I-K)\setminus\{0\}$ and all $y \in \mathcal{N}(I - K')$. This, however, is equivalent to

$$0 = \left\langle \sum_{i=1}^{m} \lambda_i x_i, y_j \right\rangle = \sum_{i=1}^{m} \langle x_i, y_j \rangle \lambda_i \quad \text{for all} \quad j \in \{1, \ldots, m\}$$

for any basis $x_1, \dots, x_m$ of $\mathcal{N}(I-K)$, any basis $y_1, \dots, y_m$ of $\mathcal{N}(I-K')$, and some vector $(\lambda_1, \dots, \lambda_m) \neq 0$. Finally, this is equivalent to the fact that $(\langle x_i, y_j \rangle)_{i,j\in\{1,\dots,m\}}$ is singular. ∎

**Remark 2.30.** If the operator $K$ is selfadjoint in the dual system $\langle X, X \rangle$, then the Riesz index of $K$ can either be 0 or 1. Because, if it's not 0, then $\dim(\mathcal{N}(I-K)) > 0$ and the condition of Theorem 2.29 (b) is satisfied for every basis of $\mathcal{N}(I-K) = \mathcal{N}(I-K')$ with $y_i = x_i$, since for any set of linearly independent elements $x_1, \dots, x_m$, the Gramian matrix $(\langle x_i, x_j \rangle)_{i,j\in\{1,\dots,m\}}$ is regular.

## 2.5. Spectral theory for compact operators

In this section, we are interested for what values $\lambda \in \mathbb{C}$ the equation

$$\lambda x(s) - \int_0^1 k(s,t)x(t)\,dt = 0$$

has a nontrivial solution $x$. This leads us to the discussion of the spectrum of compact operators.

**Definition 2.31.** Let $X$ be a normed space and $T \in L(X)$. The **spectrum** of $T$ is defined by

$$\sigma(T) := \{\lambda \in \mathbb{C} : \lambda I - T : X \to X \quad \text{has no continuous inverse}\}$$

$\lambda \in \mathbb{C}$ is called **eigenvalue** of $T$ if $\mathcal{N}(\lambda I - T) \neq \{0\}$. Each element of $\mathcal{N}(\lambda I - T)$ is called **eigenvector** or **eigenfunction** of $T$ to the eigenvalue $\lambda$.

The spectrum of $T$ consists of those **spectral values** $\lambda \in \mathbb{C}$, for which it does **not** hold that the equation

$$\lambda x - Tx = f \tag{2.48}$$

has a unique solution for all right-hand sides $f$ and that it depends continuoulsy on $f$. The value $\lambda$ can be in $\sigma(T)$ for different reasons: either the requirement for uniqueness, or existence, or continuous dependence of a solution of (2.48), or several of these requirements may be violated. $\lambda$ is an eigenvalue if the requirement for uniqueness is violated.

**Example 2.32.** In the spectrum there can also be values that are no eigenvalues: the operator $K : C[0,1] \to C[0,1]$ defined by

$$(Kx)(s) := \int_0^s x(t)\,dt$$

is compact. Taking the derivative yields that $(Kx)' = x$. This implies that $\mathcal{N}(K) = \{0\}$. Due to Theorem 2.17, $\mathcal{R}(K) = \{f \in C^1[0,1] : f(0) = 0\}$ is not closed. Therefore, $K^{-1}$ does not exist. Thus, $0 \in \sigma(K)$, but 0 is no eigenvalue.

In the next theorem we collect some properties of the spectrum of compact operators:

**Theorem 2.33.** Let $X$ be a normed space and let $K : X \to X$ be compact. Then it holds:

(a) If $\dim X = \infty$, then $0 \in \sigma(K)$.

(b) If $\lambda \in \sigma(K) \setminus \{0\}$, then $\lambda$ is an eigenvalue of $K$ with finite geometrical multiplicity, i.e., $\dim(\mathcal{N}(\lambda I - K)) < \infty$.

(c) $\sigma(K)$ is at most countable with 0 as the only possible accumulation point in $\mathbb{C} \cup \{\infty\}$.

**Proof:**

(a) Let $\dim X = \infty$ and assume that $0 \notin \sigma(K)$, i.e., $K^{-1} \in L(X, X)$. Then, due to Theorem 2.8 (d), $I = K^{-1}K$ is compact which is a contradiction to Lemma 2.15. Therefore, the assumption was wrong.

(b) We assume that $\lambda \neq 0$ is no eigenvalue of $K$, i.e., $\mathcal{N}(\lambda I - K) = \{0\}$. But then $\lambda I - K$ has Riesz index 0 and, hence, $\mathcal{R}(\lambda I - K) = X$. According to Theorem 2.19, $\lambda I - K$ has then a continuous inverse on $\mathcal{R}(\lambda I - K) = X$. Thus, $\lambda \notin \sigma(K)$.

This shows that each $\lambda \in \sigma(K) \setminus \{0\}$ is an eigenvalue, i.e, $\mathcal{N}(\lambda I - K) \neq \{0\}$. Then Theorem 2.19 implies that $\dim(\mathcal{N}(\lambda I - K)) \leq \dim(\mathcal{N}((\lambda I - K)^\nu)) < \infty$.

(c) Let $\lambda_1, \lambda_2, \ldots \in \sigma(K) \setminus \{0\}$ be pairwise different with $\lambda_n \to \lambda \in \mathbb{C}$, i.e., $\lambda$ is an accumulation point of $\sigma(K)$. Let us assume that $\lambda \neq 0$.

According to (b), each $\lambda_n$ is an eigenvalue. Therefore, an eigenvector $x_n$ exists to $\lambda_n$. Note that each set of eigenvectors $\{x_1, \ldots, x_n\}$ is linearly independent as the following argument shows: let $(\alpha_1, \ldots, \alpha_n)$ be such that $\sum_{i=1}^n \alpha_i x_i = 0$. Then, due to the definition of $x_i$, also $\sum_{i=1}^n \alpha_i \lambda_i^j x_i = 0$ for all $j = 1, \ldots, n$. Since $\left(\lambda_i^j\right)_{i,j=1,\ldots,n}$ is a Vandermonde matrix, this implies that $\alpha_i = 0$ for all $i = 1, \ldots, n$.

Let $X_n := \operatorname{span}\{x_1, \ldots, x_n\}$. Then $X_1 \subsetneq X_2 \subsetneq X_3 \subsetneq \ldots$. Due to Lemma 2.14, for all $n \in \mathbb{N}$ an element $z_n \in X_n$ exists with $\|z_n\| = 1$ and $\|z_n - x\| \geq \frac{1}{2}$ for all $x \in X_{n-1}$. Since $z_n = \sum_{i=1}^n \alpha_i^{(n)} x_i$ for some $(\alpha_i^{(n)})_{i=1,\ldots,n}$, we obtain that

$$Kz_n - \lambda_n z_n = \sum_{i=1}^n \alpha_i^{(n)}(Kx_i - \lambda_n x_i) = \sum_{i=1}^{n-1} \alpha_i^{(n)}(\lambda_i - \lambda_n)x_i \in X_{n-1}\,.$$

Obviously, $Kz_m \in X_m \subseteq X_{n-1}$ for all $m < n$. Therefore,

$$Kz_m - (Kz_n - \lambda_n z_n) \in X_{n-1} \quad \text{for all} \quad m < n$$

and, hence,

$$\begin{aligned} \|Kz_n - Kz_m\| &= \|\lambda_n z_n - (Kz_m - (Kz_n - \lambda_n z_n))\| \\ &= |\lambda_n| \left\| z_n - \frac{1}{\lambda_n}(Kz_m - (Kz_n - \lambda_n z_n)) \right\| \geq \frac{|\lambda_n|}{2} > \frac{|\lambda|}{4} \end{aligned}$$

for $n$ sufficiently large and $m < n$. Thus, $(Kz_n)$ has no convergent subsequence, which is a contradiction to the compactness of $K$. Therefore, the assumption $\lambda \neq 0$ was wrong. The contradiction proof above also works for the case $|\lambda_n| \to \infty$. One only has to replace $\frac{|\lambda|}{4}$ in the last estimate by 1. This shows that 0 is the only possible accumulation point.

But then the sets $\Lambda_n := \{\lambda \in \sigma(K) : |\lambda| \geq \frac{1}{n}\}$, $n \in \mathbb{N}$, are finite-dimensional, since otherwise $\Lambda_n$ would have an accumulation point in $\{z \in \mathbb{C} : |z| \geq \frac{1}{n}\} \cup \{\infty\}$. Thus, $\sigma(K) \subseteq \{0\} \cup \bigcup_{n\in\mathbb{N}} \Lambda_n$, i.e., $\sigma(K)$ is at most countable.

∎

For selfadjoint operators in Hilbert spaces one can prove additional results about $\sigma(K)$. For the time being, we assume that $H \neq \{0\}$ is a real or complex Hilbert space with inner product $\langle \cdot, \cdot \rangle$. Then $\langle H, H \rangle$ is a dual system, where in the complex case one has to use a sesquilinear form instead of a bilinear form (see Remark 2.28). According to Definition 2.23, an operator $K : H \to H$ is **selfadjoint** if

$$\langle Kx, y \rangle = \langle x, Ky \rangle \quad \text{for all} \quad x, y \in H\,.$$

Due to Remark 2.30, $\lambda I - K$ has Riesz index 0 or 1 for all $\lambda \neq 0$ if $K$ is selfadjoint and compact.

Thus, integral operators in $L^2(G)$ with a symmetric (or Hermite) $L^2$-kernel $k$, i.e., where $k(s,t) = k(t,s)$ (or $k(s,t) = \overline{k(t,s)}$) a.e. in $G \times G$, are selfadjoint (see Theorem 2.25).

Such operators have at least one eigenvalue whose absolute value equals $\|K\|$:

**Theorem 2.34.** Let $K : H \to H$ be compact and selfadjoint. Then

$$\{-\|K\|, \|K\|\} \cap \sigma(K) \neq \emptyset\,.$$

**Proof:** If $K = 0$, the assertion is trivial. Let now $K \neq 0$ and let $(x_n)$ be a sequence in $H$ with $\|x_n\| = 1$ and $\|Kx_n\| \to \|K\|$. Such a sequence exists according to the definition of $\|K\|$. Noting that $\langle K^2 x, x \rangle = \|Kx\|^2$,

$$\begin{aligned} 0 &\leq \left\| K^2 x_n - \|Kx_n\|^2 x_n \right\|^2 \\ &= \left\| K^2 x_n \right\|^2 - 2\|Kx_n\|^2 \langle K^2 x_n, x_n \rangle + \|Kx_n\|^4 \|x_n\|^2 \\ &\leq \|K\|^2 \|Kx_n\|^2 - \|Kx_n\|^4 \underset{n\to\infty}{\longrightarrow} 0\,. \end{aligned}$$

Thus, $K^2 x_n - \|Kx_n\|^2 x_n \to 0$. Together with $\|Kx_n\| \to \|K\|$ this implies that

$$K^2 x_n - \|K\|^2 x_n \to 0\,. \tag{2.49}$$

Since $K^2$ is compact, $(K^2 x_n)$ has a convergent subsequence $(K^2 x_{n_k})$. Together with (2.49) this implies the convergence of $(x_{n_k})$. Let $x := \lim_{k\to\infty} x_{n_k}$. Then $\|x\| = 1$ and

$$0 = (K^2 x - \|K\|^2 x) = (K - \|K\| I)(K + \|K\| I)x\,.$$

Therefore, either $x$ is eigenvector of $K$ to the eigenvalue $-\|K\|$ or $z := Kx + \|K\| x \neq 0$, is eigenvector of $K$ to the eigenvalue $\|K\|$. ∎

**Remark 2.35.** The following two assertions hold for a selfadjoint operator $K$:

(a) The eigenvalues are real valued: if $x \neq 0$ and $Kx = \lambda x$, then

$$\lambda \langle x, x \rangle = \langle Kx, x \rangle = \langle x, Kx \rangle = \bar{\lambda} \langle x, x \rangle \quad \Longrightarrow \quad \lambda \in \mathbb{R}\,.$$

(b) Eigenvectors to different eigenvalues are orthogonal to each other: if $\lambda_1 \neq \lambda_2$, $Kx_1 = \lambda_1 x_1$, and $Kx_2 = \lambda_2 x_2$, then

$$\lambda_1 \langle x_1, x_2 \rangle = \langle Kx_1, x_2 \rangle = \langle x_1, Kx_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle \quad \Longrightarrow \quad \langle x_1, x_2 \rangle = 0\,.$$

So far we know that a compact selfadjoint operator $K \neq 0$ has at least one and at most countably many non-zero eigenvalues. The eigenspace to each of these eigenvalues is finite-dimensional. If we choose an orthonormal basis for each eigenspace, we obtain an orthonormal system:

**Definition 2.36.** Let $K : H \to H$ be compact and selfadjoint with non-zero eigenvalues $\lambda_1, \lambda_2, \lambda_3, \ldots$, where each eigenvalue is repeated according to its multiplicity. Let $x_1, x_2, \ldots$ be an orthonormal system such that $Kx_i = \lambda_i x_i$ for all $i \in \mathbb{N}$. If $K$ has only finitely many eigenvalues, then $(\lambda_i)$ and $(x_i)$ denote finite sequences. The system $(\lambda_i, x_i)_{i\in\mathbb{N}}$ is called an **eigensystem** of $K$.

We say *an* and not *the* eigensystem, since the eigenvectors $x_i$ are not uniquely determined, even not when the eigenspaces are one-dimensional.

We will see that the knowledge of an eigensystem of a compact selfadjoint operator allows its reconstruction. If it is an integral operator, one can even represent the kernel using the eigensystem.

In a first step we show that the vectors $x_i$ span the range of the operators (in the sense of a Hilbert space basis).

In the following, we always assume that $K \neq 0$. If $K$ has only finitely many eigenvalues, then all series have to be interpreted as finite sums.

**Theorem 2.37.** Let $(\lambda_i, x_i)_{i\in\mathbb{N}}$ be an eigensystem of the compact selfadjoint operator $K : H \to H$. Then $\{x_i : i \in \mathbb{N}\}$ is an orthonormal basis of $\overline{\mathcal{R}(K)}$.

**Proof:** The orthonormality of the $\{x_i\}$ and the fact that $x_i \in \mathcal{R}(K)$ immediately follows from the definition of an eigensystem.

Let $E$ be the subspace of the Hilbert space $H$ spanned by $\{x_i : i \in \mathbb{N}\}$.

If $x \in E^\perp$, i.e., $\langle x, x_i \rangle = 0$ for all $i \in \mathbb{N}$, then

$$\langle Kx, x_i \rangle = \langle x, Kx_i \rangle = \lambda_i \langle x, x_i \rangle = 0 \quad \text{for all} \quad i \in \mathbb{N}\,.$$

Thus, $K(E^\perp) \subseteq E^\perp$, and, hence, $K|_{E^\perp}$ is a compact selfadjoint map from $E^\perp$ into itself. If $E^\perp = \{0\}$, we are done, since then $E = H$. If $E^\perp \neq \{0\}$, then Theorem 2.34 implies that $K|_{E^\perp}$ and also $K$ has an eigenvalue whose absolute value equals $\|K|_{E^\perp}\|$. But then it must hold that $\|K|_{E^\perp}\| = 0$, since otherwise the corresponding eigenvector would have to be an element of $E$ and $E^\perp$ at the same time. Thus,

$$\langle x, x_i \rangle = 0 \quad \text{for all} \quad i \in \mathbb{N} \quad \Longrightarrow \quad Kx = 0\,.$$

Let now $y \in \mathcal{R}(K)$ be such that $\langle y, x_i \rangle = 0$ for all $i \in \mathbb{N}$ and let $x \in X$ be such that $Kx = y$. Then

$$0 = \langle Kx, x_i \rangle = \langle x, Kx_i \rangle = \lambda_i \langle x, x_i \rangle = 0 \quad \text{for all} \quad i \in \mathbb{N} \quad \Longrightarrow \quad y = 0\,.$$

This implies that the orthonormal system $\{x_i : i \in \mathbb{N}\}$ is complete in $\mathcal{R}(K)$, i.e., $E = \overline{\mathcal{R}(K)}$. ∎

Elements of $R(K)$, i.e., $Kx$ for any $x$, can be expanded in a series with respect to the orthonormal system $\{x_i : i \in \mathbb{N}\}$:

**Corollary 2.38.** Let $K$ and $(\lambda_i, x_i)_{i\in\mathbb{N}}$ be as in Theorem 2.37. Then it holds for all $x \in H$ that

$$Kx = \sum_{i=1}^{\infty} \lambda_i \langle x, x_i \rangle x_i\,. \tag{2.50}$$

**Proof:** Since $Kx \in \mathcal{R}(K)$, Theorem 2.37 implies that

$$Kx = \sum_{i=1}^{\infty} \langle Kx, x_i \rangle x_i = \sum_{i=1}^{\infty} \langle x, Kx_i \rangle x_i = \sum_{i=1}^{\infty} \lambda_i \langle x, x_i \rangle x_i\,.$$

∎

One can interpret (2.50) as a diagonalization of the operator $K$: using the basis $\{x_i\}$, the application of the operator $K$ corresponds to the multiplication of the $i$-th coordinate by $\lambda_i$. Obviously, one can use this diagonalization to solve equations involving $K$ by decoupling them into (usually infinitely many) equations in the one-dimensional subspaces spanned by the eigenvectors:

**Corollary 2.39.** Let $K$ and $(\lambda_i, x_i)_{i\in\mathbb{N}}$ be as in Theorem 2.37. Then it holds for $\lambda \neq 0$ and $f \in H$:

If $\lambda \notin \sigma(K)$, then the unique solution of

$$\lambda x - Kx = f \tag{2.51}$$

is given by

$$x = \sum_{i=1}^{\infty} \frac{\langle f, x_i \rangle}{\lambda - \lambda_i} x_i + \sum_{i\in\Lambda} \frac{\langle f, n_i \rangle}{\lambda} n_i = \frac{1}{\lambda} f + \sum_{i=1}^{\infty} \frac{\lambda_i \langle f, x_i \rangle}{\lambda(\lambda - \lambda_i)} x_i \,. \tag{2.52}$$

If $\lambda \in \sigma(K) \setminus \{0\}$, then (2.51) is solvable if and only if $\langle f, y \rangle = 0$ for all eigenvectors $y$ to the eigenvalue $\lambda$. Each such solution has the form

$$x = \sum_{\substack{i=1 \\ \lambda_i \neq \lambda}}^{\infty} \frac{\langle f, x_i \rangle}{\lambda - \lambda_i} x_i + \sum_{i\in\Lambda} \frac{\langle f, n_i \rangle}{\lambda} n_i + y \,, \qquad y \in \mathcal{N}(\lambda I - K) \,. \tag{2.53}$$

Here, $\{n_i\}_{i\in\Lambda}$ is an orthonormal basis of $\mathcal{N}(K)$.

The equation

$$Kx = f \tag{2.54}$$

is solvable if and only if $f \in \mathcal{N}(K)^{\perp}$ and if

$$\sum_{i=1}^{\infty} \frac{|\langle f, x_i \rangle|^2}{\lambda_i^2} < \infty \,. \tag{2.55}$$

In this case, each solution of (2.54) has the form

$$x = \sum_{i=1}^{\infty} \frac{\langle f, x_i \rangle}{\lambda_i} x_i + y \,, \qquad y \in \mathcal{N}(K) \,. \tag{2.56}$$

**Proof:** First, note that, due to Remark 2.35 (b) and Theorem 2.37,

$$\{n_i : i \in \Lambda\} \cup \{x_i : i \in \mathbb{N}\}$$

is a basis of $H$. Together with Corollary 2.38 we obtain for all $x, f \in H$ that

$$\lambda x - Kx = \sum_{i=1}^{\infty} (\lambda - \lambda_i) \langle x, x_i \rangle x_i + \lambda \sum_{i\in\Lambda} \langle x, n_i \rangle n_i \tag{2.57}$$

$$f = \sum_{i=1}^{\infty} \langle f, x_i \rangle x_i + \sum_{i\in\Lambda} \langle f, n_i \rangle n_i. \tag{2.58}$$

If $\lambda \notin \sigma(K)$, comparing the Fourier coefficients in (2.57) and (2.58) immediately yields that $x$ as in the first expression of (2.52) is the unique solution of (2.51). The second

one follows from the facts that

$$\sum_{i\in\Lambda} \frac{\langle f, n_i\rangle}{\lambda} n_i = \frac{1}{\lambda} f - \frac{1}{\lambda}\sum_{i=1}^{\infty} \langle f, x_i\rangle x_i \quad \text{and} \quad \frac{1}{\lambda-\lambda_i} - \frac{1}{\lambda} = \frac{\lambda_i}{\lambda(\lambda-\lambda_i)}\,.$$

If $\lambda \in \sigma(K)\setminus\{0\}$, then comparing the Fourier coefficients in (2.57) and (2.58) yields that a solution exists if and only if $\langle f, x_i\rangle = 0$ for all $i \in \mathbb{N}$ with $\lambda_i = \lambda$. But then $\langle f, y\rangle = 0$ for all eigenvectors $y$ to the eigenvalue $\lambda$. (This is already known from the Fredholm alternative). Moreover, the solutions of (2.51) are given by (2.53).

Let us now assume that $\lambda = 0$. Then a comparison of the Fourier coefficients in (2.57) and (2.58) yields that (2.54) has a solution if and only if $\lambda_i \langle x, x_i\rangle = \langle f, x_i\rangle$ for all $i \in \mathbb{N}$ and if $f \in \mathcal{N}(K)^\perp$. But, due to Bessel's inequality $\sum_{i=1}^{\infty} |\langle x, x_i\rangle|^2 \le \|x\|^2 < \infty$, this is true if and only if (2.55) holds. The solution is then given by (2.56). ∎

**Remark 2.40.** The set $\Lambda$ in (2.52) can be empty, finite or infinite. In a non-separable Hilbert space it can even be uncountable. However, since for any $y \in H\setminus\{0\}$ only countably many $\langle y, n_i\rangle \neq 0$, $\sum_{i\in\Lambda}$ is always a common infinte series or finite sum.

An important example for (2.54) is a Fredholm integral equation of the first kind with $L^2$-kernel in $L^2(G)$. For such equations the Fredholm alternative does not hold. In this case, the solvability is guaranteed if the so-called **Picard condition** (2.55) holds. But even then the solution does not depend continuously on the data if $\dim(\mathcal{R}(K)) = \infty$ as the following argument shows: assume that $y^\delta = y + \delta x_i$. Then $y^\delta \in \mathcal{R}(K)$, $\|y - y^\delta\| = \delta$, and, due to (2.56), the error in the solution is given by $\frac{\delta}{\lambda_i}$. Since $\lambda_i \to 0$, the error can be arbitrarily large.

One can even represent the kernel of the integral operator as a series:

**Theorem 2.41.** Let $k \in L^2(G\times G)$ be a Hermite kernel with induced integral operator $K : L^2(G) \to L^2(G)$ and let $(\lambda_i, x_i)_{i\in\mathbb{N}}$ be an eigensystem for the (then selfadjoint compact) operator $K$. Then

$$k(s,t) = \sum_{i=1}^{\infty} \lambda_i x_i(s)\overline{x_i(t)}\,, \tag{2.59}$$

where equality and convergence in (2.59) hold in the $L^2(G\times G)$-sense. Moreover,

$$\|k\|_2^2 = \int_{G\times G} |k(s,t)|^2\, d(s,t) = \sum_{i=1}^{\infty} \lambda_i^2\,. \tag{2.60}$$

**Proof:** Let $\varphi_{ij}(s,t) := x_i(s)\overline{x_j(t)}$ for all $i,j \in \mathbb{N}$ and $s,t \in G$. As in the proof of Theorem 2.3 one can see that, due to Theorem 2.37, $\{\varphi_{ij} : i,j \in \mathbb{N}\}$ is an orthonormal basis of a subspace of $L^2(G\times G)$. Let

$$x \in \mathcal{N}(K) = \mathcal{R}(K)^\perp\,, \quad y \in L^2(G)\,, \quad \alpha(s,t) := y(s)\overline{x(t)}\,, \quad \beta(s,t) := x(s)\overline{y(t)}\,.$$

Then

$$\begin{aligned}
\langle\, k, \alpha \,\rangle &= \int_{G\times G} k(s,t)\overline{y(s)}x(t)\, d(s,t) \\
&= \int_G \int_G k(s,t)x(t)\, dt\, \overline{y(s)}\, ds \;=\; \langle\, Kx, y \,\rangle \;=\; 0\,, \\
\langle\, k, \beta \,\rangle &= \int_{G\times G} k(s,t)\overline{x(s)}y(t)\, d(s,t) \\
&= \int_G y(t) \overline{\int_G k(t,s)x(s)\, ds}\, dt \;=\; \langle\, y, Kx \,\rangle \;=\; 0\,,
\end{aligned}$$

where we used that $k(s,t) = \overline{k(t,s)}$. Thus, $k$ is an element of the subspace spanned by the functions $\varphi_{ij}$. Hence,

$$k(s,t) = \sum_{i,j\in\mathbb{N}} \langle\, k, \varphi_{ij} \,\rangle\, \varphi_{ij}(s,t) \quad \text{in } L^2\text{-sense}\,. \tag{2.61}$$

However,

$$\begin{aligned}
\langle\, k, \varphi_{ij} \,\rangle &= \int_{G\times G} k(s,t)\overline{\varphi_{ij}(s,t)}\, d(s,t) \;=\; \int_G \int_G k(s,t)\overline{x_i(s)}x_j(t)\, dt\, ds \\
&= \int_G (Kx_j)(s)\overline{x_i(s)}\, ds \;=\; \lambda_j \langle\, x_j, x_i \,\rangle \;=\; \lambda_j \delta_{ij}
\end{aligned}$$

for all $i, j \in \mathbb{N}$. Together with (2.61) this yields (2.59). Finally, (2.60) follows from Parseval's identity. ∎

**Remark 2.42.** In a similar way, one could show that (2.59) holds for all $s \in G$ with $\int_G |k(s,t)|^2\, dt < \infty$, i.e., $k(s,\cdot) \in L^2(G)$. The convergence of the series is then meant in $L^2(G)$ with respect to $t$.

**Remark 2.43.** Formula (2.60) implies that the eigenvalues of an integral operator with $L^2$-kernel are in $l^2$. This is, of course, not true for a general compact operator.

An integral operator with $L^2$-kernel belongs to the class of Hilbert-Schmidt operators, a subclass of compact operators, defined as follows: a linear operator $L : H \to H$ is called **Hilbert-Schmidt operator** if

$$\|L\|_2^2 := \sum_{i=1}^{\infty} \|L\varphi_i\|^2 < \infty\,,$$

where $(\varphi_i)$ is a complete orthonormal system. $\|L\|_2$ does not depend on the special choice of the basis. It always holds that $\|L\| \le \|L\|_2$.

If $K$ is a selfadjoint integral operator with $L^2$-kernel $k$ and eigensystem $(\lambda_i, x_i)$, then (2.60) implies that

$$\|K\|_2^2 = \sum_{i=1}^{\infty} \|Kx_i\|^2 = \sum_{i=1}^{\infty} \lambda_i^2 = \int_{G\times G} |k(s,t)|^2\, d(s,t) < \infty\,.$$

For matrices $L$ the Hilbert-Schmidt norm $\|L\|_2$ is identical to the Frobenius norm. The set of Hilbert-Schmidt operators endowed with the inner product

$$(L, M) := \sum_{i=1}^{\infty} \langle\, L\varphi_i, M\varphi_i \,\rangle \ ,$$

where $(\varphi_i)$ is an arbitrary complete orthonormal system, is a Hilbert space.

For continuous kernels one can even prove a stronger convergence in (2.50):

**Theorem 2.44.** Let $k \in C(G \times G)$ be a Hermite kernel with induced integral operator $K : L^2(G) \to L^2(G)$ and let $(\lambda_i, x_i)_{i\in\mathbb{N}}$ be an eigensystem for the operator $K$. Then

$$(Kx)(s) = \sum_{i=1}^{\infty} \lambda_i \,\langle\, x, x_i \,\rangle\, x_i(s) \tag{2.62}$$

for all $x \in L^2(G)$ and $s \in G$, where the convergence is absolute and uniform.

**Proof:** Since $k$ is continuous, a constant $M > 0$ exists such that

$$\int_G |k(s,t)|^2 \, dt \le M \tag{2.63}$$

for all $s \in G$. Theorem 2.41 and Remark 2.42 now imply that (2.59) holds for all $s \in G$ with convergence of the series in the $L^2$-sense with respect to $t$. Therefore, $(Kx)(s) = \langle\, k(s,\cdot), \overline{x} \,\rangle$ satisfies the series representation (2.62) for all $s \in G$.

The convergence of the series is absolute and uniform in $s$ as the following argument shows: using (2.59) and Parseval's identity, we obtain that

$$\int_G |k(s,t)|^2 \, dt = \|k(s,\cdot)\|_2^2 = \sum_{i=1}^{\infty} \lambda_i^2 |x_i(s)|^2$$

for alle $s \in G$. Together with the Cauchy-Schwarz inequality, (2.63) and Theorem 2.37 this implies that

$$\begin{aligned} 0 \ &\le \ \left( \sum_{i=n}^{\infty} |\lambda_i \,\langle\, x, x_i \,\rangle\, x_i(s)| \right)^2 \ \le \ \sum_{i=n}^{\infty} |\,\langle\, x, x_i \,\rangle\,|^2 \sum_{i=1}^{\infty} \lambda_i^2 |x_i(s)|^2 \\ &= \ \sum_{i=n}^{\infty} |\,\langle\, x, x_i \,\rangle\,|^2 \int_G |k(s,t)|^2 \, dt \ \le \ M \sum_{i=n}^{\infty} |\,\langle\, x, x_i \,\rangle\,|^2 \ \underset{n\to\infty}{\longrightarrow} \ 0 \end{aligned}$$

uniformly in $s \in G$. ∎

It follows from the proof above that, actually, one does not need the continuity of $k$, but only the weaker condition (2.63). A slightly stronger condition than the continuity of $k$ is needed to also guarantee absolute and uniform convergence of the kernel representation (2.59). For that proof we need some results about positive semi-definite operators:

**Definition 2.45.** Let $K \in L(H)$ be selfadjoint. $K$ is called **positive semi-definite** if $\langle Kx, x \rangle \geq 0$ for all $x \in H$. $K$ is called **positive definite** if $\langle Kx, x \rangle > 0$ for all $x \in H \setminus \{0\}$.

**Lemma 2.46.**

(a) Let $K \in L(H)$ be compact and selfadjoint with eigensystem $(\lambda_i, x_i)$. Then $K$ is positive semi-definite if and only if $\lambda_i > 0$ for all $i \in \mathbb{N}$.

(b) Let $k \in C(G \times G)$ be a Hermite kernel with induced integral operator $K$ in $L^2(G)$. If $K$ is positive semi-definite, then $k(s, s) \geq 0$ for all $s \in G$.

**Proof:**

(a) Let us assume that $\lambda_n < 0$ for some $n \in \mathbb{N}$. Then it follows for a corresponding eigenvector $x_n$ that $\langle Kx_n, x_n \rangle = \lambda_n \|x_n\|^2 < 0$, which is a contradiction to the definition of positive semi-definiteness. Since, due to the definition of an eigensystem, all $\lambda_i \neq 0$, this shows that $\lambda_i > 0$ for all $i \in \mathbb{N}$.

If, on the other hand, all $\lambda_i > 0$, then Corollary 2.38 implies that

$$\langle Kx, x \rangle = \sum_{i=1}^{\infty} \lambda_i \langle x, x_i \rangle \langle x_i, x \rangle = \sum_{i=1}^{\infty} \lambda_i |\langle x, x_i \rangle|^2 \geq 0$$

for all $x \in H$. Thus, $K$ is positive semi-definite.

(b) Since $k$ is a Hermite kernel, $k(s, s) \in \mathbb{R}$ for all $s \in G$. Assume that $k(s_0, s_0) < 0$ for some $s_0 \in G$. Then, due to the continuity of $k$,

$$\begin{aligned} \operatorname{Re} k(s,t) &\leq \frac{k(s_0,s_0)}{2} < 0 \qquad \text{for all} \qquad (s,t) \in U_\delta\,, \quad \text{with} \\ U_\delta &:= \{(s,t) \in G \times G : \|s - s_0\| < \delta, \|t - s_0\| < \delta\} \end{aligned}$$

and $\delta > 0$ sufficiently small. Let

$$x(t) := \begin{cases} 1\,, & \|t - s_0\| < \delta\,, \\ 0\,, & \|t - s_0\| \geq \delta\,. \end{cases}$$

Then $x$ is real valued. Since, due to the selfadjointness of $K$, also $\langle Kx, x \rangle \in \mathbb{R}$,

$$\begin{aligned} \langle Kx, x \rangle &= \operatorname{Re} \langle Kx, x \rangle = \int_G \int_G \operatorname{Re} k(s,t) x(t)\, dt\, x(s)\, ds \\ &= \int_{U_\delta} \operatorname{Re} k(s,t)\, d(s,t) \leq \frac{k(s_0,s_0)}{2} |U_\delta| < 0 \end{aligned}$$

holds, which is a contradiction to the positive semi-definiteness. Thus, $k(s, s) \geq 0$ for all $s \in G$.

**Theorem 2.47 (Mercer).** Let $k \in C(G \times G)$ be a Hermite kernel and assume that the induced integral operator $K : L^2(G) \to L^2(G)$ is positive semi-definite. Moreover, $(\lambda_i, x_i)_{i\in\mathbb{N}}$ is an eigensystem for the operator $K$. Then the equality in the representation (2.59) holds for all $s, t \in G$ and the convergence of the series is absolute and uniform.

**Proof:** Let us consider the kernel

$$r_n(s,t) := k(s,t) - \sum_{i=1}^{n} \lambda_i x_i(s)\overline{x_i(t)}\,, \qquad s,t \in G\,, \tag{2.64}$$

for all $n \in \mathbb{N}$. Since eigenfunctions of an integral operator with continuous kernel are continuous, the kernels $r_n$ are also continuous. Let $R_n$ denote the induced integral operators, then they are compact and selfadjoint.

Let $\lambda \in \sigma(R_n) \setminus \{0\}$ with corresponding eigenvector $x$. Due to (2.50), it then holds that

$$\lambda x = R_n x = Kx - \sum_{i=1}^{n} \lambda_i \langle\, x, x_i \,\rangle\, x_i = \sum_{i=n+1}^{\infty} \lambda_i \langle\, x, x_i \,\rangle\, x_i\,.$$

Therefore, $\langle\, x, x_j \,\rangle = \frac{1}{\lambda}\langle\, \lambda x, x_j \,\rangle = 0$ for $j \le n$ so that $\lambda x = Kx$, i.e., $\lambda \in \sigma(K) \setminus \{0\}$ and $x$ is an eigenvector of $K$. Since $x$ is orthogonal to $x_1, \ldots, x_n$, $x \in \operatorname{span}\{x_{n+1}, x_{n+2}, \ldots\}$. On the other hand, $\lambda_j x_j = Kx_j = R_n x_j$ for $j > n$ so that $\lambda_j \in \sigma(R_n) \setminus \{0\}$ and $x_j$ is an eigenvector of $R_n$. Altogether this implies that $(\lambda_i, x_i)_{i>n}$ is an eigensystem for $R_n$.

Lemma 2.46 (a) implies that $R_n$ is positive semi-definite and Lemma 2.46 (b) implies that

$$r_n(s,s) \ge 0\,, \qquad s \in G\,.$$

Together with (2.64) we obtain that

$$\sum_{i=1}^{n} \lambda_i |x_i(s)|^2 \le k(s,s)\,, \qquad s \in G\,,\, n \in \mathbb{N}\,. \tag{2.65}$$

Integrating this formula with respect to $s$ and using that $\|x_i\| = 1$ yields

$$0 < \sum_{i=1}^{n} \lambda_i \le \int_G k(s,s)\,ds\,, \qquad n \in \mathbb{N}\,.$$

Since all $\lambda_i > 0$, this together with (2.65) implies the convergence of the series $\sum_{i=1}^{\infty} \lambda_i$ and $\sum_{i=1}^{\infty} \lambda_i |x_i(s)|^2$ for all $s \in G$.

Let now $\varepsilon > 0$ and $s, t \in G$ be arbitrary, but fixed. Then an $n_0(s) \in \mathbb{N}$ exists such that

$$\sum_{i=n+1}^{m} \lambda_i |x_i(s)|^2 \le \varepsilon \qquad \text{for all} \qquad m > n \ge n_0(s).$$

Together with (2.65) we now obtain that

$$\begin{aligned}\left(\sum_{i=n+1}^{m}\lambda_i|x_i(s)\overline{x_i(t)}|\right)^2 &\leq \left(\sum_{i=n+1}^{m}\lambda_i|x_i(s)|^2\right)\left(\sum_{i=n+1}^{m}\lambda_i|x_i(t)|^2\right)\\ &\leq \varepsilon k(t,t) \leq \varepsilon \|k\|_\infty \end{aligned} \tag{2.66}$$

for all $m > n \geq n_0(s)$. Since $n_0(s)$ does not depend on $t$, it follows from the Cauchy criterion that the series $\tilde{k}(s,t) := \sum_{i=1}^{\infty}\lambda_i x_i(s)\overline{x_i(t)}$ converges absolutely and uniformly in $t$ for all $s \in G$. But this implies that $\tilde{k}$ is continuous in $t$ for all $s$.

Now Theorem 2.44 implies that

$$\begin{aligned}\int_G \tilde{k}(s,t)x(t)\,dt &= \sum_{i=1}^{\infty}\lambda_i\,\langle\, x, x_i\,\rangle\, x_i(s) = \int_G k(s,t)x(t)\,dt\\ &\Longrightarrow \int_G (k(s,t)-\tilde{k}(s,t))x(t)\,dt = 0\end{aligned}$$

for all $s \in G$ and all $x \in L^2(G)$. Choosing $x(t) = x_s(t) := \overline{k(s,t)-\tilde{k}(s,t)}$ together with the continuity of $k$ yields that $k(s,t) = \tilde{k}(s,t)$ for all $s,t \in G$.

Therefore, the sequence of continuous functions $(\sum_{i=1}^{n}\lambda_i|x_i(s)|^2)_{n\in\mathbb{N}}$ converges monotonically to the continuous function $k(s,s)$. Hence, due to Dini's theorem, the convergence is uniform. But this implies that the Cauchy condition (2.66) holds uniformly also with respect to $s$ yielding the absolute and uniform convergence in (2.59) with respect to $(s,t)$. ∎

**Remark 2.48.** One can see from the proof above that the results of the theorem of Mercer are still valid if instead of the positive semi-definiteness of $K$ one requires that only finitely many eigenvalues are either negative or positive. In this form Mercer's theorem will be used in Chapter 5.

Using Mercer's theorem, the fact that $\|x_i\| = 1$, and elementwise integration in (2.59), which is allowed due to the uniform convergence, yields

$$\sum_{i=1}^{\infty}\lambda_i = \int_G k(s,s)\,ds\,. \tag{2.67}$$

This number is called **trace** of the integral operator. Under the conditions of Mercer's theorem the sequence of eigenvalues is not only in $l^2$, but even in $l^1$. This last property characterizes the subclass of so-called **nuclear** or **trace class** operators.

For general $L^2$-kernels the right-hand side in (2.67) makes no sense.

**Remark 2.49.** A class of integral operators having real eigenvalues are selfadjoint compact operators according to Theorem 2.34. Another class are those with a positive kernel, according to the following theorem due to Jentzsch:

Let $k \in C([0,1]^2)$, $k(s,t) > 0$ for all $s,t \in [0,1]$, with induced integral operator $K$. Then $K$ has at least one positive eigenvalue. If $\lambda$ is the largest positive eigenvalue, then

$\mathcal{N}(\lambda I - K) = \operatorname{span}\{x\}$, where $x(s) > 0$ for all $s \in [0,1]$. Moreover $|\mu| < \lambda$ for all $\mu \in \sigma(K) \setminus \{\lambda\}$.

**Remark 2.50.** A simple constructive method to solve a Fredholm integral equation (2.1), i.e., an equation (2.51), where $K$ is the integral operator induced by an $L^2$-kernel $k$ in $L^2(G)$, is the method of successive approximation:

Thereto, one rewrites (2.51) into the equivalent fixpoint form

$$x = \frac{1}{\lambda}(Kx + f)$$

and performs an iteration according to

$$x_{n+1} := \frac{1}{\lambda}(Kx_n + f)\,, \qquad n \in \mathbb{N}\,, \tag{2.68}$$

with $x_0 := 0$. It is an immediate consequence of the fixpoint theorem due to Banach that $(x_n)$ converges to a solution of (2.51) if

$$|\lambda| > \|K\|\,. \tag{2.69}$$

Note that then the iteration operator $x \mapsto \frac{1}{\lambda}(Kx + f)$ satisfies a Lipschitz condition with Lipschitz constant $\left\|\frac{1}{\lambda}K\right\| < 1$. (2.69) can only be satisfied for $\lambda \notin \sigma(K)$.

A simple calculation shows that $x_n$ defined by (2.68) is given by $x_n = \sum_{i=0}^{n-1} \lambda^{-i-1} K^i f$. Thus, in case of (2.69), the solution $x$ of (2.51) is given by the Neumann series

$$x = \sum_{i=0}^{\infty} \lambda^{-i-1} K^i f\,. \tag{2.70}$$

The operators $K^i$ are again integral operators with so-called iterated kernels $k_i$ satisfying the recursion formula

$$k_i(s,t) = \int_G k(s,\tau) k_{i-1}(\tau,t)\,d\tau\,, \qquad i \geq 2\,, \tag{2.71}$$

with $k_1 = k$. Assuming that (2.69) holds, (2.70) implies that the series

$$r\left(s,t,\frac{1}{\lambda}\right) := \sum_{i=1}^{\infty} \lambda^{-i} k_i(s,t) \tag{2.72}$$

converges in the $L^2$-sense.

Using (2.71), one can show that the convergence above is even absolute and uniform if $k$ is continuous and satisfies instead of (2.69) the stronger condition

$$|\lambda| > \|k\|_2\,|G|\,,$$

i.e., $r(s,t,\frac{1}{\lambda})$ is then continuous.

(2.70) implies that the solution $x$ of (2.51) is given by

$$x(s) = \frac{1}{\lambda} f(s) + \frac{1}{\lambda} \int_G r\left(s,t,\frac{1}{\lambda}\right) f(t)\,dt\,, \qquad s \in G\,.$$

Therefore, the function $r$ is called **resolvent kernel**.

The series (2.70) does not only converge under condition (2.69) but also under the (in the non-selfadjoint case) slightly weaker condition

$$|\lambda| > \rho(K) := \sup\{|\mu| : \mu \in \sigma(K)\}$$

uniformly with respect to $f$ on bounded sets, i.e., in the operator norm. $\rho$ is called **spectral radius** of $K$.

So far we have only dealt with compact integral operators. In practice also non-compact operators appear. We present some examples and prove the non-compactness by showing that the spectrum does not satisfy the properties of compact operators.

**Example 2.51.** Let $X$ be a space of continuous functions on $[0, +\infty)$ containing the functions

$$x_\alpha(s) := \frac{s}{\alpha^2 + s^2} + \sqrt{\frac{\pi}{2}} e^{-\alpha s}\,, \qquad s \geq 0\,,$$

for all $\alpha > 0$. Moreover, we assume that $X$ is such that the operator $T$, defined by

$$Tx(s) := \int_0^\infty \sqrt{\frac{2}{\pi}} \sin(st) x(t)\, dt \in X\,, \quad s > 0\,, \tag{2.73}$$

where the integral is an improper integral, and

$$Tx(0) := \lim_{s \to 0} (Tx)(s)$$

maps $X$ into $X$.

One can show that

$$\int_0^\infty e^{-\alpha t} \sin(st)\, dt = \frac{s}{s^2 + \alpha^2} \quad \text{and} \quad \int_0^\infty \frac{t \sin(st)}{\alpha^2 + t^2}\, dt = \frac{\pi}{2} e^{-\alpha s}\,, \qquad \alpha > 0\,,\ s > 0\,.$$

For the second integral one needs the residue theorem (exercise !). This implies that

$$Tx_\alpha = x_\alpha\,, \qquad \alpha > 0\,.$$

Therefore, 1 is an eigenvalue of $T$ with infinite multiplicity, since all $x_\alpha$ are linearly independent for $\alpha > 0$. According to Theorem 2.33 (b), $T$ cannot be compact.

If one replaces the infinte integration intervall in (2.73) by the finite interval $[a, b]$, $0 \leq a < b < \infty$, the integral operator will be compact on $C[a, b]$ or $L^2[a, b]$. Thus, such an exchange of the integration interval radically changes the properties of the operator.

**Example 2.52.** Let $X$ be a space of functions on $\mathbb{R}$ containing all functions $\cos(\alpha s)$ and $\sin(\alpha s)$ for $\alpha > 0$ and where the operator $T : X \to X$, defined by

$$(Tx)(s) := \int_{-\infty}^{+\infty} e^{-|s-t|} x(t)\, dt$$

makes sense. A straight forward calculation shows that

$$\int_{-\infty}^{+\infty} e^{-|s-t|} e^{i\alpha t}\, dt = \frac{2e^{i\alpha s}}{1 + \alpha^2}\,, \qquad \alpha > 0\,,\ s \in \mathbb{R}\,.$$

Splitting into the real and imaginary part yields that for all $\alpha > 0$ and $s \in \mathbb{R}$

$$\begin{aligned} \int_{-\infty}^{+\infty} e^{-|s-t|} \cos(\alpha t)\, dt &= \frac{2}{1+\alpha^2} \cos(\alpha s)\,, \\ \int_{-\infty}^{+\infty} e^{-|s-t|} \sin(\alpha t)\, dt &= \frac{2}{1+\alpha^2} \sin(\alpha s)\,. \end{aligned}$$

Therefore, $\frac{2}{1+\alpha^2}$ is an eigenvalue of $T$ for all $\alpha > 0$, i.e., the whole interval $(0, 2)$ consists of eigenvalues. Due to Theorem 2.33 (c), $T$ cannot be compact. Again the operator will be compact with a radically changed spectrum if the infinte integration interval is replaced by a finite one.

**Remark 2.53.** There are many important examples of non-compact integral operators:

(a) Let $k \in L^1(\mathbb{R})$. The convolution operator (see Example 2.52)

$$(Tx)(s) := \int_{-\infty}^{+\infty} k(s-t)x(t)\, dt$$

is well-defined on $L^2(\mathbb{R})$ (why ?). The spectrum is given by

$$\overline{\{\lambda \in \mathbb{R} : \lambda = k(\mu) \text{ for some } \mu \in \mathbb{R}\}}\,,$$

which, in general, will be an uncountable set. If $\lambda$ is such that $k(\mu) = \lambda$ for all $\mu \in [a, b]$ for some $a < b \in \mathbb{R}$, then the eigenspace to this eigenvalue $\lambda$ is infinite-dimensional.

(b) The strongly singular integral operator

$$(Tx)(s) := \int_{-\infty}^{+\infty} \frac{x(t)}{s-t}\, dt$$

can be well-defined on $L^2(\mathbb{R})$ using the Fourier transform. $T$ is called **Hilbert transform**. Its spectrum consists only of $i\pi$ and $-i\pi$. Both values are eigenvalues with infinite-dimensional eigenspaces. The finite Hilbert transform on $L^2[-1, 1]$ is defined by

$$(Tx)(s) := \int_{-1}^{1} \frac{x(t)}{s-t}\, dt\,.$$

Its spectrum is given by $\{i\mu : \mu \in [-\pi, \pi]\}$, i.e., it is also non-compact.

It follows from Definition 2.12 that in one dimension ($N = 1$) the kernel $(s-t)^{-\alpha}$ is weakly singular for $\alpha < 1$ and, hence, the induced integral operator is compact. This example shows that already in the limiting case $\alpha = 1$ the compactness is lost.

# 3. Numerical solution of Fredholm equations

In this chapter, we discuss methods for the numerical solution of linear Fredholm integral equations of the second kind, namely: approximation with degenerate kernels, projection and collocation methods, and quadrature rule methods.

We consider the equation (2.1), i.e.,

$$\lambda x(s) - \int_G k(s,t)x(t)\,dt = f(s)\,, \qquad s \in G\,,$$

with $\lambda \neq 0$ and $G$ as in Chapter 2. The kernel $k$ is either continuous, weakly singular or quadratically integrable so that the induced integral operator $K$ is compact on $C(G)$ or $L^2(G)$, respectively.

## 3.1. Degenerate kernel approximation

As an approximation problem for the above equation we consider equations

$$\lambda x_n(s) - \int_G k_n(s,t)x_n(t)\,dt = f(s)\,, \qquad s \in G\,, \tag{3.1}$$

where the kernels $k_n$ are degenerate (see Definition 2.2), i.e.,

$$k_n(s,t) = \sum_{i=1}^{n} \varphi_i(s)\psi_i(t) \qquad \text{a.e.}$$

The functions $\varphi_i$ and $\psi_i$, in general, also depend on $n$. The kernels $k_n$ should be constructed in such a way that

$$\lim_{n\to\infty} \|K - K_n\| = 0\,, \tag{3.2}$$

where $K_n$ are the integral operators induced by $k_n$ and $\|\cdot\|$ is the operator norm. Note that (3.2) can only be satisfied for compact operators $K$ (see Theorem 2.8 and Theorem 2.3). The convergence of the approximate solutions is based on the following theorem:

**Theorem 3.1.** Let $K$ and $K_n$ be as above (considered as operators from $C(G)$ or $L^2(G)$ into iteself) such that (3.2) holds.

(a) If $\lambda \notin \sigma(K)$, then $\lambda I - K_n$ is continuously invertible for $n \in \mathbb{N}$ sufficiently large. If $x$ and $x_n$ are the unique solutions of (2.1) and (3.1), respectively, then

$$\|x - x_n\| \leq \frac{\|(\lambda I - K)^{-1}\|}{1 - \|(\lambda I - K)^{-1}\|\,\|K - K_n\|}\,\|Kx - K_n x\|\,. \tag{3.3}$$

(b) Let $\lambda \notin \sigma(K_n)$ for some $n \in \mathbb{N}$ with

$$\|K - K_n\| < \left\|(\lambda I - K_n)^{-1}\right\|^{-1}.$$

Then $\lambda \notin \sigma(K)$. If $x$ and $x_n$ are the unique solutions of (2.1) and (3.1), respectively, then

$$\|x - x_n\| \leq \frac{\|(\lambda I - K_n)^{-1}\|}{1 - \|(\lambda I - K_n)^{-1}\| \, \|K - K_n\|} \|Kx_n - K_n x_n\| \,. \tag{3.4}$$

**Proof:** The proof immediately follows from the theorem on inverses of neighbouring operators applied to $\lambda I - K$ and $\lambda I - K_n$, respectively. ∎

**Remark 3.2.** In the theorem above only condition (3.2) is important. The concrete form of $K_n$ or $K$ is not essential. In contrast to (3.3), (3.4) is a computable error estimate, since only the approximation $x_n$ and not the exact solution $x$ is used.

It was already shown in Remark 2.4, how the solution $x_n$ of (3.1) can be calculated by solving a linear system. Note that (2.5) is always uniquely solvable if $\lambda \notin \sigma(K_n)$ and if the functions $\varphi_1, \ldots, \varphi_n$ are linearly independent.

We now present some methods for constructing sequences of degenerate kernels so that (3.2) holds.

**Method 3.3 (Expansion by eigenfunctions).** If $k$ is a Hermite $L^2$-kernel, then Theorem 2.41 implies that

$$k(s,t) = \sum_{i=1}^{\infty} \lambda_i x_i(s) \overline{x_i(t)}$$

in the $L^2$-sense. Defining

$$k_n(s,t) := \sum_{i=1}^{n} \lambda_i x_i(s) \overline{x_i(t)} \,,$$

we obtain that

$$\|K - K_n\|^2_{L(L^2(G))} \leq \sum_{i=n+1}^{\infty} \lambda_i^2 \,.$$

Since $(\lambda_i) \in l^2$, (3.2) holds.

In this case the linear system (2.5) has the special form $a_{ij} = \lambda_i \delta_{ij}$. Therefore, the approximate solution $x_n$ is given by

$$x_n = \frac{1}{\lambda} f + \sum_{i=1}^{n} \frac{\lambda_i \langle f, x_i \rangle}{\lambda(\lambda - \lambda_i)} x_i$$

if $\lambda \neq \lambda_i$ for $i = 1, \ldots, n$. This is identical to the second formula of (2.52) applied to $K_n$.

**Method 3.4 (Expansion by orthonormal systems).** Let $k \in L^2(G \times G)$ and let $(\varphi_i)$ be an arbitrary complete orthonormal system in $L^2(G)$. Since $k(\cdot, t) \in L^2(G)$ for almost all $t \in G$, it holds that

$$k(s,t) = \sum_{i=1}^{\infty} \varphi_i(s) \psi_i(t) \qquad \text{a.e.}$$

with $\psi_i(t) := \langle\, k(\cdot,t), \varphi_i \,\rangle = \int_G k(s,t)\overline{\varphi_i(s)}\,ds$, i.e.,

$$\psi_i = \overline{K^*\varphi_i}\,. \tag{3.5}$$

We now approximate $k$ by degenerate kernels

$$k_n(s,t) := \sum_{i=1}^{n} \varphi_i(s)\psi_i(t)\,.$$

Due to (3.5) the coefficients of the matrix $A$ in (2.6) are given by

$$a_{ij} = \langle\, \varphi_i, K^*\varphi_j \,\rangle = \langle\, K\varphi_i, \varphi_j \,\rangle\,.$$

Now

$$((K-K_n)x)(s) = \int_G \sum_{i=n+1}^{\infty} \varphi_i(s)\psi_i(t)x(t)\,dt = \sum_{i=n+1}^{\infty} \langle\, x, K^*\varphi_i \,\rangle\, \varphi_i(s)$$

implies that

$$\|K-K_n\|^2_{L(L^2(G))} \le \sum_{i=n+1}^{\infty} \|K^*\varphi_i\|_2^2\,. \tag{3.6}$$

Since $K^*$ as integral operator with an $L^2$-kernel is a Hilbert-Schmidt operator, $\sum_{i=1}^{\infty}\|K^*\varphi_i\|_2^2 < \infty$ (see Remark 2.42) so that the estimate on the right-hand side of (3.6) goes to 0 with $n\to\infty$, i.e., (3.2) holds.

Instead of expanding the kernel with respect to the variable $s$, one could look for an expansion with respect to $t$ or with respect to both variables using the complete orthonormal system $(\varphi_i(s)\overline{\varphi_j(t)})$ in $L^2(G\times G)$. Obviously, Method 3.3 is a special case of Method 3.4.

**Method 3.5 (Taylor series approximation).** We only deal with the one-dimensional case $G := [0,1]$ and assume that the kernel $k(s,t)$ only depends on the product $s\cdot t$, i.e., $k(s,t) = r(s\cdot t)$, where $r$ has a Taylor series expansion in $[0,1]$ at 0, i.e.,

$$r(u) = \sum_{i=0}^{\infty} \alpha_i u^i\,, \qquad u\in[0,1]\,, \qquad \alpha_i = \frac{r^{(i)}(0)}{i!}\,.$$

Defining

$$k_n(s,t) := \sum_{i=1}^{n} \alpha_{i-1}s^{i-1}t^{i-1}$$

yields the following formulas for the coefficients in (2.6):

$$a_{ij} = \frac{\alpha_{i-1}}{i+j-1}\,, \qquad \bar{f}_i = \alpha_{i-1}\int_0^1 s^{i-1}f(s)\,ds\,.$$

A straightforward calculation shows that (3.2) holds if

$$\sum_{i=1}^{\infty} |\alpha_i|^2(2i+1)^{\varepsilon-\frac{1}{2}} < \infty \quad \text{for some} \quad \varepsilon > 0\,.$$

**Method 3.6 (Approximation by interpolation).** Again we only deal with the one-dimensional case $G := [0,1]$ and assume that $k \in C(G \times G)$. We interpolate the kernel with respect to $s$ for all $t$, e.g., with linear splines. Using the uniform knots $s_i = \frac{i}{n}$, $i \in \{0, \ldots, n\}$ and defining the functions

$$l_i(s) := \begin{cases} n(s - s_{i-1}), & s \in [s_{i-1}, s_i], \\ n(s_{i+1} - s), & s \in [s_i, s_{i+1}], \\ 0, & \text{else}, \end{cases}$$

we approximate $k$ by

$$k_n(s,t) := \sum_{i=0}^{n} k(s_i, t) l_i(s), \qquad s, t \in [0,1].$$

This is equivalent to

$$k_n(s,t) = n((s_i - s)\, k(s_{i-1}, t) + (s - s_{i-1})\, k(s_i, t)), \qquad s \in [s_{i-1}, s_i],\ t \in [0,1].$$

The coefficients in (2.6) are given by

$$\begin{aligned} a_{ij} &= (Kl_i)(s_j) = \int_{\max\{0, s_{i-1}\}}^{\min\{1, s_{i+1}\}} k(s_j, t) l_i(t)\, dt, \\ \bar{f}_i &= (Kf)(s_i) = \int_0^1 k(s_i, t) f(t)\, dt. \end{aligned}$$

$i, j \in \{0, \ldots, n\}$. If the kernel $k$ is smooth enough, then (3.2) holds. For instance, if $k(., t) \in C^2[0,1]$, then

$$\|K - K_n\|_{L(C[0,1])} \leq \frac{1}{8n^2} \int_0^1 \max_{s \in [0,1]} \left| \frac{\partial^2 k(s,t)}{\partial s^2} \right| dt.$$

**Remark 3.7.** Since, under the conditions of Theorem 3.1, $\|(\lambda I - K_n)^{-1}\|$ is uniformly bounded, the convergence rate of $\|x - x_n\|$ is determined by $\|K - K_n\|$ or, more precisely, by $\|Kx_n - K_n x_n\|$ and $\|Kx - K_n x\|$, respectively. To obtain a computable a-posteriori error estimate one should not forget that estimate (3.4) not only contains $\|K - K_n\|$, but also $\|(\lambda I - K_n)^{-1}\|$. The last expression can be estimated as follows:

Let $A$ be the matrix of (2.6) and let

$$(\lambda I - A)^{-1} =: (b_{ij})_{1 \leq i,j \leq n}.$$

Then (2.5) – (2.7) imply that

$$x_n(s) = \frac{1}{\lambda} \left( f(s) + \sum_{i,j=1}^{n} b_{ij} \int_G f(t) \psi_j(t)\, dt\, \varphi_i(s) \right).$$

This together with $x_n = (\lambda I - K_n)^{-1} f$ implies that

$$\left\| (\lambda I - K_n)^{-1} f \right\|_\infty \leq \frac{1}{|\lambda|} \|f\|_\infty \left( 1 + \max_{s \in G} \int_G \left| \sum_{i,j=1}^{n} b_{ij} \psi_j(t) \varphi_i(s) \right| dt \right),$$

i.e., setting $M_n := \max_{1\le i,j\le n} |b_{ij}|$,

$$\left\|(\lambda I - K_n)^{-1}\right\|_{L(C(G))} \le \frac{1}{|\lambda|}\left(1 + M_n \max_{s\in G}\int_G \sum_{i,j=1}^{n} |\psi_j(t)\varphi_i(s)|\,dt\right).$$

In a similar way, one can derive an estimate for $\|(\lambda I - K_n)^{-1}\|_{L(L^2(G))}$.

The major disadvantage of degenerate kernel methods is that the integrations needed for the calculation of the coefficients in (2.6) can be quite costly. In Method 3.4 even double integrals have to be calculated. One advantage of Method 3.6 is that the integrals for the computation of $a_{ij}$ only have to be evaluated over the support of the $j$-th basis spline.

If $k$ or $f$ have a special form, quite often the indefinite integrals necessary for the calculation of the coefficients $a_{ij}$ and $\bar{f}_i$ can be computed explicitly. Here computer algebra methods are beneficial. In all other cases, the integrals have to be approximated using quadrature rules leading to additional errors.

## 3.2. Projection methods

We start by formulating projection methods for the abstract version of (2.1), i.e., (see (2.51))

$$\lambda x - Kx = f\,,$$

where $K \in L(X)$ is compact on the Banach space $X$ (usually $X = C(G)$ or $X = L^2(G)$), $f \in X$, $\lambda \notin \sigma(K)$. For each $n \in \mathbb{N}$ let $X_n$ be a finite-dimensional (in general, $n$-dimensional) subspace of $X$, $P_n : X \to X_n$ a bounded linear projector. The $n$-th approximation of this method is defined by the equation

$$\lambda x_n - P_n K x_n = P_n f, \qquad x_n \in X_n. \tag{3.7}$$

Note that for $\lambda \neq 0$ each solution of $\lambda x - P_n K x = P_n f$ is automatically in $X_n$. The convergence analysis is again based on the theorem on inverses of neighbouring operators:

**Theorem 3.8.** Let $K \in L(X)$ be compact, $\lambda \notin \sigma(K)$, $P_n : X \to X_n$ as above. If

$$\|P_n K - K\| < \left\|(\lambda I - K)^{-1}\right\|^{-1}, \tag{3.8}$$

then $\lambda \notin \sigma(P_n K)$ and (3.7) has a unique solution. If $x$ and $x_n$ are the solutions of (2.51) and (3.7), respectively, then

$$\|x - x_n\| \le |\lambda| \left\|(\lambda I - P_n K)^{-1}\right\| \|x - P_n x\| \tag{3.9}$$

and

$$\left\|(\lambda I - P_n K)^{-1}\right\| \le \frac{\|(\lambda I - K)^{-1}\|}{1 - \|(\lambda I - K)^{-1}\|\,\|P_n K - K\|}\,. \tag{3.10}$$

If

$$\lim_{n\to\infty} P_n z = z \qquad \text{for all} \quad z \in X\,, \tag{3.11}$$

then (3.8) holds for all $n \in \mathbb{N}$ sufficiently large.

**Proof:** The last assertion follows from Lemma 2.9. There it was shown that for a compact $K$ (3.11) implies that

$$\lim_{n\to\infty} \|P_n K - K\| = 0\,. \tag{3.12}$$

The invertibility of $(\lambda I - P_n K)$ and (3.10) follows from the theorem on inverses of neighbouring operators applied to $\lambda I - K$ and $\lambda I - P_n K$. It is an immediate consequence of (2.51) and (3.7) that

$$(\lambda I - P_n K)(x - x_n) = P_n(\lambda x - Kx) + \lambda(x - P_n x) - P_n f = \lambda(x - P_n x)$$

and, hence,

$$x - x_n = \lambda(\lambda I - P_n K)^{-1}(x - P_n x)$$

yielding (3.9). ∎

**Remark 3.9.** Estimate (3.9) implies that the convergence rate of $\|x - x_n\|$ is determined by $\|x - P_n x\|$. If $X$ is a Hilbert space and $P_n$ the orthogonal projection, then the convergence rate is optimal in the sense that no better rate can be obtained by an approximation with elements from $X_n$, since then $\|x - P_n x\| = \inf_{z\in X_n} \|x - z\|$. As in Theorem 3.1 one could interchange the roles of $K$ and $P_n K$ in Theorem 3.8 to obtain an assertion, where one conclude from the unique solvability of (3.7) for $n$ sufficiently large to the one of (2.51).

**Method 3.10 (Galerkin method).** In this special projection method $X = L^2(G)$ and $X_1 \subseteq X_2 \subseteq X_3 \subseteq \ldots$ is a sequence of finite-dimensional subspaces with

$$\overline{\bigcup_{n\in\mathbb{N}} X_n} = X\,. \tag{3.13}$$

For each $n \in \mathbb{N}$ let $P_n$ be the orthogonal projector onto $X_n$. The solution of (3.7) can be calculated as follows:

Let $\{\varphi_1, \ldots, \varphi_n\}$ be a basis of $X_n$. Then (3.7) is equivalent to $P_n(\lambda x_n - Kx_n - f) = 0$ and $x_n \in X_n$, i.e.,

$$\lambda x_n - Kx_n - f \in X_n^{\perp}, \qquad x_n \in X_n\,,$$

which is again equivalent to

$$\langle\, \lambda x_n - Kx_n - f, \varphi_j \,\rangle = 0 \qquad \text{for} \quad j \in \{1, \ldots, n\}\,, \qquad x_n \in X_n\,. \tag{3.14}$$

Since $x_n \in X_n$ has a representation

$$x_n = \sum_{i=1}^{n} \bar{x}_i \varphi_i\,, \tag{3.15}$$

(3.14) implies that the coefficients $\bar{x} := (\bar{x}_1, \ldots, \bar{x}_n)^\top$ can be computed as solution of the linear system

$$(\lambda B_n - M_n^\top)\bar{x} = \bar{f}_n \tag{3.16}$$

with

$$B_n := (\langle\, \varphi_i, \varphi_j \,\rangle)_{1\le i,j\le n}\,, \qquad M_n := (\langle\, K\varphi_i, \varphi_j \,\rangle)_{1\le i,j\le n}\,,$$
$$\bar{f}_n := (\langle\, f, \varphi_1 \,\rangle\,, \ldots, \langle\, f, \varphi_n \,\rangle)^\top\,.$$

If (3.8) holds, equation (3.16) has a unique solution. Due to (3.13), $P_n z \to z$ for all $z$ in the dense subset $\bigcup_{n\in\mathbb{N}} X_n$. Since $\|P_n\| = 1$ for all $n \in \mathbb{N}$, the Banach-Steinhaus theorem implies (3.11) and, hence, (3.12).

An essential disadvantage of the Galerkin method is that integrals have to be calculated for the computation of the elements of the matrices $B_n, M_n$ and the right-hand side $\bar{f}_n$.

**Method 3.11 (Collocation).** Here $X = C(G)$. For $n \in \mathbb{N}$ let $s_1, \ldots, s_n$ be pairwise different points in $G$ and let $\varphi_1, \ldots, \varphi_n \in X$ be such that the matrix

$$B_n := (\varphi_i(s_j))_{1\le i,j\le n} \tag{3.17}$$

is regular. Moreover, let $X_n := \operatorname{span}\{\varphi_1, \ldots, \varphi_n\}$ and let $P_n : X \to X_n$ be the projector defined as follows: $P_n z$ is the element $z_n \in X_n$ with $z_n(s_i) = z(s_i)$ for all $i \in \{1, \ldots, n\}$. Note that, due to the regularity of $B_n$, $z_n$ is unique.

It follows from simple facts of interpolation theory that $P_n$ is bounded. The norm of $P_n$ is given by

$$\|P_n\| = \max_{s\in G} \sum_{i=1}^{n} |l_i(s)|\,,$$

where $l_1, \ldots, l_n \in X_n$ is the so-called cardinal basis, i.e., $l_i(s_j) = \delta_{ij}$, $i, j \in \{1, \ldots, n\}$.

Then the approximate solution $x_n$ according to (3.7) is again given by (3.15), where the coefficients $\bar{x} := (\bar{x}_1, \ldots, \bar{x}_n)^\top$ are determined as follows:

$$\lambda x_n(s_j) - (Kx_n)(s_j) = f(s_j)\,, \qquad j \in \{1, \ldots, n\}\,.$$

This again yields equation (3.16), where $B_n$ is as in (3.17) and

$$M_n := ((K\varphi_i)(s_j))_{1\le i,j\le n}\,, \qquad \bar{f}_n := (f(s_1), \ldots, f(s_n)^\top\,.$$

Depending on the choice of the functions $\varphi_i$ one speaks of polynomial or spline collocation. In polynomial collocation one very often uses the zeros of Tschebyscheff polynomials as collocation points. Note that then condition (3.11) is not satisfied. However, it is still possible that (3.12) holds (see [2]). The major problem in this method is again the calculation of integrals. Here computer algebra methods might be helpful.

## 3.3. Quadrature rule methods

**Method 3.12 (Nyström method).** We present this method only for the one-dimensional case $G := [0,1]$, i.e., we consider

$$\lambda x(s) - \int_0^1 k(s,t)x(t)\,dt = f(s)\,, \qquad s \in [0,1]\,, \tag{3.18}$$

where $k$, $f$, and $x$ are continuous; the corresponding integral operator $K$ is defined on $C[0,1]$.

The first step of the Nyström method consists of approximating the integral operator in (3.18) using a quadrature rule: let

$$Q_n x := \sum_{j=1}^{n} \omega_j x(t_j) \approx \int_0^1 x(t)\,dt \tag{3.19}$$

be a quadrature rule for $x \in C[0,1]$ with nodes $t_1, \ldots, t_n \in [0,1]$ and weights $\omega_1, \ldots, \omega_n$; actually, the $t_j$ and $\omega_j$ also depend on $n$. Moreover, we define

$$(K_n x)(s) := Q_n(k(s,\cdot)x) = \sum_{j=1}^{n} \omega_j k(s,t_j)x(t_j) \tag{3.20}$$

for $x \in C[0,1]$ and $s \in [0,1]$. Then we look for a solution of the equation

$$\lambda x_n - K_n x_n = f\,. \tag{3.21}$$

The **full discretization** step consists in choosing $s = t_i$, $i \in \{1, \ldots, n\}$, in (3.21), i.e.,

$$\lambda x_n(t_i) - \sum_{j=1}^{n} \omega_j k(t_i,t_j)x_n(t_j) = f(t_i)\,, \qquad i \in \{1,\ldots,n\}\,. \tag{3.22}$$

Setting the vector $\bar{x}_n = (\bar{x}_{n,1}, \ldots, \bar{x}_{n,n}) := (x_n(t_1), \ldots, x_n(t_n))$ and defining

$$M_n := \omega_j k(t_i,t_j)_{1\le i,j\le n}\,, \qquad \bar{f}_n := (f(t_1), \ldots, f(t_n))^\top\,,$$

(3.22) is equivalent to the linear system

$$\lambda \bar{x}_n - M_n \bar{x}_n = \bar{f}_n\,. \tag{3.23}$$

Defining

$$x_n(s) := \frac{1}{\lambda}\left(f(s) + \sum_{j=1}^{n} \omega_j k(s,t_j)\bar{x}_{n,j}\right)\,, \qquad s \in [0,1]\,,$$

a straightforward calculation shows that $x_n$ really solves (3.21).

The importance of this fact is that one gets the solution $x_n$ of the infinite-dimensional problem (3.21) by solving the finite-dimensional problem (3.23).

Formally, (3.21) has the same form as the abstract version of (3.1) so that one might think that Theorem 3.1 is applicable. Unfortunately, this is not the case, since (3.2) does not hold anymore. This gave the impulse for studying collectively compact operators (see [1]) that allow convergence results, even when the approximation operators do not converge in norm.

**Definition 3.13.** Let $X$ be a Banach space and let $\mathcal{K}$ be a set of linear operators from $X$ into iteself. $\mathcal{K}$ is called **collectively compact** if for any bounded set $B \subseteq X$ the set $\overline{\mathcal{K}}(B) := \overline{\bigcup_{K\in\mathcal{K}} K(B)}$ is compact.

Since every operator of a collectively compact set is compact, it is also bounded. As in the proof of Theorem 2.8 (b) one can see that for a collectively compact set $\mathcal{K}$ it even holds that

$$\sup_{K\in\mathcal{K}} \|K\| < \infty\,.$$

The following convergence concept will be important:

**Definition 3.14.** Let $(K_n)$ be a sequence of linear operators and let $K$ be a linear operator on $X$. $(K_n)$ **converges collectively compact** to $K$, denoted by $K_n \xrightarrow{\text{cc}} K$, if $(K_n)$ converges pointwise to $K$ and $\{K_n : n \in \mathbb{N}\}$ is collectively compact.

This convergence concept has the following important properties in common with norm convergence that do not hold for simple pointwise convergence:

**Lemma 3.15.** If $K_n \xrightarrow{\text{cc}} K$, then $K$ is compact.

**Proof:** Let $B \subseteq X$ be bounded. Since $Kx = \lim_{n\to\infty} K_n x$ for all $x \in B$, it follows that $\overline{K(B)} \subseteq \overline{\bigcup_{n\in\mathbb{N}} K_n(B)}$. This together with the collective compactness of $\{K_n : n \in \mathbb{N}\}$ yields the compactness of $K$. ∎

The following assertion serves, in a way, as a substitute for (3.2):

**Lemma 3.16.** If $K_n \xrightarrow{\text{cc}} K$, then

$$\lim_{n\to\infty} \|(K - K_n)K\| = 0 \quad \text{and} \quad \lim_{n\to\infty} \|(K - K_n)K_n\| = 0\,. \tag{3.24}$$

**Proof:** Since $K_n \to K$ pointwise, Lemma 2.9 implies uniform convergence on the compact sets $\overline{K(U)}$ and $\overline{\bigcup_{n\in\mathbb{N}} K_n(U)}$, where $U$ is the unit ball in $X$. This implies that $(K - K_n)K$ and $(K - K_n)K_n$ converge uniformly to 0 on $U$ which yields (3.24). ∎

We also need a version of the theorem on inverses of neighbouring operators based on (3.24) instead of (3.2):

**Lemma 3.17.** Let $A$ be compact, $T \in L(X)$, $\lambda \neq 0$, and $\lambda \notin \sigma(T)$. Moreover, assume that

$$q := \left\|(\lambda I - T)^{-1}(A - T)A\right\| < |\lambda| \tag{3.25}$$

holds. Then $\lambda \notin \sigma(A)$ and the following estimates hold:

$$\left\|(\lambda I - A)^{-1}\right\| \leq \frac{1 + \|(\lambda I - T)^{-1}\| \, \|A\|}{|\lambda| - q}\,, \tag{3.26}$$

$$\begin{aligned} &\left\|(\lambda I - A)^{-1}x - (\lambda I - T)^{-1}x\right\| \\ &\qquad \leq \frac{\|(\lambda I - T)^{-1}\| \cdot \|Ax - Tx\| + q\,\|(\lambda I - T)^{-1}x\|}{|\lambda| - q}\,. \end{aligned} \tag{3.27}$$

**Proof:** Let $B := (\lambda I - T)^{-1}(A - T)A$. Since, according to (3.25), $q = \|B\| < |\lambda|$, $(\lambda I - B)$ is continuously invertible. Noting that

$$\begin{aligned} (\lambda I - B) &= \lambda I - (\lambda I - T)^{-1}(A - T)A \\ &= (\lambda I - T)^{-1}(\lambda(\lambda I - T) - (A - T)A) \\ &= (\lambda I - T)^{-1}(\lambda I - T + A)(\lambda I - A)\,, \end{aligned}$$

we obtain that $\mathcal{N}(\lambda I - A) \subseteq \mathcal{N}(\lambda I - B) = \{0\}$. Since $A$ is compact, Theorem 2.19 implies that $\lambda \notin \sigma(A)$.

Multiplying the equation above with $(\lambda I - B)^{-1}$ from the left and with $(\lambda I - A)^{-1}$ from the right, yields $(\lambda I - A)^{-1} = (\lambda I - B)^{-1}(\lambda I - T)^{-1}((\lambda I - T) + A)$, i.e.,

$$(\lambda I - A)^{-1} = (\lambda I - B)^{-1}(I + (\lambda I - T)^{-1}A)\,. \tag{3.28}$$

Since $\|B\| = q < |\lambda|$, an expansion into a Neumann series implies that

$$\left\|(\lambda I - B)^{-1}\right\| \leq \frac{1}{|\lambda| - q}\,, \tag{3.29}$$

which together with (3.28) implies (3.26).

(3.28) also implies that

$$\begin{aligned} &(\lambda I - A)^{-1} - (\lambda I - T)^{-1} \\ &\quad = (\lambda I - B)^{-1}(I + (\lambda I - T)^{-1}A - (\lambda I - B)(\lambda I - T)^{-1}) \\ &\quad = (\lambda I - B)^{-1}(I + (\lambda I - T)^{-1}A - \lambda(\lambda I - T)^{-1} + B(\lambda I - T)^{-1}) \\ &\quad = (\lambda I - B)^{-1}((\lambda I - T)^{-1}(\lambda I - T + A - \lambda I) + B(\lambda I - T)^{-1}) \\ &\quad = (\lambda I - B)^{-1}((\lambda I - T)^{-1}(A - T) + B(\lambda I - T)^{-1}) \end{aligned}$$

so that

$$\begin{aligned} &\left\|(\lambda I - A)^{-1}x - (\lambda I - T)^{-1}x\right\| \\ &\qquad \leq \left\|(\lambda I - B)^{-1}\right\| \left(\left\|(\lambda I - T)^{-1}\right\| \|Ax - Tx\| + \|B\| \left\|(\lambda I - T)^{-1}x\right\|\right) \end{aligned}$$

for all $x \in X$. Together with (3.25) and (3.29) this implies estimate (3.27). ∎

This lemma may be used to obatin a convergence result for the Nyström method:

**Theorem 3.18.** Let $X$ be a Banach space, let $K : X \to X$ be compact, $\lambda \neq 0$, $\lambda \notin \sigma(K)$, and $f \in X$. Moreover, let $(K_n)$ be a sequence of linear operators with $K_n \xrightarrow{\text{cc}} K$ and define

$$q_n := \left\| (\lambda I - K)^{-1}(K - K_n)K_n \right\| , \qquad n \in \mathbb{N} . \tag{3.30}$$

Then it holds that

$$\lim_{n\to\infty} q_n = 0 . \tag{3.31}$$

If $n \in \mathbb{N}$ is such that $q_n < |\lambda|$, then $\lambda \notin \sigma(K_n)$ and

$$\left\| (\lambda I - K_n)^{-1} \right\| \leq \frac{1 + \|(\lambda I - K)^{-1}\| \, \|K_n\|}{|\lambda| - q_n} .$$

Let $x$ and $x_n$ denote the unique solutions of $\lambda x - Kx = f$ and $\lambda x_n - K_n x_n = f$. Then

$$\|x_n - x\| \leq \frac{\|(\lambda I - K)^{-1}\| \, \|K_n f - Kf\| + q_n \, \|x\|}{|\lambda| - q_n} , \tag{3.32}$$

especially $(x_n)$ converges to $x$.

**Proof:** Assertion (3.31) immediately follows from Lemma 3.16. The rest follows from Lemma 3.17 with $A := K_n$ and $T := K$. ∎

**Theorem 3.19.** Let $X := C[0,1]$ and let $K$ be the integral operator of (3.18). Moreover, let $Q_n$ be a quadrature rule (3.19) for all $n \in \mathbb{N}$ so that

$$\lim_{n\to\infty} Q_n x = \int_0^1 x(t)\, dt \quad \text{for all} \quad x \in C[0,1] \tag{3.33}$$

holds. Using these $Q_n$, $K_n$ is defined by (3.20). Then

$$K_n \xrightarrow{\text{cc}} K .$$

**Proof:** It follows from (3.33) and the Banach-Steinhaus theorem that

$$\sup_{n\in\mathbb{N}} \|Q_n\| < \infty .$$

Since $\|K_n x\|_\infty \leq \|Q_n\| \, \|k\|_\infty \, \|x\|_\infty$ for all $x \in C[0,1]$, this yields that for any bounded

set $B$ the set $\bigcup_{n\in\mathbb{N}} K_n(B)$ is bounded. But this set is also equicontinuous, since

$$\begin{aligned} |(K_n x)(s) - (K_n x)(\sigma)| &= \left| \sum_{j=1}^{n} \omega_j (k(s,t_j) - k(\sigma,t_j)) x(t_j) \right| \\ &\leq \|Q_n\| \sup_{t\in[0,1]} |k(s,t) - k(\sigma,t)| \, \|x\|_\infty \end{aligned}$$

for all $x \in B$, $s, \sigma \in [0,1]$, and $n \in \mathbb{N}$, and since $k$ is uniformly continuous. Therefore, due to the theorem of Arzela-Ascoli $\overline{\bigcup_{n\in\mathbb{N}} K_n(B)}$ is compact for any bounded set $B$. Thus, $\{K_n : n \in \mathbb{N}\}$ is collectively compact.

It remains to be shown that $(K_n)$ converges pointwise to $K$, i.e.,

$$\|K_n x - Kx\|_\infty \to 0 \quad \text{for all} \quad x \in C[0,1] \,.$$

Let $x \in C[0,1]$ and $V_x := \{k(s,\cdot)x : s \in [0,1]\} \subseteq C[0,1]$. Due to the Arzela-Ascoli theorem, the set $\overline{V_x}$ is compact. Now Lemma 2.9 applied to

$$Y := \mathbb{R} \,, \quad T_n := Q_n \,, \quad Tx := \int_0^1 x(t)\, dt \,, \quad \text{and} \quad A := \overline{V_x} \,,$$

together with condition (3.33) implies that

$$\|K_n x - Kx\|_\infty = \sup_{s\in[0,1]} \left| Q_n(k(s,\cdot)x) - \int_0^1 k(s,t)x(t)\, dt \right| \to 0 \,, \tag{3.34}$$

i.e., $(K_n)$ converges pointwise to $K$. ∎

Combining both theorems above we obtain a result for the Nyström method:

**Corollary 3.20.** Let the conditions of Theorem 3.19 hold and let $\lambda \neq 0$, $\lambda \notin \sigma(K)$. Then it holds: The approximation $x_n$ obtained by method 3.12 is uniquely determined for $n$ sufficiently large. Moreover, $(x_n)$ converges uniformly to the unique solution of (3.18) and satisfies the error estimate (3.32) with $q_n$ as in (3.30) and $K_n$ as in (3.20).

**Proof:** Since $x_n$ solves (3.21), Theorem 3.19 implies that Theorem 3.18 is applicable. This yields the assertion. ∎

**Remark 3.21.** This corollary says that the Nyström method converges if $\lambda \notin \sigma(K)$ and if the quadrature rule is convergent, i.e., if (3.33) holds. But this is equivalent to pointwise convergence of $Q_n$ holding for polynomials and that

$$\sup_{n\in\mathbb{N}} \sum_{j=1}^{n} |\omega_j^{(n)}| < \infty \,.$$

If the quadrature rules are based on polynomial interpolation, the second condition is automatically satisfied if all weights are non-negative. This is definitely the case for Gaussian quadrature rules.

As one can see from (3.32), the error estimate depends on error estimates for quadrature rules and, hence, on the smoothness of $x$ and $k$. These error estimates depending on $x$ and $k$ show up in the expression $\|K_n x - Kx\|$ (see (3.34)) and also in $\|(K_n - K)K_n\|$ and, hence, in $q_n$.

**Remark 3.22.** The major advantage of the Nyström method is that no integrals have to be evaluated for the calculation of the coefficients of the linear system (3.23). However, very often one needs a higher dimension to obtain the same quality of approximation compared to, e.g., the Galerkin method. To avoid too large systems, one usually uses iterative methods for the solution of (3.21) as follows: since

$$\begin{aligned} \lambda x_n - K_m x_n &= (\lambda x_n - K_n x_n) + (K_n - K_m)x_n \\ &= f + \frac{1}{\lambda}(K_n - K_m)(f + K_n x_n) \end{aligned}$$

for all $m < n$, this suggests the following iteration method, assuming an initial guess $x_n^{(0)}$:

$$(\lambda I - K_m)x_n^{(k+1)} = f + \frac{1}{\lambda}(K_n - K_m)(f + K_n x_n^{(k)})$$

For $m \ll n$ this procedure has the advantage that a much smaller system has to be solved. Under certain conditions one can prove that $x_n^{(k)} \to x_n$ as $k \to \infty$. For details see [2].

# 4. Volterra equations

We consider linear Volterra equations of the second kind, i.e.,

$$\lambda x(s) - \int_0^s k(s,t)x(t)\,dt = f(s)\,, \qquad s \in [0, s_0]\,. \tag{4.1}$$

This corresponds to an equation (2.51), i.e., $\lambda x - Kx = f$, where $K$ is a Volterra integral operator, defined by

$$(Kx)(s) := \int_0^s k(s,t)x(t)\,dt \tag{4.2}$$

with given $L^2$-kernel $k$ on $L^2([0,s_0]^2)$. A kernel $k$ with the property

$$k(s,t) = 0 \qquad \text{for} \qquad t > s$$

is called **Volterra kernel**.

## 4.1. Solvability

Obviously (4.1) is a special case of a Fredholm equation of the second kind so that all results from Chapter 2 are applicable. It turns out, however, that a Volterra integral operator has no eigenvalues different from 0 so that the Neumann series (2.70) converges for all $\lambda \neq 0$. The same is true for the series (2.72) of the resolvent kernel.

**Theorem 4.1.** Let $k \in L^2([0,s_0]^2)$ be a Volterra kernel, $K$ the corresponding Volterra integral operator on $L^2[0,s_0]$, defined as in (4.2). Then $\sigma(K) = \{0\}$.

**Proof:** Due to Theorem 2.33, $0 \in \sigma(K)$. For all $s,t \in [0,s_0]$ we define

$$A(s) := \left(\int_0^s |k(s,t)|^2\,dt\right)^{\frac{1}{2}}, \quad B(t) := \left(\int_t^{s_0} |k(s,t)|^2\,ds\right)^{\frac{1}{2}}, \quad D(s) := \int_0^s A(u)^2\,du\,.$$

Since $k \in L^2([0,s_0]^2)$, $A, B \in L^2[0,s_0]$. Therefore, a $C > 0$ exists with

$$D(s_0) = \int_0^{s_0} A(s)^2\,ds \le C\,, \qquad \int_0^{s_0} B(t)^2\,dt \le C\,. \tag{4.3}$$

Let $k_n$, $n \in \mathbb{N}$, be the kernel of the integral operator $K^n$. Then, due to (2.71), the following recursion formula holds: $k_1 = k$ and

$$k_i(s,t) = \begin{cases} \displaystyle\int_t^s k(s,\tau)k_{i-1}(\tau,t)\,d\tau\,, & t \le s\,, \\ 0\,, & t > s\,, \end{cases} \qquad i \ge 2\,,\ s \in [0,s_0]\,. \tag{4.4}$$

We show by induction that

$$|k_i(s,t)|^2 \le A(s)^2 B(t)^2 \frac{(D(s) - D(t))^{i-2}}{(i-2)!}\,, \qquad t \le s\,,\ i \ge 2\,. \tag{4.5}$$

If $i = 2$ and $t \le s$, then

$$\begin{aligned} |k_2(s,t)|^2 &\le \left( \int_t^s |k(s,\tau)||k(\tau,t)|\, d\tau \right)^2 \\ &\le \int_0^s |k(s,\tau)|^2\, d\tau \int_t^{s_0} |k(\tau,t)|^2\, d\tau \;=\; A(s)^2 B(t)^2\,, \end{aligned}$$

i.e., (4.5) holds for $i = 2$. Let now (4.5) hold for some $i \ge 2$. Then (4.4) implies that

$$\begin{aligned} |k_{i+1}(s,t)|^2 &\le \int_0^s |k(s,\tau)|^2\, d\tau \int_t^s |k_i(\tau,t)|^2\, d\tau \\ &\le A(s)^2 \int_t^s A(\tau)^2 B(t)^2 \frac{(D(\tau)-D(t))^{i-2}}{(i-2)!}\, d\tau\,. \end{aligned} \tag{4.6}$$

for $t \le s$. Since $A^2 \in L^1[0,s_0]$, $D$ is absolutely continuous and $D'(s) = A(s)^2$ a.e. Thus, one can use the substitution rule with the variable $u := D(\tau) - D(t)$ to obtain

$$\int_t^s A(\tau)^2 \frac{(D(\tau)-D(t))^{i-2}}{(i-2)!}\, d\tau = \frac{u^{i-1}}{(i-1)!}\Big|_0^{D(s)-D(t)} = \frac{(D(s)-D(t))^{i-1}}{(i-1)!}\,.$$

This together with (4.6) shows that (4.5) holds for $i+1$.

Since $D$ is monotonically increasing (4.5) and (4.3) imply that

$$\begin{aligned} \left\|K^i x\right\|_2^2 &= \int_0^{s_0} \left( \int_0^s k_i(s,t)x(t)\, dt \right)^2 ds \;\le\; \|x\|_2^2 \int_0^{s_0} \int_0^s |k_i(s,t)|^2\, dtds \\ &\le \|x\|_2^2 \int_0^{s_0} \int_0^s A(s)^2 B(t)^2 \frac{(D(s)-D(t))^{i-2}}{(i-2)!}\, dtds \\ &\le \frac{D(s_0)^{i-2}}{(i-2)!} \|x\|_2^2 \int_0^{s_0} A(s)^2\, ds \int_0^{s_0} B(t)^2\, dt \;\le\; \frac{C^i}{(i-2)!} \|x\|_2^2\,. \end{aligned}$$

for all $x \in L^2[0,s_0]$ and $i \ge 2$ and some $C > 0$. Thus,

$$0 \le \lim_{i\to\infty} \sqrt[i]{\|K^i\|} \le \lim_{i\to\infty} \frac{\sqrt{C}}{\sqrt[2i]{(i-2)!}} = 0\,. \tag{4.7}$$

Assume that $\lambda \in \sigma(K)$ exists with $\lambda \ne 0$. Then, due to Theorem 2.33 (b), $\lambda$ is an eigenvalue, i.e., $x \ne 0$ exists with $\lambda x = Kx$ and, hence, $\lambda^i x = K^i x$ for all $i \in \mathbb{N}$. This implies that

$$|\lambda^i|\, \|x\| = \left\|K^i x\right\| \le \left\|K^i\right\| \|x\| \quad \Longrightarrow \quad |\lambda| \le \sqrt[i]{\|K^i\|}\,.$$

Together with (4.7) this yields that $\lambda = 0$, which is a contradiction to the assumption. Thus, $\sigma(K) = \{0\}$. ∎

Theorem 4.1 together with Theorem 2.19 implies that the Volterra equation (4.1) with $L^2$-kernel $k$ has a unique solution in $L^2[0,s_0]$ for all $\lambda \ne 0$ and for all $f \in L^2[0,s_0]$. The same is true for the adjoint equation with respect to $L^2[0,s_0]$ given by (see Theorem 2.25 (a))

$$\lambda x(s) - \int_s^{s_0} k(t,s)x(t)\, dt = f(s)\,, \qquad s \in [0,s_0]\,.$$

If one allows solutions of (4.1) that are not in $L^2[0,s_0]$, then (4.1) is not necessarily uniquely solvable anymore. Thus, the assertion of Theorem 4.1 clearly depends on the spaces.

**Example 4.2.** Let

$$k(s,t) := \begin{cases} t^{s-t}\,, & 0 < t \le s \le 1\,, \\ 0\,, & \text{else}\,. \end{cases} \tag{4.8}$$

$k$ is bounded and, hence, a Volterra kernel in $L^2([0,1]^2)$. Let

$$x_0(s) := \begin{cases} s^{s-1}\,, & s \in (0,1]\,, \\ 0\,, & s = 0\,. \end{cases}$$

Then

$$\int_0^s k(s,t)x_0(t)\,dt = \int_0^s t^{s-t}t^{t-1}\,dt = \int_0^s t^{s-1}\,dt = s^{s-1} = x_0(s)\,.$$

Thus, equation

$$x(s) - \int_0^s k(s,t)x(t)\,dt = 0\,, \qquad s \in [0,1]\,,$$

has the non-trivial solution $x_0$. Therefore, 1 is an eigenvalue of this Volterra integral operator if it is considered in a space containg $x_0$. Due to Theorem 4.1, $x_0$ cannot be an element of $L^2[0,1]$, which can be easily shown directly. Equation (4.1) with $\lambda = 1$ and $k$ as in (4.8) has a unique solution $x_1 \in L^2[0,1]$, but infinitly many solutions $x_1 + \alpha x_0$, $\alpha \in \mathbb{R}$, that are not in $L^2[0,1]$.

Under a simple condition, Volterra equations of the first kind may be reduced to equations of the second kind:

**Theorem 4.3.** Let $k \in C([0,s_0]^2)$ and assume that $k$ is continuously differentiable with respect to the first variable. Moreover, let $f \in C^1[0,s_0]$ with $f(0) = 0$ and assume that

$$k(s,s) \neq 0 \qquad \text{for all} \qquad s \in [0,s_0]\,.$$

Then $x \in C[0,s_0]$ solves

$$\int_0^s k(s,t)x(t)\,dt = f(s)\,, \qquad s \in [0,s_0]\,, \tag{4.9}$$

if and only if $x$ solves

$$x(s) + \frac{1}{k(s,s)}\int_0^s \frac{\partial k}{\partial s}(s,t)x(t)\,dt = \frac{f'(s)}{k(s,s)}\,, \qquad s \in [0,s_0]\,. \tag{4.10}$$

**Proof:** By differentiation it follows from (4.9) that

$$k(s,s)x(s) + \int_0^s \frac{\partial k}{\partial s}(s,t)x(t)\,dt = f'(s)\,,$$

i.e., (4.9) implies (4.10). Conversely, since

$$k(s,s)x(s) + \int_0^s \frac{\partial k}{\partial s}(s,t)x(t)\,dt = \frac{d}{ds}\left(\int_0^s k(s,t)x(t)\,dt\right),$$

(4.10) and integration yields

$$\int_0^s k(s,t)x(t)\,dt = \int_0^s f'(t)\,dt = f(s) - f(0) = f(s)\,,$$

i.e., (4.10) implies (4.9). ∎

**Remark 4.4.** This means that under the conditions of Theorem 4.3 one can conclude that an equation of the first kind (4.9) is uniquely solvable for every $f \in C^1[0,s_0]$ with $f(0) = 0$, since (4.10) as a Volterra equation of the second kind has a unique solution $x \in L^2[0,s_0]$. Since

$$x(s) = \frac{f'(s)}{k(s,s)} - \frac{1}{k(s,s)}\int_0^s \frac{\partial k}{\partial s}(s,t)x(t)\,dt$$

and since the right-hand side is continuous, which can be seen analogously to the proof of Theorem 2.1, $x$ is even continuous.

The range of the integral operator in (4.9) is infinite-dimensional, the nullspace in $C[0,s_0]$ equals $\{0\}$. Therefore, Theorem 2.17 implies that the range of the operator in (4.9) is not closed in $C[0,s_0]$. Thus, continuous functions $f$ must exist, where (4.9) is not solvable in $C[0,s_0]$. Analogously one can see that $f \in L^2[0,s_0]$ must exist so that (4.9) has no solution in $L^2[0,s_0]$.

In case of unique solvability, solutions of equations of the second kind depend continuously on the right-hand side. This does not hold for solutions of equations of the first kind, i.e., they are **ill-posed**. Note that the reduction of an equation (4.9) to an equation (4.10) is no remedy for this fact, since the transition from $f$ to $f'$ on the right-hand side is discontinuous in $C[0,s_0]$. However, it means that, under the conditions of Theorem 4.3, problem (4.9) is as ill-posed as differentiation.

Theorem 4.3 is not applicable if $k(s,s) = 0$ for some $s$ as, e.g., in the Volterra equation (1.5) in Example 1.1. If, however, $k(s,s) = 0$ for all $s \in [0,s_0]$, then one can show by an additional differentiation, assuming that $f'(0) = 0$, that (4.9) is equivalent to the equation of the second kind,

$$x(s) + \frac{1}{\frac{\partial k}{\partial s}(s,s)}\int_0^s \frac{\partial^2 k}{\partial s^2}(s,t)x(t)\,dt = \frac{1}{\frac{\partial k}{\partial s}(s,s)}f''(s)\,, \qquad s \in [0,s_0]\,.$$

Obviously, $\frac{\partial^2 k}{\partial s^2}$ has to be continuous and $\frac{\partial k}{\partial s}(s,s) \neq 0$ for all $s \in [0,s_0]$. The discontinuous dependence of the solution of (4.9) on $f$ then consists in the transition of $f$ to $f''$, which is stronger than for the transition to (4.10). One can see that a degree of ill-posedness of Volterra equations of the first kind in some special cases is possible by the number of how often one has to differentiate the right-hand side to end up in an equation of the second kind.

If $k(0,0) = 0$, but $k(s,s) \neq 0$ for $s > 0$, then a transition of (4.9) to an equation of the second kind is never possible in the interval $[0,s_0]$ with $s_0 > 0$.

A special (weakly singular) Volterra integral equation of the first kind is Abel's integral equation (1.12) or its generalization

$$\int_0^s \frac{g(s,t)}{(s-t)^\alpha} x(t)\,dt = f(s)\,, \qquad s \in [0, s_0]\,, \tag{4.11}$$

with continuous $g$, $g(s,s) \neq 0$, and $\alpha \in (0,1)$. An explicit solution formula can be given for this equation that is based on the following transition of equation (4.11) to an equivalent equation with continuous kernel.

**Theorem 4.5.** Let $g \in C([0,s_0]^2)$, $\alpha \in (0,1)$, $f, x \in C[0,s_0]$. Then $x$ is a solution of (4.11) if and only if

$$\int_0^\tau k(\tau,t)x(t)\,dt = \tilde{f}(\tau)\,, \qquad \tau \in [0,s_0]\,, \tag{4.12}$$

with

$$k(\tau,t) := \int_0^1 \frac{g(t+r(\tau-t),t)}{(1-r)^{1-\alpha}r^\alpha}\,dr\,, \qquad \tau,t \in [0,s_0]\,, \tag{4.13}$$

and

$$\tilde{f}(\tau) := \int_0^\tau \frac{f(t)}{(\tau-t)^{1-\alpha}}\,dt \qquad \tau \in [0,s_0]\,. \tag{4.14}$$

**Proof:** Note that the improper integrals in (4.13) and (4.14) exist and are continuous functions. Let $x$ be a solution of (4.11). Then multiplication by $(\tau-s)^{\alpha-1}$ and integration with respect to $s$ from 0 to $\tau$, $\tau \in [0,s_0]$, yields:

$$\begin{aligned}\tilde{f}(\tau) &= \int_0^\tau \frac{f(s)}{(\tau-s)^{1-\alpha}}\,ds = \int_0^\tau \frac{1}{(\tau-s)^{1-\alpha}} \int_0^s \frac{g(s,t)}{(s-t)^\alpha} x(t)\,dtds \\ &= \int_0^\tau \int_t^\tau \frac{g(s,t)x(t)}{(\tau-s)^{1-\alpha}(s-t)^\alpha}\,dsdt = \int_0^\tau k(\tau,t)x(t)\,dt\,,\end{aligned}$$

i.e., (4.12) holds. The last identity follows from

$$\int_t^\tau \frac{g(s,t)}{(\tau-s)^{1-\alpha}(s-t)^\alpha} ds = \int_0^1 \frac{g(t+r(\tau-t),t)}{((1-r)(\tau-t))^{1-\alpha}(r(\tau-t))^\alpha}(\tau-t)\,dr = k(\tau,t)\,.$$

Conversely, let $x$ be a solution of (4.12), i.e.,

$$\begin{aligned}0 &= \int_0^\tau \int_0^1 \frac{g(t+r(\tau-t),t)}{(1-r)^{1-\alpha}r^\alpha}\,dr\,x(t)\,dt - \int_0^\tau \frac{f(t)}{(\tau-t)^{1-\alpha}}\,dt \\ &= \int_0^\tau \int_t^\tau \frac{g(s,t)}{(\tau-s)^{1-\alpha}(s-t)^\alpha} x(t)\,dsdt - \int_0^\tau \frac{f(s)}{(\tau-s)^{1-\alpha}}\,ds \\ &= \int_0^\tau \int_0^s \frac{g(s,t)}{(\tau-s)^{1-\alpha}(s-t)^\alpha} x(t)\,dtds - \int_0^\tau \frac{f(s)}{(\tau-s)^{1-\alpha}}\,ds \\ &= \int_0^\tau (\tau-s)^{\alpha-1}h(s)ds \qquad \text{with} \qquad h(s) := \int_0^s \frac{g(s,t)}{(s-t)^\alpha} x(t)\,dt - f(s)\end{aligned}$$

for all $\tau \in [0, s_0]$. Multiplication by $(\sigma - \tau)^{-\alpha}$ and integration with respect to $\tau$ from 0 to $\sigma$, $\sigma \in [0, s_0]$, implies that

$$\begin{aligned} 0 = \int_0^\sigma \int_0^\tau \frac{h(s)}{(\sigma-\tau)^\alpha(\tau-s)^{1-\alpha}}\, ds d\tau &= \int_0^\sigma \int_s^\sigma \frac{h(s)}{(\sigma-\tau)^\alpha(\tau-s)^{1-\alpha}}\, d\tau ds \\ &= \int_0^\sigma h(s)\, ds \int_0^1 \frac{dr}{r^\alpha(1-r)^{1-\alpha}}\,. \end{aligned}$$

Thus, $\int_0^\sigma h(s)ds = 0$ for all $\sigma \in [0, s_0]$. Therefore, $h \equiv 0$, since $h$ is continuous. From the definition of $h$ we finally conclude that $x$ is a solution of (4.11). ∎

**Remark 4.6.** After transforming (4.11) to a Volterra integral equation of the first kind with continuous kernel one can reduce it to an equivalent integral equation of the second kind if $g(s, s) \neq 0$ for all $s \in [0, s_0]$, $g$ continuously differentiable with respect to the first variable, and if $\tilde{f} \in C^1[0, s_0]$. Note that then also $k(t, t) \neq 0$ for all $t \in [0, s_0]$ and that $k$ is continuously differentiable with respect to the first variable. According to Theorem 4.3, (4.12) can be reduced to

$$x(\tau) + \frac{1}{k(\tau,\tau)} \int_0^\tau \frac{\partial k}{\partial \tau}(\tau, t)x(t)\, dt = \frac{\tilde{f}'(\tau)}{k(\tau,\tau)}\,, \qquad \tau \in [0, s_0]\,.$$

**Example 4.7.** Using Theorem 4.5, equation (4.11) can be solved explicitly if $g \equiv 1$. Since

$$\int_0^1 r^\beta (1-r)^\gamma\, dr = \frac{\Gamma(\beta+1)\Gamma(\gamma+1)}{\Gamma(\beta+\gamma+2)}\,, \qquad \beta\,,\, \gamma > -1\,,$$

and since

$$\Gamma(t)\Gamma(1-t) = \frac{\pi}{\sin(\pi t)}\,, \qquad t \in (0, 1)\,,$$

we obtain in this case that

$$k(\tau, t) = \int_0^1 r^{-\alpha}(1-r)^{\alpha-1}\, dr = \frac{\Gamma(1-\alpha)\Gamma(\alpha)}{\Gamma(1)} = \frac{\pi}{\sin(\pi\alpha)}$$

so that (4.12) has the form

$$\int_0^\tau x(t)\, dt = \frac{\sin(\pi\alpha)}{\pi} \int_0^\tau \frac{f(t)}{(\tau-t)^{1-\alpha}}\, dt\,, \qquad \tau \in [0, s_0]\,.$$

Differentiation on both sides yields the solution formula

$$x(s) = \frac{\sin(\pi\alpha)}{\pi} \frac{d}{ds} \int_0^s \frac{f(t)}{(s-t)^{1-\alpha}}\, dt\,, \qquad s \in [0, s_0]\,,$$

holding for all $f$, where the right-hand side makes sense. If $f \in C^1[0, s_0]$, a further transformation is possible:

$$\begin{aligned} \frac{d}{ds} \int_0^s \frac{f(t)}{(s-t)^{1-\alpha}}\, dt &= \frac{d}{ds} \int_0^1 s^\alpha \frac{f(rs)}{(1-r)^{1-\alpha}}\, dr \\ &= \alpha s^{\alpha-1} \int_0^1 \frac{f(rs)}{(1-r)^{1-\alpha}}\, dr + s^\alpha \int_0^1 \frac{f'(rs)r}{(1-r)^{1-\alpha}}\, dr \\ &= \alpha s^{-1} \int_0^s \frac{f(t)}{(s-t)^{1-\alpha}}\, dt + s^{-1} \int_0^s \frac{tf'(t)}{(s-t)^{1-\alpha}}\, dt\,. \end{aligned}$$

Thus,

$$x(s) = \frac{\sin(\pi\alpha)}{\pi s} \int_0^s \frac{\alpha f(t) + t f'(t)}{(s-t)^{1-\alpha}}\, dt\,, \qquad s \in (0, s_0]\,.$$

As a special case, we get as solution of Abel's integral equation (1.12) (see Example 1.5)

$$x(s) = \frac{\sqrt{2g}}{\pi s} \int_0^s \frac{\frac{f(t)}{2} + t f'(t)}{\sqrt{s-t}}\, dt\,, \qquad s \in (0, s_0]\,, \tag{4.15}$$

where we assume that $f(0) = 0$ and that $f \in C^1[0, s_0]$. The uniqueness of this solution follows from the already mentioned connection with a Volterra integral equation of the second kind.

From formula (4.15) one might guess the ill-posedness of (1.12): the right-hand side has to be differentiated. The discontinuous dependence on $f$ is then weakend by the integration, however, due to the singularity $(s-t)^{-\frac{1}{2}}$ not strong enough. One can show that solving (1.12) is half as ill-posed as differentiation.

**Remark 4.8.** An analytic method for the solution of Volterra equations with a difference kernel is based on the Laplace transform. We only discuss basics and do not deal with problems of the domain of the Laplace transform and present no proofs: we consider

$$\lambda x(s) - \int_0^s k(s-t)x(t)\, dt = f(s)\,, \qquad s \in [0, s_0]\,, \tag{4.16}$$

with continuous $k$, $f$, and $\lambda \neq 0$. The Laplace convolution of $g, h : \mathbb{R}_0^+ \to \mathbb{R}$ is defined by

$$(g * h)(s) := \int_0^s g(s-t)h(t)\, dt\,, \qquad s \in \mathbb{R}_0^+\,.$$

An important property of the Laplace transform, defined in Example 1.3, is that

$$L(g * h) = (Lg)(Lh)\,.$$

Since (4.16) is equivalent to

$$\lambda x(s) - (k * x)(s) = f(s)\,, \qquad s \in [0, s_0]\,,$$

we obtain that

$$\lambda Lx - (Lk)(Lx) = (Lf) \qquad \Longrightarrow \qquad Lx = \frac{Lf}{\lambda - Lk}\,,$$

holding for those arguments where $Lk \neq \lambda$. Therefore, $Lx$ can be explicitly calculated. It is not a good idea to numerically compute $x$ from its Laplace transform, since this is a very ill-posed problem. This method is only appropriate if $Lx$ is a function that appears in tables on Laplace transforms or if the application of the inversion formula

$$x(s) = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} e^{s(\alpha+i\beta)} (Lx)(\alpha + i\beta)\, d\beta$$

makes sense. Here, $\alpha$ is larger than the real parts of all singularities.

## 4.2. Numerical solution

Volterra integral equations contain as a special case initial value problems for ordinary differential equations (see Example 1.1). It is, therefore, obvious to try to transfer numerical methods for the solution of initial value problems to the solution of Volterra integral equations. We only treat one example of a Runge-Kutta method.

**Example 4.9.** We consider the nonlinear Volterra integral equation

$$x(s) - \int_0^s k(s,t,x(t))\,dt = f(s)\,, \qquad s \geq 0\,, \tag{4.17}$$

with a continuous kernel $k$ and continuous $f$. Obviously, (4.17) also contains the linear case.

An explicit Runge-Kutta method for the solution of the initial value problem

$$x'(s) = k(s,x(s))\,, \qquad s \geq 0\,, \qquad x(0) = a\,, \tag{4.18}$$

or equivalently

$$x(s) - \int_0^s k(t,x(t))\,dt = a\,, \qquad s \geq 0\,, \tag{4.19}$$

has the form

$$x_{j+1} := x_j + h\sum_{i=1}^{m} \gamma_i \varphi_i(s_j, x_j)\,, \qquad j \in \mathbb{N}\,,$$

where $h$ is the step size, $x_j$ is an approximation of $x(s_j)$ with $s_j := jh$, $j \in \mathbb{N}$, $x_0 := a$, $m \geq 2$, and

$$\varphi_i(s,x) := k\left(s + \alpha_i h, x + h\sum_{j=1}^{i-1} \beta_{ij}\varphi_j(s,x)\right), \quad i = 1,\ldots,m\,.$$

Here, $\alpha_1 := 0$ and all other parameters, collected in the so-called **Runge-Kutta scheme**

$$\begin{array}{c|ccccc} \alpha_2 & \beta_{21} & & & & \\ \vdots & \vdots & \ddots & & & \\ \alpha_m & \beta_{m1} & \cdots & \beta_{m,m-1} & & \\ \hline & \gamma_1 & \cdots & \gamma_{m-1} & \gamma_m \end{array}$$

have to be determined in such a way that a prescribed order of consistency $p$ is obtained, i.e.,

$$|\tau_h(s)| = O(h^p) \qquad \text{with} \qquad h \to 0$$

for all $s$, where $\tau_h$ is the local discretization error

$$\tau_h := \frac{x(s+h) - x(s)}{h} - \sum_{i=1}^{m} \gamma_i \varphi_i(s, x(s))$$

and $x$ is the exact solution of (4.18). Equations for the coefficients are obtained via a Taylor series expansion of $\tau_h$ in $h$ and a comparison of coefficients so that all terms for

a power less than $p$ vanish. Of course, it is the aim to make $p$ as large as possible for a given $m$.

For the special case $m = 4$, one can show that the best order is $p = 4$. It is obtained with the standard fourth-order scheme:

$$\begin{array}{c|cccc} \frac{1}{2} & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ 1 & 0 & 0 & 1 & \\ \hline & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \tag{4.20}$$

or with the Kuntzmann scheme:

$$\begin{array}{c|cccc} \frac{2}{5} & \frac{2}{5} & & & \\ \frac{3}{5} & -\frac{3}{20} & \frac{3}{4} & & \\ 1 & \frac{19}{44} & -\frac{15}{44} & \frac{40}{44} & \\ \hline & \frac{55}{360} & \frac{125}{360} & \frac{125}{360} & \frac{55}{360} \end{array}$$

These ideas that lead to a solution of the special Volterra equation (4.19) can be transferred to the general situation (4.17): an essential difference then is that the kernel also depends on $s$. An $m$ step Runge-Kutta like fomula looks as follows: let $h$ be the step size, $s_j = t_j := jh$, $j \in \mathbb{N}_0$, and $\tilde{x}$ denotes the approximate solution, computed as follows:

$$\tilde{x}(s_j + \alpha_n h) := \varphi_j(s_j + \alpha_n h) + h \sum_{i=0}^{n-1} \beta_{ni} k(s_j + \alpha_n h, t_j + \alpha_i h, \tilde{x}(t_j + \alpha_i h))\,, \tag{4.21}$$

where $0 = \alpha_0 < \alpha_1 < \alpha_2 < \ldots < \alpha_m = 1$, $j \in \mathbb{N}_0$, $n = 1, \ldots, m$, and

$$\varphi_j(s) := f(s) + h \sum_{l=0}^{j-1} \sum_{i=0}^{m} \gamma_i k(s, t_l + \alpha_i h, \tilde{x}(t_l + \alpha_i h))\,. \tag{4.22}$$

Thus, $\tilde{x}(s_{j+1}) = \tilde{x}(s_j + \alpha_m h)$. The following initial settings are used:

$$\tilde{x}(0) = f(0)\,, \qquad \varphi_0(s) = f(s)\,.$$

This procedure may be motivated as follows: if $s \in [s_j, s_{j+1}]$, then (4.17) implies that

$$x(s) = f(s) + \sum_{l=0}^{j-1} \int_{s_l}^{s_{l+1}} k(s, t, x(t))\, dt + \int_{s_j}^{s} k(s, t, x(t))\, dt\,.$$

The function $\varphi_j$ in (4.22) is an approximation of the first two terms, where the integrals have been replaced by quadrature rules with nodes $t_l + \alpha_i h$. The last integral corresponds to the sum in (4.21). The Runge-Kutta aspect is that the computation of $\tilde{x}(s_{j+1})$ is performed in several steps. Similar considerations as for initial value problems lead to formulas for the coefficients. For $m = 4$ one gets the following scheme (compare (4.20)):

$$\begin{array}{lll|llll|llll|llll} \alpha_0 & = & 0 & & & & & & & & & & & & \\ \alpha_1 & = & \frac{1}{2} & \beta_{10} & = & \frac{1}{2} & & & & & & & & & \\ \alpha_2 & = & \frac{1}{2} & \beta_{20} & = & 0 & \beta_{21} & = & \frac{1}{2} & & & & & & \\ \alpha_3 & = & 1 & \beta_{30} & = & 0 & \beta_{31} & = & 0 & \beta_{32} & = & 1 & & & \\ \hline \alpha_4 & = & 1 & \beta_{40} & = & & \beta_{41} & = & & \beta_{42} & = & & \beta_{43} & = & \\ & & & \gamma_0 & = & \frac{1}{6} & \gamma_1 & = & \frac{1}{3} & \gamma_2 & = & \frac{1}{3} & \gamma_1 & = & \frac{1}{3} \end{array}$$

A large number of numerical methods for the solution of Volterra equations can be found in [3]. A method that is used quite often is the so-called **product integration**:

**Example 4.10.** The principle of product integration is as follows: find functions $v_0, \dots, v_n$ such that for given nodes $t_0, \dots, t_n \in [0, s_0]$

$$\int_0^s k(s,t)x(t)\,dt = \sum_{i=0}^{n} v_i(s)x(t_i)\,, \qquad s \in [0, s_0]\,, \tag{4.23}$$

for all piecewise polynomial functions $x$ of a certain order. The construction of such formulas is similar to the one of quadrature rules. We treat the simple case of equidistant nodes $t_i := ih$, $i \in \{0, \dots, n\}$, with $h := \frac{s_0}{n}$, where (4.23) should hold for all continuous piecewise linear functions $x$.

Let $x$ be a piecewise linear function such that $x(t_k) = \delta_{ik}$. Then we obtain that

$$\begin{aligned}\sum_{k=0}^{n} v_k(s)x(t_k) \;=\; v_i(s) \;\;=\;\; & \frac{1}{h}\int_{\min\{\max\{(i-1)h,0\},s\}}^{\min\{ih,s\}} k(s,t)(t+(1-i)h)\,dt \\ & + \frac{1}{h}\int_{\min\{ih,s\}}^{\min\{(i+1)h,s\}} k(s,t)((1+i)h-t)\,dt\,.\end{aligned}$$

This implies that

$$v_0(s) = \begin{cases} \dfrac{1}{h}\displaystyle\int_0^h k(s,t)(h-t)\,dt\,, & s \le h\,, \\ \dfrac{1}{h}\displaystyle\int_0^s k(s,t)(h-t)\,dt\,, & h < s\,,\ n > 1\,, \end{cases}$$

$$v_i(s) = \begin{cases} 0\,, & 0 \le s \le (i-1)h\,, \\ \dfrac{1}{h}\displaystyle\int_{(i-1)h}^{s} k(s,t)(t+(1-i)h)\,dt\,, & (i-1)h < s \le ih\,, \\ \dfrac{1}{h}\displaystyle\int_{(i-1)h}^{ih} k(s,t)(t+(1-i)h)\,dt & \\ \quad + \dfrac{1}{h}\displaystyle\int_{ih}^{s} k(s,t)((1+i)h-t)\,dt\,, & ih < s \le (i+1)h\,, \\ \dfrac{1}{h}\displaystyle\int_{(i-1)h}^{ih} k(s,t)(t+(1-i)h)\,dt & \\ \quad + \dfrac{1}{h}\displaystyle\int_{ih}^{(i+1)h} k(s,t)((1+i)h-t)\,dt\,, & (i+1)h < s \le s_0\,, \end{cases} \tag{4.24}$$

with $i \in \{1, \dots, n\}$. Note that the last case can only occur for $i < n-1$ and the one before last only for $i < n$.

When $v_i$ is as above, then (4.23) corresponds to the trapezoidal rule. Therefore, one can show that it has convergence order $O(h^2)$ for sufficiently smooth functions $x$. If (4.23) should be exact for piecewise quadratic functions, then one obtains a generalization of Simpson's rule with convergence order $O(h^4)$. Computer algebra tools are helpful for the computation of the functions $v_i$.

After approximating the integral operator via (4.23) one can combine that with collocation at the nodes $t_0, \dots, t_n$ to approximate the solution $x$ of equation (4.17) ($s \in [0, s_0]$)

at the nodes by the solution of the system

$$x(t_j) - \sum_{i=0}^{n} v_i(t_j)x(t_i) = f(t_j)\,, \quad j \in \{0,\dots,n\}\,. \tag{4.25}$$

Note that, if $v_i$ is given as in (4.24), then $v_i(t_j) = 0$ for $j < i$. Thus, the system matrix in (4.25) is then a triangular matrix which can be simply solved.

Product integration is quite popular for weakly singular kernels. The singularity only shows up in the computation of the functions $v_i$.

This method can also be applied to so-called Hammerstein equations, i.e.,

$$x(s) - \int_0^s k(s,t)\phi(t,x(t))\,dt = f(s)\,, \qquad s \in [0,s_0]\,,$$

where $\phi$ is a a non-linear function. Instead of (4.25) one then obtains the nonlinear system

$$x(t_j) - \sum_{i=0}^{n} v_i(t_j)\phi(t_i,x(t_i)) = f(t_j)\,, \qquad j \in \{0,\dots,n\}\,,$$

If one uses $v_i$ as in (4.24), one can solve it step by step, where in each step only one nonlinear equation has to be solved for $x(t_j)$.

# 5. Sturm-Liouville theory

## 5.1. Initial and boundary value problems

In this section we will transform initial and boudary value problems for second order ordinary differential equations with variable coefficients into integral equations so that we can use the results from Chapters 2 and 4.

We consider equation

$$\bar{p}(s)x''(s) + \bar{r}(s)x'(s) + \bar{q}(s)x(s) = \bar{f}(s) \tag{5.1}$$

on the intervall $[a,b]$, where $\bar{p}$, $\bar{r}$, $\bar{q}$, and $\bar{f}$ are real continuous functions and $\bar{p}$ has no zeros. Let

$$p(s) := \exp\left(\int_a^s \frac{\bar{r}(t)}{\bar{p}(t)}\,dt\right), \qquad q(s) := \frac{p(s)}{\bar{p}(s)}\bar{q}(s)\,, \qquad f(s) := \frac{p(s)}{\bar{p}(s)}\bar{f}(s)$$

for all $s \in [a,b]$. Since $p'(s) = p(s)\frac{\bar{r}(s)}{\bar{p}(s)}$, multiplying (5.1) by $\frac{p(s)}{\bar{p}(s)}$ yields

$$f(s) = p(s)x''(s) + p'(s)x'(s) + q(s)x(s) = (px')'(s) + q(s)x(s)\,.$$

Thus, (5.1) can be transformed into the problem

$$(px')'(s) + q(s)x(s) = f(s)\,, \qquad s \in [a,b]\,, \tag{5.2}$$

with $p \in C^1[a,b]$ and $q, f \in C[a,b]$, where $p$ does not vanish anywhere and $p, q$ are real. In the following we use the abbreviation

$$(Lx)(s) := (px')'(s) + q(s)x(s)\,, \qquad s \in [a,b]\,, \tag{5.3}$$

and always assume that $p$, $q$, and $f$ have the above properties.

**Theorem 5.1.** Let $\alpha, \beta \in \mathbb{R}$. Then there exists a unique solution $x \in C^2[a,b]$ of the initial value problem

$$\begin{aligned} (Lx)(s) &= f(s)\,, \qquad s \in [a,b]\,, \\ x(a) &= \alpha\,, \quad x'(a) = \beta\,. \end{aligned} \tag{5.4}$$

The solution depends continuously on $(\alpha, \beta, f)$. $x$ satisfies the Volterra integral equation of the second kind

$$x(s) + \int_a^s k(s,t)x(t)\,dt = g(s)\,, \qquad s \in [a,b]\,, \tag{5.5}$$

with

$$k(s,t) := q(t)\int_t^s \frac{d\tau}{p(\tau)}\,, \qquad s,t \in [a,b]\,,\ t \le s\,,$$

and

$$g(s) := \alpha + p(a)\beta\int_a^s \frac{d\tau}{p(\tau)} + \int_a^s f(t)\int_t^s \frac{d\tau}{p(\tau)}\,dt\,, \qquad s \in [a,b]\,. \tag{5.6}$$

(5.5) has a unique continuous solution.

**Proof:** By integration (5.4) immediately yields

$$x(a) = \alpha\,, \qquad p(s)x'(s) - p(a)\beta + \int_a^s q(t)x(t)\,dt = \int_a^s f(t)\,dt\,, \quad s \in [a,b]\,. \tag{5.7}$$

Note that, if $x \in C^1[a,b]$ solves (5.7), then $x \in C^2[a,b]$, since

$$x'(s) = \frac{1}{p(s)}\left(p(a)\beta + \int_a^s f(t)\,dt - \int_a^s q(t)x(t)\,dt\right) \tag{5.8}$$

and since the right-hand side above is continuously differentiable. On the other hand, differentiation of (5.7) immediately yields (5.4) so that (5.7) and (5.4) are equivalent.

Due to Fubini's theorem, integration of (5.8) leads to

$$\begin{aligned} x(s) - \alpha &= p(a)\beta \int_a^s \frac{d\tau}{p(\tau)} + \int_a^s \int_a^\tau \frac{f(t)}{p(\tau)}\,dt d\tau - \int_a^s \int_a^\tau \frac{q(t)x(t)}{p(\tau)}\,dt d\tau \\ &= p(a)\beta \int_a^s \frac{d\tau}{p(\tau)} + \int_a^s \int_t^s \frac{f(t)}{p(\tau)}\,d\tau dt - \int_a^s \int_t^s \frac{q(t)x(t)}{p(\tau)}\,d\tau dt\,, \end{aligned}$$

which implies (5.5). If $x$ is a continuous solution of (5.5), then, due to the continuous differentiability of $s \mapsto g(s) - \int_a^s k(s,t)x(t)\,dt$, we automatically get that $x \in C^1[a,b]$. Furthermore, differentiation yields (5.8). Altogether we have shown that the following two problems are equivalent:

(A) Find $x \in C^2[a,b]$ solving (5.4).

(B) Find $x \in C[a,b]$ solving (5.5).

Let $K$ be the integral operator with kernel $k$ in $C[a,b]$. Then Theorem 4.1 implies that only 0 can be an element of $\sigma(K)$. Since if $\lambda \in \sigma(K)\backslash\{0\}$, then, due to Theorem 2.33 (b), $\lambda$ would be an eigenvalue of $K$ and, hence, also an eigenvalue of the integral operator induced by $k$ in $L^2[a,b]$. But this would be a contradiction to Theorem 4.1. Thus, $(I+K)$ is continuously invertible. Therefore, (5.5) has exactly one solution depending continuoulsy on $g$. Since, due to (5.6), $g$ depends continuously on $(\alpha, \beta, f)$, the assertion follows from the equivalence of the problems (A) and (B). ∎

Now we turn to special boundary value problems for equation (5.2). Thereto, let $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$ with

$$\alpha_1^2 + \beta_1^2 \neq 0 \quad \text{and} \quad \alpha_2^2 + \beta_2^2 \neq 0\,. \tag{5.9}$$

We use the notations

$$B_1(x) := \alpha_1 x(a) + \beta_1 x'(a)\,, \quad B_2(x) := \alpha_2 x(b) + \beta_2 x'(b)\,, \qquad x \in C^1[a,b]\,, \tag{5.10}$$

and consider the boundary value problem

$$\begin{aligned} &(Lx)(s) = f(s)\,, \qquad s \in [a,b]\,, \\ &B_1(x) = B_2(x) = 0\,, \end{aligned} \tag{5.11}$$

with $L$ as in (5.3), $B_1, B_2$ as in (5.10), and $p, q, f, \alpha_1, \alpha_2, \beta_1, \beta_2$ as above. Note that the boundary conditions above are homogeneous.

We will see in the next section that for the solution of (5.11) it will be advantageous to study the following eigenvalue problem,

$$\begin{aligned} &(Lx)(s) = \lambda x(s)\,, \qquad s \in [a,b]\,, \\ &B_1(x) = B_2(x) = 0 \end{aligned} \tag{5.12}$$

called Sturm-Liouville problem. The solution of (5.11) will be expressed as series with respect to the eigenfunctions of (5.12).

The existence of nontrivial solutions of (5.12) does not only depend on $L$, but also on the boundary conditions that somehow define the domain of $L$: let

$$D_B := \{x \in C^2[a,b] : B_1(x) = B_2(x) = 0\}\,.$$

$D_B$ is a normed space with the usual norm in $C[a,b]$. $L_B$ denotes the restriction of $L$ onto $D_B$, i.e.,

$$L_B := L\Big|_{D_B} : D_B \to C[a,b]\,. \tag{5.13}$$

**Theorem 5.2.** Let $L_B$ be defined by (5.13) and let $\langle\,\cdot,\cdot\,\rangle$ be the usual inner product in $L^2[a,b]$. Then $L_B$ is symmetric, i.e.,

$$\langle\, L_B x, y\,\rangle = \langle\, x, L_B y\,\rangle\,, \qquad x, y \in D_B\,, \tag{5.14}$$

**Proof:** Since $B_1(x) = B_2(x) = 0 = B_1(y) = B_2(y)$, we obtain that

$$x(a)\overline{y'(a)} = x'(a)\overline{y(a)} \qquad \text{and} \qquad x(b)\overline{y'(b)} = x'(b)\overline{y(b)}\,.$$

Therefore,

$$\begin{aligned} \langle\, L_B x, y\,\rangle &= \int_a^b ((px')'(s) + q(s)x(s))\,\overline{y(s)}\,ds \\ &= p(b)x'(b)\overline{y(b)} - p(a)x'(a)\overline{y(a)} - \int_a^b (px')(s)\overline{y'(s)}\,ds + \int_a^b q(s)x(s)\overline{y(s)}\,ds \\ &= p(b)(x'(b)\overline{y(b)} - x(b)\overline{y'(b)}) - p(a)((x'(a)\overline{y(a)} - x(a)\overline{y'(a)}) \\ &\quad + \int_a^b x(s)(p(s)\overline{y'(s)})'\,ds + \int_a^b x(s)q(s)\overline{y(s)}\,ds \\ &= \int_a^b x(s)\overline{((py')'(s) + q(s)y(s))}x(s)\,ds \;=\; \langle\, x, L_B y\,\rangle \end{aligned}$$

for all $x, y \in D_B$, since $p$ and $q$ are real. ∎

In the following we assume that

$$p(s) < 0 \qquad \text{for all} \qquad s \in [a,b]\,. \tag{5.15}$$

Note that this is no restriction of the generality, since we assumed that $p \in C^1[a,b]$ does not vanish anywhere.

Defining below a Green's function for $L_B$ we will be able to express the solution of (5.11) as an integral operator applied to $f$.

**Definition 5.3.** Let $L_B$ be defined as in (5.13). A continuous function $g : [a,b]^2 \to \mathbb{R}$ is called **Green's function for** $L_B$ if $g$ is twice continuously differentiable in

$$\triangle_l := \{(s,t) \in [a,b]^2 : s \le t\} \quad \text{and} \quad \triangle_u := \{(s,t) \in [a,b]^2 : s \ge t\}$$

and if the following three conditions hold:

$$B_1(g(\cdot,t)) = B_2(g(\cdot,t)) = 0 \qquad \text{for all} \qquad t \in [a,b]\,, \tag{5.16}$$

$$\frac{\partial}{\partial s}\left(p(s)\frac{\partial g}{\partial s}(s,t)\right) + q(s)g(s,t) = 0 \quad \text{for all} \quad t \in [a,b] \quad \text{and} \quad s \in [a,b] \setminus \{t\}\,, \tag{5.17}$$

$$\lim_{h\to 0^+} \frac{\partial g}{\partial s}(t+h,t) - \lim_{h\to 0^+} \frac{\partial g}{\partial s}(t-h,t) = \frac{1}{p(t)} \quad \text{for all} \quad t \in (a,b)\,. \tag{5.18}$$

Note that for every fixed $t$ a Green's function solves (5.11) (with $f = 0$) as a function of $s$, except for $s = t$. There, $g$ is not differentiable with respect to $s$, but satisfies the jump condition (5.18). For the equation in (5.17) we shortly write

$$(L_s g)(s,t) = 0\,, \qquad s \ne t\,, \tag{5.19}$$

where $L_s$ denotes the differential operator $L$ applied to the variable $s$.

The next theorem says that such a Green's function exists under a certain condition:

**Theorem 5.4.** Let $L_B$ be as in (5.13). Then there exists a Green's function if the homogeneous problem

$$Lx = 0\,, \qquad B_1(x) = B_2(x) = 0\,, \tag{5.20}$$

only has the trivial solution $x = 0$.

**Proof:** Let $u, v \in C^2[a,b]$, $u \ne 0$, $v \ne 0$ with

$$(Lu)(s) = 0\,, \qquad s \in [a,b]\,, \qquad B_1(u) = 0\,, \tag{5.21}$$

$$(Lv)(s) = 0\,, \qquad s \in [a,b]\,, \qquad B_2(v) = 0\,. \tag{5.22}$$

Such functions exist: according to Theorem 5.1, $u$ could be chosen as solution of the initial value problem $Lu = 0$, $u(a) = -\beta_1$, $u'(a) = \alpha_1$. Due to (5.9), $u \ne 0$. Analogously one can see that $v$ exists. Let now

$$g(s,t) := \begin{cases} Cu(s)v(t)\,, & s \le t\,, \\ Cv(s)u(t)\,, & s \ge t\,, \end{cases}$$

where $C$ has to be chosen appropriately (see below). Due to (5.21) and (5.22), conditions (5.16) and (5.17) hold independently from the choice of $C$. Moreover, $g$ is twice continuously differentiable in $\triangle_l$ and in $\triangle_u$ and continuous in $[a,b]^2$. We show that

$$p(v'u - u'v) \quad \text{is constant and different from 0 in} \quad [a,b]\,. \tag{5.23}$$

For all $s \in [a,b]$ it holds that

$$\begin{aligned}\frac{d}{ds}\left(p(v'u-u'v)\right)(s) &= \left(p'(v'u-u'v)+p(v''u+v'u'-u''v-u'v')\right)(s)\\ &= \left(u((pv')'+qv)-v((pu')'+qu)\right)(s)\\ &= u(s)(Lv)(s)-v(s)(Lu)(s) \;=\; 0\,.\end{aligned}$$

Thus, $p(v'u-u'v)$ is constant in $[a,b]$. Assume that this constant equals 0. Then, since $p$ does not vanish anywhere, $v'u-u'v \equiv 0$. Especially, $v'(a)u(a) = u'(a)v(a)$ and $v'(b)u(b) = u'(b)v(a)$. But then (5.21) and (5.22) imply that $B_1(v) = B_2(u) = 0$, i.e., $u, v$ are solutions of (5.20) and, hence, due to the assumption of the theorem, identically 0. This is a contradiction to the construction of $u$ and $v$. Therefore, (5.23) holds and we may set

$$C := \frac{1}{p(v'u-u'v)}\,.$$

With this choice we obtain for all $t \in (a,b)$:

$$\lim_{h\to 0^+}\frac{\partial g}{\partial s}(t+h,t) = Cv'(t)u(t) \qquad \text{and} \qquad \lim_{h\to 0^+}\frac{\partial g}{\partial s}(t-h,t) = Cu'(t)v(t)$$

implying (5.18). Thus, $g$ is a Green's function for $L_B$. ∎

**Remark 5.5.** Note that the condition that (5.20) only has the trivial solution need not be satisfied: e.g., the problem $-x'' = 0$, $x(0) = 0$, $x(1) - x'(1) = 0$ has the solution $x(s) = s$.

The proof of Theorem 5.4 is constructive. One can actually determine a Green's function by solving the initial value problems (5.21) and (5.22). Doing this, one can also see if (5.20) has non-trivial solutions. In the notation of the proof this is the case if and only if $v'u - u'v \equiv 0$.

Let us apply this construction to the following problem:

$$Lx := -x'' + x\,, \qquad B_1(x) := x(0)\,, \qquad B_2(x) := x(1)\,.$$

Then we obtain that

$$u(s) = \sinh(s)\,, \qquad v(s) = \sinh(1-s)\,,$$

$$p(v'u-u'v) = \sinh(1) \quad \Longrightarrow \quad C = \frac{1}{\sinh(1)}\,.$$

A Green's function is given by

$$g(s,t) = \frac{1}{\sinh(1)}\begin{cases}\sinh(s)\sinh(1-t)\,, & s \le t\,,\\ \sinh(1-s)\sinh(t)\,, & s > t\,.\end{cases}$$

The next theorem tells us, how to find a solution of the boundary value problem (5.11) for every $f \in C[0,1]$ using a Green's function:

**Theorem 5.6.** Let $L_B$ be defined as in (5.13) and assume that (5.20) only has the trivial solution $x = 0$. Moreover, let $G$ be the integral operator in $C[a,b]$ induced by a Green's function $g$ of $L_B$ (called **Green's operator**). Then: $G = L_B^{-1}$

**Proof:** It suffices to show the following three steps:

$$f \in C[a,b] \implies Gf \in D_B\,, \tag{5.24}$$
$$f \in C[a,b] \implies LGf = f\,, \tag{5.25}$$
$$x \in D_B \implies GLx = x\,. \tag{5.26}$$

Let $f \in C[a,b]$ and let

$$x(s) := (Gf)(s) = \int_a^s g(s,t)f(t)\,dt + \int_s^b g(s,t)f(t)\,dt\,, \qquad s \in [a,b]\,.$$

Since $g$ is twice continuously differentiable on both integration regions,

$$x'(s) = g(s,s)f(s) + \int_a^s \frac{\partial g}{\partial s}(s,t)f(t)\,dt + \int_s^b \frac{\partial g}{\partial s}(s,t)f(t)\,dt - g(s,s)f(s) \tag{5.27}$$

for all $s \in [a,b]$. Thus,

$$\begin{aligned} B_1(x) &= \alpha_1 x(a) + \beta_1 x'(a) = \int_a^b (\alpha_1 g(a,t) + \beta_1 \frac{\partial g}{\partial s}(a,t))f(t)\,dt \\ &= \int_a^b B_1(g(\cdot,t))f(t)\,dt = 0\,, \\ B_2(x) &= \alpha_2 x(b) + \beta_2 x'(b) = \int_a^b (\alpha_2 g(b,t) + \beta_2 \frac{\partial g}{\partial s}(b,t))f(t)\,dt \\ &= \int_a^b B_2(g(\cdot,t))f(t)\,dt = 0\,. \end{aligned}$$

Formula (5.27) for $x'$ and the smoothness conditions of $g$ imply that $x \in C^2[a,b]$ and, hence, $x \in D_B$. Thus, (5.24) holds. Furthermore,

$$\begin{aligned} x''(s) &= \lim_{h\to 0+} \frac{\partial g}{\partial s}(s+h,s)f(s) + \int_a^s \frac{\partial^2 g}{\partial s^2}(s,t)f(t)\,dt \\ &\quad - \lim_{h\to 0+} \frac{\partial g}{\partial s}(s-h,s)f(s) + \int_s^b \frac{\partial^2 g}{\partial s^2}(s,t)f(t)\,dt \\ &= \frac{f(s)}{p(s)} + \int_a^b \frac{\partial^2 g}{\partial s^2}(s,t)f(t)\,dt \end{aligned}$$

and, hence, (5.19) implies that

$$\begin{aligned} (Lx)(s) &= p(s)x''(s) + p'(s)x'(s) + q(s)x(s) \\ &= f(s) + \int_a^b p(s)\frac{\partial^2 g}{\partial s^2}(s,t)f(t)\,dt + \int_a^b p'(s)\frac{\partial g}{\partial s}(s,t)f(t)\,dt \\ &\quad + \int_a^b q(s)g(s,t)f(t)\,dt \\ &= f(s) + \int_a^b (L_s g)(s,t)f(t)\,dt = f(s) \end{aligned}$$

for all $s \in [a,b]$. Thus, (5.25) holds.

Let now $x \in D_B$ and $f := L_B x = Lx \in C[a,b]$. Then (5.25) implies that $LGf = f$, i.e., $LGLx = Lx$ and, hence,

$$L(GLx - x) = 0\,.$$

Since, due to (5.24), $GLx - x \in D_B$, this implies that

$$L_B(GLx - x) = 0\,.$$

Therefore, $GLx - x$ is a solution of (5.20). But then, due to the assumption that (5.20) only has the trivial solution, $GLx - x = 0$, i.e., (5.26) holds. ∎

**Remark 5.7.** Under the conditions of Theorem 5.6 problem (5.11) has a unique solution given by

$$Gf = \int_a^b g(\cdot,t)f(t)\,dt\,. \tag{5.28}$$

This also shows that the Green's function is then unique.

Maybe problem (5.20) has a nontrivial solution or it is hard to find the Green's function for $L_B$, but it is much easier to find the Green's function $\tilde{g}$ of $\tilde{L}_B$ with $\tilde{L}x := (px')' + \tilde{q}x$. Since $Lx = f$ is equivalent to $\tilde{L}x = f + (\tilde{q} - q)x$, (5.11) is equivalent to the integral equation

$$x(s) = \int_a^b \tilde{g}(s,t)(\tilde{q}(t) - q(t))x(t)\,dt + \int_a^b \tilde{g}(s,t)f(t)\,dt\,. \tag{5.29}$$

The price for the easier computation of the Green's function is that instead of the direct evaluation in (5.28) one has to solve the integral equation (5.29).

The Green's function $g$ satisfies the differential equation (5.19) for all fixed $t$ and $s \neq t$. In $s = t$ it is not differentiable. Formally, based on the theory of distributions, instead of (5.19) one often writes

$$(L_s g)(s,t) = \delta(s-t)\,, \qquad s,t \in [a,b]\,, \tag{5.30}$$

i.e., one assigns a *value* to $L_s g$ also at $s = t$. The $\delta$-distribution has the property that

$$\int_a^b \delta(s-t)f(t)\,dt = f(s)$$

should hold for sufficiently smooth $f$. Together with (5.30) this then formally yields that

$$(LGf)(s) = L\int_a^b g(s,t)f(t)\,dt = \int_a^b (L_s g)(s,t)f(t)\,dt = \int_a^b \delta(s-t)f(t)\,dt = f(s)$$

for all $s \in [a,b]$, i.e., that (5.25) holds. Equation (5.30) that can not be satisfied in the classical sense also formally implies that

$$\int_{t-h}^{t+h} \frac{\partial}{\partial s}\left(p(s)\frac{\partial g}{\partial s}(s,t)\right) ds + \int_{t-h}^{t+h} q(s)g(s,t)\,ds = \int_{t-h}^{t+h} \delta(s-t)\,ds = 1$$

for $h > 0$. Thus,

$$p\frac{\partial g}{\partial s}(\cdot,t)\Big|_{t-h}^{t+h} + \int_{t-h}^{t+h} q(s)g(s,t)\,ds = 1\,.$$

Since the second integral vanishes with $h \to 0$, we obtain that:

$$p(t)\left(\lim_{h\to 0+}\frac{\partial g}{\partial s}(t+h,t) - \lim_{h\to 0+}\frac{\partial g}{\partial s}(t-h,t)\right) = 1\,,$$

i.e., (5.18). This consideration can, of course, only be a heuristic motivation for (5.18). Checking the proof of Theorem 5.6, one can see that the jump in $\frac{\partial g}{\partial s}$, i.e., the exact formulation of (5.30), was essential for $L(Gf) = f$ to hold. This jump allows the reproduction of the inhomogeneity $f$.

## 5.2. Sturm-Liouville problem

Now we will study the eigenvalue problem (5.12), called **Sturm-Liouville problem**.

**Definition 5.8.** Let $L_B$ be defined as in (5.13). $\lambda \in \mathbb{C}$ is called **eigenvalue** of $L_B$ if $x \in D_B \setminus \{0\}$ exists with

$$L_B x = \lambda x\,.$$

$x$ is called **eigenfunction** of $L_B$.

So far we did not need condition (5.15) that will be required in the next theorem. If (5.15) is not satisfied, but $p > 0$, then the assertions are still true, however positive eigenvalues will be negative then and vice versa.

**Theorem 5.9.** Let $L_B$ be defined as in (5.13), assume that (5.15) holds and that $0$ is no eigenvalue of $L_B$, $g$ is the Green's function of $L_B$ with induced integral operator $G$ on $L^2[a,b]$. Then the following assertions hold:

(a) $G$ is selfadjoint.

(b) $\lambda \neq 0$ is an eigenvalue of $L_B$ if and only if $\frac{1}{\lambda}$ is an eigenvalue of $G$. Each eigenfunction to the eigenvalue $\lambda$ of $L_B$ is eigenfunction of $G$ to the eigenvalue $\frac{1}{\lambda}$ and vice versa. All eigenvalues of $L_B$ are real.

(c) $L_B$ (and $G$) has only finitely many negative eigenvalues.

(d) The set of eigenfunctions to each eigenvalue of $L_B$ (and $G$) is one-dimensional.

**Proof:**

(a) Let $x, y \in C[a,b]$. Then Theorem 5.6 and (5.14) imply that

$$\langle\, Gx, y \,\rangle = \langle\, GL_BGx, L_BGy \,\rangle = \langle\, L_BGL_BGx, Gy \,\rangle = \langle\, x, Gy \,\rangle\,.$$

Since $C[a,b]$ is dense in $L^2[a,b]$, this yields that $\langle\, Gx, y \,\rangle = \langle\, x, Gy \,\rangle$ holds for all $x, y \in L^2[a,b]$.

(b) This assertion immediately follows from Theorem 5.6, since $L_Bx = \lambda x$ is equivalent to $GL_Bx = \lambda Gx$, i.e., $\frac{1}{\lambda}x = Gx$ with $x \in C[a,b]$. According to Theorem 2.1, $\mathcal{R}(G) \subseteq C[a,b]$ so that all eigenfunctions of $G$ to eigenvalues different from 0 are in $C[a,b]$ and, hence, due to (5.24), in $D_B$. Since $G$ is selfadjoint, all eigenvalues are real.

(c) Let $x \in D_B$ be arbitrary, but fixed, with $\|x\|_2 = 1$. We will construct a lower bound for $\langle\, Lx, x \,\rangle$ that is independent from $x$: by partial integration we obtain that

$$\begin{aligned}\langle\, Lx, x \,\rangle &= \int_a^b (p(s)x'(s))'\overline{x(s)}\,ds + \int_a^b q(s)|x(s)|^2\,ds \\ &= px'\bar{x}\Big|_a^b - \int_a^b p(s)|x'(s)|^2\,ds + \int_a^b q(s)|x(s)|^2\,ds\,. \qquad (5.31)\end{aligned}$$

We will estimate each term on the right-hand side of (5.31) from below: since $\|x\|_2 = 1$, it holds that $\int_a^t |x(\tau)|^2\,d\tau \le 1$ for all $t \in (a,b]$. But then an $s \in (a,t)$ exists with $|x(s)| \le \frac{1}{\sqrt{t-a}}$, because otherwise $\int_a^t |x(\tau)|^2\,d\tau > \int_a^t \frac{1}{t-a}\,d\tau = 1$. With such an $s$ we obtain that

$$\begin{aligned}|x(a)| &\le |x(s)-x(a)| + |x(s)| \le \left|\int_a^s x'(\tau)\,d\tau\right| + \frac{1}{\sqrt{t-a}} \\ &\le \sqrt{\int_a^s 1\,d\tau \int_a^s |x'(\tau)|^2\,d\tau} + \frac{1}{\sqrt{t-a}} \le \sqrt{t-a}\,\|x'\|_2 + \frac{1}{\sqrt{t-a}}\end{aligned}$$

for all $t \in (a,b]$. Then this estimate also holds for the value $t = t_0$, where the right-hand side is minimized. This is the case for $t_0 := \min\{b, a + \|x'\|_2^{-1}\}$. If $t_0 = b$, then $a + \|x'\|_2^{-1} \ge t_0$, i.e., $(t_0 - a)\,\|x'\|_2 \le 1$ and, hence, $|x(a)| \le \frac{2}{\sqrt{b-a}}$. If $t_0 = a + \|x'\|_2^{-1}$, then $|x(a)| \le 2\sqrt{\|x'\|_2}$. Combining both cases, we obtain that

$$|x(a)|^2 \le 4\,\|x'\|_2 + \frac{4}{b-a}\,.$$

If $x(a) = 0$, then obviously $(px'\bar{x})(a) = 0$. Otherwise, due to (5.9),

$$0 = B_1(x) = \alpha_1 x(a) + \beta_1 x'(a) \quad \Longrightarrow \quad \beta_1 \neq 0 \quad \text{and} \quad x'(a) = -\frac{\alpha_1}{\beta_1}x(a)$$

and, hence,

$$\begin{aligned}|(px'\bar{x})(a)| &\le |p(a)|\left|\frac{\alpha_1}{\beta_1}x(a)\right|\overline{|x(a)|} = |p(a)|\left|\frac{\alpha_1}{\beta_1}\right||x(a)|^2 \\ &\le C_1\,\|x'\|_2 + C_2 \qquad (5.32)\end{aligned}$$

for some constants $C_1, C_2 > 0$ depending on $p, \alpha_1, \beta_1$. Analogously, one can show that constants $C_3, C_4 > 0$ depending on $p, \alpha_2, \beta_2$ exist such that

$$|(px'\bar{x})(b)| \leq C_3 \|x'\|_2 + C_4 \,. \tag{5.33}$$

Due to (5.15) and the continuity of $q$, constants $C_5 > 0$ and $C_6 \in \mathbb{R}$ exist with

$$-p(s) \geq C_5 \qquad \text{and} \qquad q(s) \geq C_6 \,, \qquad s \in [a,b] \,. \tag{5.34}$$

Combining (5.31), (5.32), (5.33), (5.34), and $\|x\|_2 = 1$ yields

$$\begin{aligned} \langle\, Lx, x \,\rangle &\geq -C_1 \|x'\|_2 - C_2 - C_3 \|x'\|_2 - C_4 + C_5 \|x'\|_2^2 + C_6 \\ &= \left( \sqrt{C_5}\, \|x'\|_2 - \frac{C_1 + C_3}{2\sqrt{C_5}} \right)^2 + C_6 - C_2 - C_4 - \frac{(C_1 + C_3)^2}{4C_5} \\ &\geq C_6 - C_2 - C_4 - \frac{(C_1 + C_3)^2}{4C_5} =: C \in \mathbb{R} \,. \end{aligned} \tag{5.35}$$

Let us now assume that $\lambda$ is an eigenvalue of $L_B$ with corresponding eigenfunction $x$ with $L^2$-Norm 1. Then (5.35) implies that

$$C \leq \langle\, Lx, x \,\rangle = \langle\, \lambda x, x \,\rangle = \lambda \|x\|^2 = \lambda \,.$$

This means that all eigenvalues of $L_B$ are bounded from below by $C$. If $\lambda < 0$ is an eigenvalue of $L_B$, then $C < 0$ and $\frac{1}{\lambda} \leq \frac{1}{C} < 0$. But $\frac{1}{\lambda}$ is an eigenvalue of $G$. Therefore, due to Theorem 2.33 (c), there can only exist finitely many negative eigenvalues, because otherwise there would exist an accumulation point of the eigenvalues of $G$ in $(-\infty, \frac{1}{C}]$.

(d) Let $\lambda \neq 0$ be an eigenvalue of $L_B$ and assume that $x$ and $y$ are two eigenfunctions. Since $B_1(x) = B_1(y) = 0$, (5.9) implies that the system

$$\begin{pmatrix} x(a) & x'(a) \\ y(a) & y'(a) \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0$$

has the nontrivial solution $\alpha_1, \beta_1$. This means that the vectors $(x(a), x'(a))$ and $(y(a), y'(a))$ are linearly dependent. Therefore, $\gamma$ and $\delta$ exist with $\gamma^2 + \delta^2 \neq 0$ so that

$$\gamma(x(a), x'(a)) + \delta(y(a), y'(a)) = (0,0) \,.$$

But then $u := \gamma x + \delta y$ solves the initial value problem

$$\begin{aligned} &(Lu - \lambda u)(s) = 0 \,, \qquad s \in [a,b] \,, \\ &u(a) = 0 \,,\, u'(a) = 0 \,. \end{aligned}$$

Now Theorem 5.1 implies that $u \equiv 0$. Thus, $x$ and $y$ are linearly dependent.

∎

**Corollary 5.10.** Let $L_B$ be defined as in (5.13), assume that (5.15) holds and that $0$ is no eigenvalue of $L_B$. Then the following assertions hold:

$L_B$ has countably many eigenvalues $\lambda_n$, with $\infty$ as the only accumulation point. For each $\lambda_n$ there exists (up to the sign) only one eigenfunction $\varphi_n$ with $\|\varphi_n\| = 1$. The set $\{\varphi_n : n \in \mathbb{N}\} \subseteq D_B$ is a complete orthonormal system in $L^2[a,b]$. Moreover, the Green's function of $L_B$ is given by

$$g(s,t) = \sum_{n=1}^{\infty} \frac{1}{\lambda_n} \varphi_n(s) \overline{\varphi_n(t)}\,, \tag{5.36}$$

for all $s,t \in [a,b]$, where the convergence is absolute and uniform. Problem (5.11) is uniquely solvable for all $f \in C[a,b]$ and this unique solution may be represented via the absolutely and uniformly convergent series

$$x(s) = \sum_{n=1}^{\infty} \frac{1}{\lambda_n} \langle\, f, \varphi_n \,\rangle\, \varphi_n(s)\,. \tag{5.37}$$

Finally,

$$\sum_{n=1}^{\infty} \frac{1}{\lambda_n} < \infty\,.$$

**Proof:** Due to Theorem 5.9 (a) and (c), we may apply Theorem 2.47 in combination with Remark 2.48 to the integral operator $G$ with kernel $g$. Then the assertions follow together with Theorem 5.6 and Theorem 5.9 (b) and (d).

Note that, due to Theorem 5.6, $\mathcal{R}(G) = D_B$. Therefore, Theorem 2.37 implies that the functions $\{\varphi_n\}$ are a complete orthonormal system in $\overline{\mathcal{R}(G)} = L^2[a,b]$.

The representation (5.37) follows from Theorem 2.44 and the fact that, due to Theorem 5.6, $Gf$ is the unique solution of (5.11). ∎

**Remark 5.11.** The corollary above says that, based on the solution theory for integral equations, one can derive a solution formula for inhomogeneous Sturm-Liouville boundary value problems using eigenvalues and eigenfunctions, leading to the absolutely and uniformly convergent series (5.37).

It is also possible to interpret (5.37) in such a way that the series

$$x(s) = \sum_{n=1}^{\infty} \langle\, x, \varphi_n \,\rangle\, \varphi_n(s)$$

is absolutely and uniformly convergent for all $x \in D_B$. For the special case $Lx := -x''$ this yields assertions about the absolute and uniform convergence of Fourier series (see Example 5.12). In addition, Corollary 5.10 provides the possibility to construct several different complete orthonormal systems in $L^2[a,b]$, namely eigenfunctions of Sturm-Liouville problems.

**Example 5.12.** Let $Lx := -x''$, $B_1(x) := x(0)$ and $B_2(x) := x(\pi)$, i.e.,

$$D_B = \{x \in C^2[0,\pi] : x(0) = x(\pi) = 0\}\,.$$

Since the general solution of $Lx = \lambda x$ is given by

$$x(s) = a\cos\left(\sqrt{\lambda}s\right) + b\sin\left(\sqrt{\lambda}s\right)\,,$$

it follows from the boundary conditions that nontrivial solutions of $L_B x = \lambda x$ can only exist if $a = 0$ and $\sqrt{\lambda} \in \mathbb{N}$. Thus, the eigenvalues of $L_B$ are the numbers $n^2$ with $n \in \mathbb{N}$. The corresponding normed eigenfunctions are given by $\sqrt{\frac{2}{\pi}}\sin(ns)$. Therefore, it follows with (5.36) that the Green's function $g$ of $L_B$ is given by

$$g(s,t) = \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\sin(ns)\sin(nt)}{n^2}\,,$$

where this series converges absolutely and uniformly in $[0,\pi]^2$. If one computes $g$ via the construction suggested in the proof of Theorem 5.4, one obtains

$$g(s,t) = \frac{1}{\pi}\begin{cases} s(\pi - t)\,, & s \le t\,,\\ t(\pi - s)\,, & s \ge t\,.\end{cases}$$

The solution $x$ of

$$\begin{aligned} -x''(s) &= f(s)\,, \qquad s \in [0,\pi]\,,\\ x(0) &= x(\pi) = 0 \end{aligned}$$

with $f \in C[a,b]$ can be directly calculated via integration or written as an absolutely and uniformly convergent series (see (5.37))

$$x(s) = \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{1}{n^2}\int_0^{\pi} f(t)\sin(nt)\,dt\,\sin(ns)\,.$$

This is a representation in terms of the Fourier coefficients of $f$. By the way, it is a consequence of (2.67) that

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = \int_0^{\pi} g(s,s)\,ds = \frac{1}{\pi}\int_0^{\pi} s(\pi - s)\,ds = \frac{\pi^2}{6}\,.$$

**Remark 5.13.** Let us assume that the conditions of Theorem 5.9 are satisfied and that $g$ is the Green's function of $L_B$, then one can transform problem

$$Lx = \lambda x + f\,, \qquad B_1(x) = B_2(x) = 0\,, \tag{5.38}$$

with $f \in C[a,b]$ and $\lambda \in \mathbb{R}$ into the equivalent integral equation

$$x - \lambda Gx = Gf \quad \Longleftrightarrow \quad \frac{1}{\lambda}x - Gx = \frac{1}{\lambda}Gf \in \mathcal{R}(G) \subseteq \mathcal{N}(G)^{\perp}\,. \tag{5.39}$$

If $\lambda_n$ and $\varphi_n$ are as in Corollary 5.10 and if $\lambda$ is no eigenvalue of $L_B$, then Corollary 2.39 implies that the solution of (5.39) and, hence, also of (5.38), is given by

$$x = \sum_{n=1}^{\infty}\frac{\left\langle \frac{1}{\lambda}Gf, \varphi_n \right\rangle}{\frac{1}{\lambda} - \frac{1}{\lambda_n}}\varphi_n\,. \tag{5.40}$$

Note that $\{\frac{\varphi_n}{\lambda_n} : n \in \mathbb{N}\}$ is an eigensystem for $G$. Since

$$\left\langle \frac{1}{\lambda} Gf, \varphi_n \right\rangle = \frac{1}{\lambda} \langle f, G\varphi_n \rangle = \frac{1}{\lambda\lambda_n} \langle f, \varphi_n \rangle ,$$

(5.40) implies that

$$x = \sum_{n=1}^{\infty} \frac{\langle f, \varphi_n \rangle}{\lambda_n - \lambda} \varphi_n .$$

Using Example 5.12, we obtain that, if $\lambda \neq n^2$ for $n \in \mathbb{N}$ and if $f \in C[0, \pi]$, then the unique solution of

$$\begin{aligned} -x''(s) &= \lambda x(s) + f(s) , \qquad s \in [0, \pi] , \\ x(0) &= x(\pi) = 0 , \end{aligned} \tag{5.41}$$

is given by

$$x(s) = \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{n^2 - \lambda} \int_0^{\pi} f(t) \sin(nt)\, dt \sin(ns) .$$

In a similar way one can use (2.53) to find conditions for the solvability of (5.41) and a representation of solutions if $\lambda = n^2$.

Using this study of problem (5.38), one can see what one can do if 0 is an eigenvalue of $L_B$: one then considers the operator $\tilde{L}x = L_B x - \lambda x$, where $\lambda$ should not be an eigenvalue of $L_B$. Since only countably many eigenvalues exist accumulating in $\infty$, one can easily find such a $\lambda$ (compare Remark 5.7). 0 is now no eigenvalue of the operator $\tilde{L}_B$. Using its Green's function and the fact that $L_B x = f$ is equivalent to $\tilde{L}_B x = f - \lambda x$, one can solve $L_B x = f$.

A similar theory as for Sturm-Liouville problems treated above can be developed for more general problems of the kind

$$\begin{aligned} \frac{1}{r(s)}[(p(s)x'(s))' + q(s)x(s)] &= f(s) , \qquad s \in [a, b] , \\ \alpha_1 x(a) + \beta_1 x'(a) + \gamma_1 x(b) + \delta_1 x'(b) &= 0 , \\ \alpha_2 x(a) + \beta_2 x'(a) + \gamma_2 x(b) + \delta_2 x'(b) &= 0 , \end{aligned}$$

where $r$ is positive and continuous and the boundary conditions are independent. One then uses the inner product

$$\langle f, g \rangle_r := \int_a^b f(t)\overline{g(t)} r(t)\, dt$$

in the space $L^2([a, b]; r) := \{f : \langle f, f \rangle_r < \infty\} = L_2[a, b]$. After finding the Green's function $g$, the induced integral operator is given by

$$(Gf)(s) := \int_a^b r(t) g(s, t) f(t)\, dt .$$

Even the case $r(a) = 0$ can be treated which is essential for Bessel's differential equation

$$x''(s) + \frac{1}{s} x'(s) + \left(\lambda - \frac{r^2}{s^2}\right) x(s) = 0$$

in $[0,1]$. It can also be written in the form

$$\frac{1}{s}\left((sx'(s))' - \frac{r^2}{s}x(s)\right) = -\lambda x(s)\,.$$

**Remark 5.14.** The problem – find $q \in C[0,1]$ such that the Sturm-Liouville problem

$$\begin{aligned} -x''(s) + q(s)x(s) &= 0\,, \qquad s \in [0,1]\,, \\ x(0) = x(1) &= 0\,, \end{aligned}$$

has a prescribed sequence $\lambda_1, \lambda_2, \lambda_3, \ldots$ as eigenvalues – is called **inverse Sturm-Liouville problem**. A possible interpretation is as follows: What mass distribution of a clamped string guarantees prescribed eigenfrequencies? The results of this chapter show that a solution is only possible if the sequence $(\lambda_n)$ satisfies the following conditions: it has at most finitely many negative values, it accumulates in $\infty$, and $\sum_{n=1}^{\infty} \frac{1}{\lambda_n}$ is convergent.

# References


[1] P. Anselone, Collectively Compact Operator Approximation Theory, Prentice-Hall, Englewood Cliffs, New Jersey, 1971.

[2] K. E. Atkinson, A Survey of Numerical Methods for the Solution of Fredholm Integral Equations of the Second Kind, SIAM, Philadelphia, 1976.

[3] C. T. Baker, The Numerical Treatment of Integral Equations, Clarendon Press, Oxford, 1977.

[4] H. W. Engl, Integralgleichungen, Springer, Wien, 1997.

[5] R. Kress, On the Fredholm alternative, Integral Equations and Operator Theory 6 (1983), 453-457.

[6] R. Kress, Linear Integral Equations, Springer, Berlin, 1989.

# Index