МЕХАНИКО-
МАТЕМАТИЧЕСКИЙ
ФАКУЛЬТЕТ
МГУ ИМЕНИ
М.В. ЛОМОНОСОВА

teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ

Functional Analysis and Theory of Operators

ШЕЙПАК
ИГОРЬ АНАТОЛЬЕВИЧ

МЕХМАТ МГУ

БЛАГОДАРИМ ЗА ПОДГОТОВКУ КОНСПЕКТА
АСПИРАНТА МЕХАНИКО-МАТЕМАТИЧЕСКОГО ФАКУЛЬТЕТА МГУ
РЫХЛОВА ВЛАДИСЛАВА ВЛАДИМИРОВИЧА

Содержание

Lecture 1. Basics of Functional Analysis. Metric Spaces

Metric Spaces. Examples of Metric Spaces

Definition 1.1. (X,ρ), *where* X *is an arbitrary set and* $\rho : X \times X \to [0,+\infty)$, *is called a* ***metric space****, if* ρ *satisfies*

1) $\rho(x,y) = 0$ *iff* $x = y$,

2) $\rho(x,y) = \rho(y,x)$,

3) $\rho(x,y) \leqslant \rho(x,z) + \rho(z,y)$ *(the triangle inequality).*

One of the central concepts in Functional Analysis is the notion of a complete metric space, defined as follows:

Definition 1.2. *A metric space* (X,ρ) *is called* ***complete*** *if for any Cauchy sequence* $\{x_n\}_{n=1}^{\infty}$ *there exists* $\lim\limits_{n\to\infty} x_n = x \in X$.

Now we demonstrate some fundamental examples of metric spaces.

Example 1.1. $\mathbb{R}^n$ *(or* $\mathbb{C}^n$*) with coordinates* $x = (x_1, x_2, \ldots, x_n)$ *endowed with a standard Euclidean metric*

$$\rho(x,y) = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}.$$

In further, when we mention some metric spaces, they are assumed to be endowed with a certain (standard in some sense) metric, so we omit the explicit notation of the given metric.

$\mathbb{R}^n$ and $\mathbb{C}^n$ above serve as examples of finite-dimensional metric spaces, while the main objects, which are studied in Functional Analysis, are infinite-dimensional metric spaces. Let us look at the following examples.

Example 1.2. *Consider the following spaces of sequences:*

a) c_{00}, *which is the space of infinite sequences* $x = (x_1, x_2, \ldots, x_n, 0, 0, \ldots)$ *with a finite number of nonzero coordinates (this number may be different for distinct elements of the space):*

$$\forall x \in c_{00}\ \exists n = n(x):\ \forall k > n\ x_k = 0.$$

b) c_0, *which is the space of infinite sequences* $x = (x_1, x_2, \ldots, x_n, \ldots)$ *such that*

$$\lim_{n\to\infty} x_n = 0.$$

c) c, which is the space of infinite sequences $x = (x_1, x_2, \ldots, x_n, \ldots)$ such that

$$\exists \lim_{n\to\infty} x_n = a \equiv a(x).$$

These are examples of infinite-dimensional metric spaces. The standard metric is given by $\rho(x,y) = \sup_{i\geqslant 1}|x_i - y_i|$. It can be easily seen that $c_{00} \subset c_0 \subset c$.

What can we say about the completeness of these spaces in examples above? $\mathbb{R}^n$ and $\mathbb{C}^n$, being finite-dimensional spaces, are obviously complete, since the convergence there is in fact the coordinate-wise convergence. Let us define the convergence in a generic metric space.

Definition 1.3. $x_n \xrightarrow{\rho} x$ *in* (X,ρ) *if* $\rho(x_n, x) \to 0$.

In the first example, the convergence with respect to the metric is just the coordinate-wise convergence.

What can we say about the space c_{00}?

Exercise 1.1. *Prove that c_{00} is not complete.*

An example proving that this space is incomplete can be constructed by adding something small to further and further coordinates, for instance,

$$x^1 = \Big(1, 0, 0, 0, \ldots\Big),$$
$$x^2 = \Big(1, \frac{1}{2}, 0, 0, \ldots\Big),$$
$$\ldots$$
$$x^n = \Big(1, \frac{1}{2}, \ldots, \frac{1}{n}, 0, \ldots\Big).$$

$\{x^n\}_{n=1}^{\infty}$ *is a Cauchy sequence:*

$$\rho(x^n, x^m) = \frac{1}{\min(n,m)+1} \to 0 \quad as \quad n,m\to\infty$$

(note that we have the supremum metric, and not ℓ_2-metric!). By the convergence with respect to metric in c_{00}, c_0, and c, it follows that $\forall k\ x_k^n \to x_k$, so the limit sequence is harmonic: $x = (1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots)$, which is not finite, therefore, it does not belong to c_{00}.

Let us proceed to the following examples.

Example 1.3. *Consider $\ell_p(n)$, $1 \leqslant p < \infty$, the space of finite-dimensional vectors $x = (x_1, \ldots, x_n)$, $x_j \in \mathbb{R}$ (or $\mathbb{C}$), with metric*

$$\rho(x,y) = \Big(\sum_{i=1}^{n} |x_i - y_i|^p\Big)^{1/p};$$

if we take the limit with respect to the parameter p*, as* $p \to \infty$*, then, for* $p = \infty$*, we have*

$$\rho(x,y) = \max_{i \geqslant 1} |x_i - y_i|.$$

It is clear that these functions $\rho(x,y)$ *are indeed metrics in the spaces* $\ell_p(n)$*: they are symmetric, nonnegative, take zero values only for coinciding elements* ($x = y$)*, and the corresponding triangle inequalities are simply the Minkowski inequalities.*

Example 1.4. *Consider* ℓ_p*,* $1 \leqslant p < \infty$*, the space of infinite sequences* $x = (x_1, \dots, x_n, \dots)$*,* $x_j \in \mathbb{R}$ *(or* $\mathbb{C}$*), such that*

$$\sum_{i=1}^{n} |x_i|^p < \infty$$

for $p < \infty$ *and*

$$\sup_{i \geqslant 1} |x_i| < \infty$$

for $p = \infty$*. The metric is given by*

$$\rho(x,y) = \Big(\sum_{i=1}^{n} |x_i - y_i|^p \Big)^{1/p}$$

for $p < \infty$ *and*

$$\rho(x,y) = \sup_{i \geqslant 1} |x_i - y_i|$$

for $p = \infty$*.*

The following example is represented by the space of functions.

Example 1.5. *Consider* $C[a,b]$*, the space of continuous functions with the (uniform) metric*

$$\rho(f,g) = \max_{[a,b]} |f(x) - g(x)|.$$

These metric spaces ($\ell_p(n)$, ℓ_p, and $C[a,b]$) are complete, though this property can be violated if we define the metric in the space of continuous functions in the following way:

Example 1.6. *Consider* $C_p[a,b]$*, the space of continuous functions, where the parameter* p *indicates that we use the integral metric*

$$\rho(f,g) = \Big(\int_a^b |f(x) - g(x)|^p \, dx \Big)^{1/p};$$

as the functions are continuous, the integral is the Riemann integral. If we take the limit as $p \to \infty$*, we immediately obtain the previous example, i.e. the space* $C[a,b]$ *of continuous functions with the uniform metric.*

For $1 \leqslant p < \infty$, these spaces are not complete.

Exercise 1.2. *Prove that $C_1[0,1]$ is not complete.*

We can construct a sequence $\{f_n\}_{n=1}^{\infty}$ of continuous functions such that $f_n \equiv 1$ for $x \leqslant 1/2$, f_n decreases to zero on $[\frac{1}{2}, \frac{1}{2}+\frac{1}{n}]$, and $f_n \equiv 0$ for $x \geqslant \frac{1}{2}+\frac{1}{n}$. This sequence is

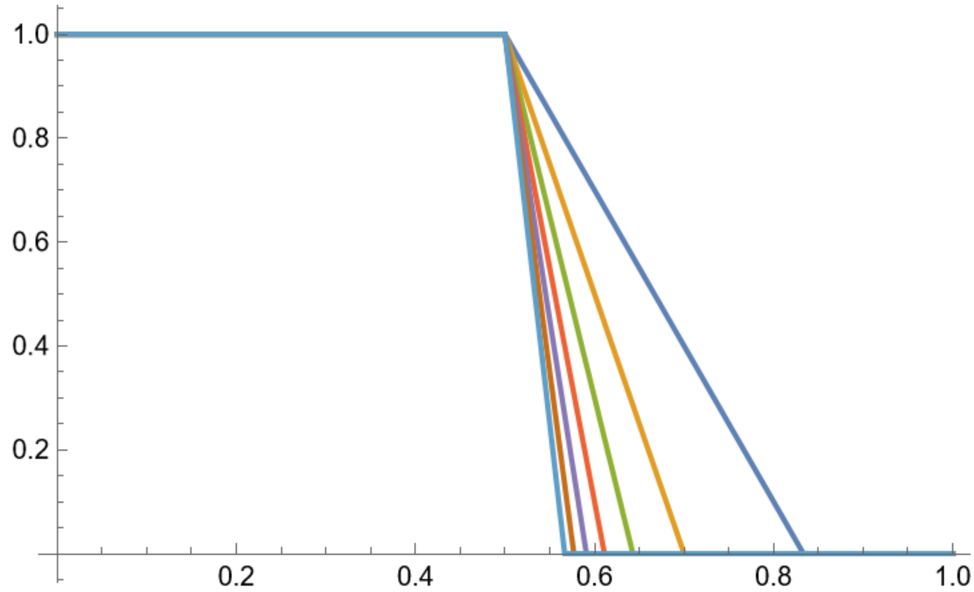

Рис. 1.1. Graphs of f_n, $n = 3,5,7,9,11,13,15$.

obviously a Cauchy sequence: $\rho(f_n, f_m)$ is dominated by the square of the triangle with vertices $(1/2, 1)$, $(1/2+1/n, 0)$, and $(1/2+1/m, 0)$, that is,

$$\rho(f_n, f_m) = \frac{1}{2}\left|\frac{1}{n} - \frac{1}{m}\right| \to 0$$

as $n, m \to \infty$. With respect to the given metric, f_n converges to an indicator function $\chi_{[0,\frac{1}{2}]}$ of the interval $[0, \frac{1}{2}]$, which is not continuous, so the space $C_1[0,1]$ is incomplete (since the metric is integral, we must identify the functions that are equal almost everywhere, but since we are in the space of continuous functions, this means that "almost everywhere" is equivalent to "everywhere", so the limit function is unique).

In the following example, we consider the spaces of differentiable (smooth) functions.

Example 1.7. *Consider $C^n[a,b]$, the space of functions f such that $\forall j = 0,1,\dots,n$: $f^{(j)} \in C[a,b]$. We can endow this space with either of metrics*

$$\rho_1(f,g) = \sum_{j=0}^{n} \max_{[0,1]} |f^{(j)}(x) - g^{(j)}(x)|$$

or

$$\rho_2(f,g) = \max_{0\leqslant j\leqslant n}\max_{[0,1]}|f^{(j)}(x) - g^{(j)}(x)|.$$

These metrics are equivalent since $\rho_2 \leqslant \rho_1 \leqslant (n+1)\rho_1$ *(so when replacing one metric with the other, we just change the geometry of our space, while the convergence properties remain the same). These spaces are complete.*

Consider more complicated examples.

Example 1.8. *Consider* (Ω, M, μ), *where* Ω *is the universal set,* M *is a* σ*-algebra, and* μ *is a* σ*-finite measure. We can define the space of measurable functions* $L_p(\Omega,\mu)$*:*

$$f \in L_p(\Omega,\mu) \ \textit{if} \int\limits_{\Omega} |f(x)|^p \, d\mu < \infty, \quad 1 \leqslant p < \infty,$$

and $f \in L_\infty(\Omega,\mu)$ *if* $\operatorname{ess\,sup}|f(x)| < \infty$, *i.e. the function is bounded almost everywhere, meaning that*

$$\operatorname{ess\,sup}|f(x)| = \inf_{\mu(A)=0}\sup_{\Omega\setminus A}|f(x)|.$$

For $1 \leqslant p < \infty$, *the metric is defined by* $\rho(f,g) = \Big(\int\limits_{\Omega}|f-g|^p\,d\mu\Big)^{1/p}$*; for* $p=\infty$ *it is defined by* $\rho(f,g) = \operatorname{esssup}|f(x)-g(x)|$. *These spaces are complete.*

Example 1.9 (Sobolev spaces, one-dimensional case)**.** *Consider*

$$W_p^n[a,b] = \{f \ \textit{such that} \ \forall j = 0,1,\ldots,n-1 \ f^{(j)} \in AC[a,b], \ f^{(n)} \in L_p[a,b]\},$$

where $AC[a,b]$ *is the space of absolutely continuous functions. For* $1\leqslant p<\infty$, *the metric can be defined as follows:*

$$\rho(f,g) = \Big(\sum_{j=0}^{n}\int\limits_{[a,b]}|f^{(j)}(x) - g^{(j)}(x)|^p\,d\mu\Big)^{1/p},$$

and for $p=\infty$, *the integral must be replaced with the essential supremum. These spaces are complete.*

Example 1.10. *Discrete metric space* X_{discr}. *Let* X *be an arbitrary set, and let the metric be defined by*

$$\rho(x,y) = \begin{cases} 1, & \textit{if } x \neq y, \\ 0, & \textit{if } x = y. \end{cases}$$

In this metric, all Cauchy sequences are simply stabilizing sequences:

$$x_1, x_2, \ldots, x_N, a, a, \ldots, a, \ldots$$

Thus, this space is obviously complete since $a \in X$. *In the topology associated with the given metric, every set is open.*

Limit and Closure Points. Closure of a Set. Separable Spaces

Let us remind the definition of open and closed subsets of the metric space.

Definition 1.4. *Let (X,ρ) be a metric space and $M \subset X$. M is **open** if $\forall x \in M\ \exists \varepsilon > 0$: $B(x,\varepsilon) \subset M$, where $B(x,\varepsilon) = \{y \in X : \rho(y,x) < \varepsilon\}$. $M \subset X$ is **closed** if $X \backslash M$ is open.*

According to this definition, a single point $\{a\} \subset X$ is an open subset of X_{discr}; any union of open sets is open, so any subset of X is an open set in the metric space X_{discr}.

Another definition of the closed subset can be given in terms of limit points of the set. Let us recall some definitions.

Definition 1.5. *x_0 is called a **limit point** of a set $M \subset (X,\rho)$ if $\forall \varepsilon > 0$ $B(x_0,\varepsilon) \cap M$ contains infinitely many points of M.*

Definition 1.6. *x_0 is called a **closure point** of a set $M \subset (X,\rho)$ if $\forall \varepsilon > 0$ $B(x_0,\varepsilon) \cap M \neq \varnothing$.*

Definition 1.7. *The **closure** of a set $M \subset (X,\rho)$ is $\overline{M} = M \cup \{\text{all limit points}\} = \{\text{all closure points of } M\}$.*

Let us recall some other definitions from Functional Analysis.

Definition 1.8. *A set $M \subset (X,\rho)$ is **dense** in X if $\overline{M} = X$.*

Definition 1.9. *A metric space (X,ρ) is called **separable** if there exists a countable or finite dense subset of X.*

Note that the condition of finiteness of the dense subset is reserved specifically for discrete metric such as in X_{discr}.

Next, we shall point out which of spaces in the examples above are separable and which are not.

1) X_{discr} is separable if X_{discr} is finite or countable.

2) $C[a,b]$ is separable since for every $f \in C[a,b]$ and any $\varepsilon > 0$ there exists a polynomial p with rational coefficients such that $\|f-p\|_{C[a,b]} < \varepsilon$ (see the Weierstrass approximation theorem):

$$p = \sum_{i=0}^{n} c_i x^i, \qquad c_i \in \mathbb{Q}.$$

3) $L_p(X,\mu)$, $1 \leqslant p < \infty$, are separable if the measure μ is σ-additive.

4) ℓ_∞ is not separable.

Exercise 1.3. *Prove that ℓ_∞ is not separable.*

Lemma 1.1. *Let (X,ρ) be a metric space. If there exists an uncountable $M \subset X$ such that $\exists d > 0\ \forall x,y \in M$: $\rho(x,y) \geqslant d$, then X is not separable.*

Proof *by contradiction. Assume that X is separable, then*

$$\exists X_0 \subset X, \text{ finite or countable, such that } \overline{X}_0 = X.$$

This is equivalent to the following property. For $\varepsilon > 0$, consider balls with centers at x of radii ε. Thus,

$$\cup_{x \in X_0} B(x,\varepsilon) = X.$$

The number of the balls in this union has the same cardinality as X_0, i.e. it is finite or countable. But M (see the condition of the lemma) is not countable, so $\exists B(x_0,\varepsilon) \supset \{x,y\}$, $x,y \in M$. Take $\varepsilon = d/3$; then

$$d \leqslant \rho(x,y) \leqslant \rho(x,x_0) + \rho(x_0,y) \leqslant 2d/3,$$

where the first inequality is due to property of the set M, and the second one is due to the triangle inequality, which gives us a contradiction. □

If we would like to use this lemma to prove that ℓ_∞ is not separable, then we have to find a subset of ℓ_∞ with the property described. Consider the set of sequences

$$M = \{x = (x_1, x_2, \dots, x_n, \dots) \text{ such that } \forall k: \ x_k \in \{0,1\}\}.$$

This set is uncountable; one can show it by employing Cantor's diagonal method (if we suppose that this set is countable, we can write it in the form of a table; then, we pick the diagonal and change any symbol of the diagonal to the opposite; there is no such an element in this table, so the set is uncountable. This method is usually used to prove that $\mathbb{R}$ is not countable in Calculus) and $\rho(x,y) = 1$ as $x \neq y$, so this set satisfies the conditions of the lemma.

Maps of Metric Spaces

Let (X,ρ) and (Y,d) be metric spaces. Consider the map $(X,\rho) \xrightarrow{f} (Y,d)$. We will focus on the following kinds of maps:

1) f is continuous at a point $x_0 \in X$,

2) f is continuous on X,

3) f is uniformly continuous on X,

4) f is Lipschitz continuous on X. (Recall that it means that

$$\exists r \geqslant 0: \sup_{x,y \in X:\ x \neq y} \frac{d(f(x), f(y))}{\rho(x,y)} = r < \infty,$$

and r is called a Lipschitz constant).

For instance, in the existence and uniqueness theorem for the solution of Ordinary Differential Equation (namely, the Cauchy problem) there are Lipschitz continuous functions considered as a right-hand side of the equation; for the Cauchy problem

$$y' = G(x,y),$$
$$y(x_0) = y_0$$

to be uniquely solvable, we must require that $G(x,y)$ is Lipschitz continuous with respect to y.

5) f is contraction:

Definition 1.10. *$f : (X,\rho) \to (Y,d)$ is called a* ***contraction*** *if f is Lipschitz continuous with $r \in [0,1)$.*

6) f is isometry:

   a) f is a **complete** isometry if f is a bijection $X \to Y$ and $d(f(x), f(y)) = \rho(x,y)$.

   b) f is a **partial** isometry if f is not a bijection, while $d(f(x), f(y)) = \rho(x,y)$ holds.

These are the most important properties of maps of metric spaces.

Properties of Complete Metric Spaces

The main property is that we can take a limit and guarantee that the limit element has the same properties as the elements of the sequence converging to it. For instance, we know that the space of (n times) differentiable functions is complete; thus, taking a limit of a sequence of differentiable functions we can only obtain a differentiable function.

Theorem 1.1 (fixed-point theorem)**.** *Let (X,ρ) be a complete metric space, and $f: X \to X$ be a contraction mapping. Then*

$$\exists! x^* \in X: \ f(x^*) = x^*.$$

Example 1.11 (of incomplete space for which this theorem is not valid)**.** *Consider a real axis with zero excluded, $\mathbb{R}\backslash\{0\}$, with a standard metric $\rho(x,y) = |x-y|$. Consider a contraction $f(x) = \frac{x}{2}$. On $\mathbb{R}$, it has 0 as a fixed point; when we exclude 0 from the space $\mathbb{R}$, it becomes incomplete, and, at the same time, it looses the fixed point of the given contraction.*

Idea of the proof. Let x_0 be an arbitrary start point. Take

$$\begin{gathered} x_1 = f(x_0), \\ x_2 = f(x_1) = f(f(x_0)), \\ \ldots, \\ x_n = f(x_{n-1}), \\ x_{n+1} = f(x_n), \\ \ldots, \end{gathered}$$

so we obtain a sequence $\{x_n\}_{n=1}^{\infty}$. We can prove that this sequence is a Cauchy sequence using the contraction properties of f, therefore, there exists

$$x^* = \lim_{n\to\infty} x_n.$$

We can prove that $f(x^*) = x^*$, and then prove that if there is another point y^* such that $y^* = f(y^*)$, then $x^* = y^*$.

To formulate the following theorem, we have to define the system of nested closed balls.

Definition 1.11. $B_n = B[x_n, r_n]$, *such that* $B_1 \supseteq B_2 \supseteq \cdots \supseteq B_n \supset B_{n+1}$ *is called a* ***system of nested closed balls****.*

Remark on notation. $B(x,\varepsilon) = \{y \in X : \rho(x,y) < \varepsilon\}$ denotes an open ball and $B[x,\varepsilon] = \{y \in X : \rho(x,y) \leqslant \varepsilon\}$ denotes a closed ball.

Theorem 1.2. *Let (X,ρ) be a metric space. It is complete iff $\forall \{B_n\}_{n=1}^{\infty}$ (system of nested closed balls) with radii $r_n \to 0$*

$$\exists! x^* = \cap_{n=1}^{\infty} B_n.$$

Proof. $\Rightarrow$. Let (X,ρ) be complete. Let $\{B_n\}_{n=1}^{\infty}$ be our system of nested closed balls with $r_n \to 0$. Consider a sequence $\{x_n\}_{n=1}^{\infty}$ of centers. This sequence is a Cauchy sequence:

$$\rho(x_n, x_m) \underset{n>m}{\leqslant} r_n \to 0 \text{ as } n \to \infty,$$

therefore, since (X,ρ) is complete,

$$\exists x^* := \lim_{n\to\infty} x_n.$$

As it is the limit of x_n, and the intersection $\cap B_n$ of all balls is closed, x^* is a limit point of this set. Thus,

$$x^* \in \cap_{n=1}^{\infty} B_n.$$

If there would be another point of this set, we would have $y^* \in \cap B_n$; then the distance $\rho(x^*,y^*)$ between x^* and y^*, by the triangle inequality, is dominated by an infinitesimal sequence:

$$\rho(x^*,y^*) \leqslant \rho(x^*,x_n)+\rho(x_n,y^*) \leqslant 2r_n \to 0,$$

so $x^* = y^*$.

$\Leftarrow$. Let any system of nested closed balls have a unique common point. Prove that our space is complete.

Let $\{x_n\}_{n=1}^{\infty}$ be an arbitrary Cauchy sequence. By the definition of the Cauchy sequence,

$$\exists n_1 \in \mathbb{N} : \ \forall n \geqslant n_1 \ \rho(x_n,x_{n_1}) < 1/2.$$

Take the first ball $B_1 := B[x_{n_1},1]$ (twice as large as in the line above). Then, by induction,

$$\exists n_2 > n_1 : \ \forall n \geqslant n_2 \ \rho(x_n,x_{n_2}) < 1/4.$$

Take the next ball $B_2 := B[x_{n_2},\frac{1}{2}]$. It can be easily verified that $B_2 \subset B_1$: let $y \in B_2$; let us find $\rho(x_{n_1},y)$. $\rho(x_{n_1},y) \leqslant \rho(x_{n_2},y)+\rho(x_{n_2},x_{n_1}) < \frac{1}{2}+12 < 1$, so $y \in B_1$. Then we construct by induction

$$B_1 \supset B_2 \supset \dots \supset B_m, \quad B_k = B[x_{n_k},\frac{1}{2^{k-1}}],$$

$k = 1,\dots,m$,

$$\exists n_{m+1} > n_m : \ \forall n \geqslant n_{m+1} \ \rho(x_{n_{m+1}},x_{n_m}) < 1/2^m,$$

and take $B_{m+1} := B[x_{n_{m+1}},\frac{1}{2^{m-1}}] \subset B_m$. Thus, $\{x_{n_m}\}$ is a Cauchy sequence, and

$$\exists \underset{m\to\infty}{x}{}_{n_m} = x^*.$$

But x_{n_m} is a subsequence of x_n. Even though,

$$\rho(x^*,x_n) \leqslant \rho(x^*,x_{n_m})+\rho(x_{n_m},x_n),$$

and each of these terms approaches zero (for the second one, it is due to the fact that we have a Cauchy sequence) as $n,m\to\infty$, therefore, $x^* = \lim_{n\to\infty} x_n$. □

Lecture 2. Metric Spaces. Normed Spaces. Seminorms and Polynormed Spaces. Banach Spaces.

Wrapping Up the Previous Lecture: Properties of Complete Metric Spaces

In the previous lecture, we have completed the proof of the theorem, which provides a criterion for completeness in terms of systems of nested closed balls. Now, we are to give some examples.

Example 2.1. *Let (X,ρ) be an incomplete metric space. We have a system of nested closed balls $\{B_n\}$, so that their radii r_n approaching zero, and $\cap_{n=1}^{\infty} B_n = \varnothing$. This example can be represented by $\mathbb{R}\backslash\{0\}$ and the balls with centers at $1/n$ and the same radii: $B_n := B[\frac{1}{n},\frac{1}{n}] = (0,\frac{2}{n}]$. These balls are closed in that space (according to the definition of the closed subset), and their intersection is empty.*

The following theorem is the last one in the section devoted to the general properties of complete metric spaces.

Definition 2.1. *A subset $M \subset (X,\rho)$ is called* ***nowhere dense*** *if $\forall B$ (ball in X) $\exists \tilde{B} \subset B$ (another ball): $M \cap \tilde{B} = \varnothing$.*

This definition is equivalent to **interior of** $\overline{M} = \varnothing$.

Theorem 2.1 (Baire category theorem)**.** *Let (X,ρ) be a complete metric space, and X be represented as a countable union of subsets $X = \cup_{n=1}^{\infty} X_n$. Then $\exists n_0$: $\overline{X_{n_0}}$ has interior points.*

This means that all X_n cannot be nowhere dense all at once.

According to Baire, X is a set of *I* category if there is a representation of X as a countable union $X = \cup_{n=1}^{\infty} X_n$ of nowhere dense sets X_n; X is a set of *II* category otherwise.

So, if (X,ρ) is complete metric space, then it belongs to the *II* category.

Proof by contradiction. Let (X,ρ) be complete, and suppose that there is a representation of X as a countable union $X = \cup_{n=1}^{\infty} X_n$ of nowhere dense sets X_n. Then, by definition of nowhere dense set, there exists a ball $B_1 = B[x_1,r_1]$, $r_1 < 1$: $B_1 \cap X_1 = \varnothing$.

Then we take the nowhere dense X_2; there exists $B_2 = B[x_2,r_2] \subset B_1$: $B_2 \cap X_2 = \varnothing$, and $r_2 < 1/2$.

If we construct nested balls $B_1 \supset B_2 \supset \cdots \supset B_n$, $B_k = B[x_k,r_k]$, $r_k < 1/2^{k-1}$ in such a manner, then

$$X_k \cap B_k = \varnothing,$$

and since X_{n+1} is nowhere dense, there exists a ball $B_{n+1} = B[x_{n+1}, r_{n+1}]$, $r_{n+1} < 1/2^n$, such that $X_{n+1} \cap B_{n+1} = \varnothing$.

We obtain a system of nested closed balls $\{B_n\}_{n=1}^{\infty}$, so that $r_n \to 0$ and $B_n \cap X_n = \varnothing$. According to the theorem from the previous lecture,

$$\exists! x^* = \cap_{n=1}^{\infty} B_n.$$

Thus, $x^* \notin \cup X_n = X$, so we arrive at the contradiction. □

Let us give some remarks concerning this theorem. First of all, the Baire category theorem tells us something only about complete metric spaces (that they belong to the second category); incomplete metric spaces can belong to either of the categories. Consider some examples:

Example 2.2. *Let (X, ρ) be an incomplete metric space. For which X can we find a representation in the form of a countable union $X = \cup_{n=1}^{\infty} X_n$ of nowhere dense sets X_n? For instance, $X = \mathbb{Q} = \cup_{r_n \in \mathbb{Q}} \{r_n\}$: each point $r_n \in \mathbb{Q}$ is nowhere dense in $\mathbb{Q}$.*

Example 2.3. *Let (X, ρ) be an incomplete metric space. For which X there is no representation in the form of a countable union $X = \cup_{n=1}^{\infty} X_n$ of nowhere dense sets X_n? The simplest example is $\mathbb{R} \setminus \{0\}$ (this is an incomplete metric space, but there is no such a representation, since otherwise we would prove that $\mathbb{R}$ is countable).*

Example 2.4. *$\exists$ countable dense in $\mathbb{R}$, countable nowhere dense in $\mathbb{R}$, and uncountable nowhere dense in $\mathbb{R}$: $\mathbb{Q}$, $\mathbb{N}$, and the Cantor set respectively.*

What can we do if our metric space is incomplete?

Definition 2.2. *(Y, d) is called a **completion** of a metric space (X, ρ) if*

1) *(Y, d) is a complete metric space,*

2) *$\exists Y_0 \subset Y : Y_0 \cong X$ (full isometry),*

3) *$\overline{Y_0} = Y$.*

Theorem 2.2 (without a proof for now)**.** *For any metric space (X, ρ), there exists a unique (up to isometry) completion.*

Normed Spaces

Definition 2.3. *Let X be a linear space over a field $\mathbb{K}$ ($\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$). A function $\|\cdot\| : X \to [0, \infty)$, $x \mapsto \|x\|$, is called a **norm** if it satisfies the following conditions:*

1) $\|x\| = 0 \;\Leftrightarrow\; x = 0$,

2) $\forall \alpha \in \mathbb{K}\ \forall x \in X$*:* $\|\alpha x\| = |\alpha| \cdot \|x\|$,

3) $\forall x, y \in X$*:* $\|x + y\| \leqslant \|x\| + \|y\|$ *(the triangle inequality).*

A set X *endowed with a norm* $\|\cdot\|$ *is called a* ***normed space****.* ***Convergence*** *in the normed space is naturally defined by*

$$x_n \to x \quad \textit{if} \quad \|x_n - x\| \to 0 \quad \textit{as} \quad n \to \infty.$$

Any normed space $(X, \|\cdot\|)$ is obviously a metric space (X, ρ) with metric $\rho(x, y) = \|x - y\|$, so the convergence here means exactly the same as the convergence with respect to the norm.

All the examples of metric spaces from the previous lecture, except for the space with discrete metric, are normed spaces. Discrete metric cannot be defined by a norm since this metric is not linear.

Question: Is every linear space X with a shift-invariant metric ρ (i.e. $\rho(x + z, y + z) = \rho(x, y)$) a normed space? (This property obviously holds for the metric defined by $\rho(x, y) = \|x - y\|$.)

The **answer** is no!

We can construct a metric space, metric of which cannot be defined by a norm. Consider a space of all sequences:

$$s \ni x = (x_1, x_2, \dots);$$

it has linear structure:

$$\alpha \cdot x = (\alpha x_1, \alpha x_2, \dots), \quad \alpha \in \mathbb{K},$$

and

$$x + y = (x_1 + y_1, x_2 + y_2, \dots,).$$

What about the convergence in this space? It is natural to define a point-wise convergence:

$$x^n \equiv (x_1^n, x_2^n, \dots, x_k^n, \dots) \to x \equiv (x_1, x_2, \dots, x_k, \dots)$$

if $\forall k$: $x_k^n \to x_k$ as $n \to \infty$.

Statement 2.1. *1) There exists a metric* ρ *such that* $\rho(x^n, x) \to 0 \Leftrightarrow x_k^n \to x_k\ \forall k$,

2) there is no norm $\|\cdot\|$ *that defines convergence in* s*.*

Hint: if ρ is a metric, then $\rho' = \frac{\rho}{1+\rho}$ is also a metric, and it defines the same convergence. Moreover, this metric is bounded from above by 1. Proof of these facts is an optional exercise.

Proof of 1). Consider a metric

$$\rho(x^n, x) = \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{|x_k^n - x_k|}{1 + |x_k^n - x_k|}$$

(it is obviously a metric, according to the exercise above). We claim that convergence with respect to this metric is equivalent to the coordinate-wise convergence:

$$\rho(x^n, x) \Leftrightarrow x_k^n \to x_k \ \forall k.$$

To prove it in $\Leftarrow$ direction, we note that the sum converges uniformly with respect to n: the sum can be dominated by $\sum_{k=1}^{\infty} \frac{1}{2^k} < \infty$. Thus, one can take a limit with respect to n under the sum sign:

$$\lim_{n\to\infty} \sum = \sum \lim_{n\to\infty}$$

due to the uniform convergence, as we remember from Calculus. Now recall that assumption here is that we have a coordinate-wise convergence; then, $(x_k^n \to x_k) \Rightarrow \rho(x^n, x) \to 0$.

Proof in $\Rightarrow$ direction can be completed by contradiction: let $\rho(x^n, x) \to 0$ and $\exists k_0$: $\exists n_j \to \infty$ $\exists c > 0$: $|x_{k_0}^{n_j} - x_{k_0}| \geqslant c$. Note that the function $f(t) = t/1+t$ is a strictly monotonic function, thus

$$\rho(x^n, x) \geqslant \frac{1}{2^{k_0}} \frac{|x_{k_0}^{n_j} - x_{k_0}|}{1 + |x_{k_0}^{n_j} - x_{k_0}|} \geqslant \frac{1}{2^{k_0}} \frac{c}{1+c} \not\to 0,$$

which gives us a contradiction. □

Proof of 2) can also be completed by contradiction. Let $\exists \|\cdot\|$. Consider

$$x^n = (0, \dots, 0, 1, 0, \dots),$$

where 1 appears at the n-th position. The norms of these elements are some nonzero numbers (since $x^n \neq 0$, see the definition of the norm): $\|x^n\| = \alpha_n$. Now we consider a sequence

$$y^n = \frac{x^n}{\alpha_n}, \quad \|y^n\| = 1.$$

What can we say about the distance between y^n and 0, i.e. $\rho(y^n, 0)$?

$$\rho(y^n, 0) = \frac{1}{2^n} \frac{1/\alpha_n}{1 + 1/\alpha_n} \to 0 \quad \text{as} \quad n \to \infty,$$

so y^n converges to 0 with respect to the metric from the point 1) above; it is equivalent to the coordinate-wise convergence. In other words, we constructed a sequence converging to 0 with a norm equal to 1, which means that this sequence does not converge with respect to the norm, so we have a contradiction. □

Seminorms and Polynormed Spaces

In further, we are going to refrain from the topology of the spaces we study, so that this course would not become a topological functional analysis; our aim is to study operators. Even though, let us now consider a little topological side note.

Definition 2.4. *Let X be a linear space over a field $\mathbb{K}$, $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$. A function $p : X \to \mathbb{R}$ is called a **seminorm** if*

1) $\forall x \in X$: $p(x) \geqslant 0$,

2) $\forall \alpha \in \mathbb{K}$, $\forall x \in X$: $p(\alpha x) = |\alpha| \cdot p(x)$,

3) $\forall x, y \in X$: $p(x+y) \leqslant p(x) + p(y)$.

The difference between norms and seminorms is that the latter can be equal to zero even for nonzero elements of our space: $x \neq 0$ and $p(x) = 0$.

Example 2.5 (of seminorms that are not norms). *1) For sequences $x = (x_1, \dots, x_n, \dots)$: $p_k(x) = |x_k|$.*

2) For $\mathbb{R}^3$: $p(x) = \sqrt{x_1^2 + x_2^2}$.

3) For $C[a,b]$: $p_x(f) = |f(x)|$ (evaluation of the value of f at a certain point $x \in [a,b]$).

Definition 2.5. *X is called a **polynormed space** (or a **locally convex space**) if there is a set of seminorms defined on X: $\{p_\alpha\}_{\alpha \in \Lambda}$ (Λ can be uncountable), and convergence is defined by*

$$x_n \to x \quad if \quad \forall \alpha \in \Lambda: \ p_\alpha(x_n - x) \to 0,$$

and the set of seminorms distinguish the points, i.e.

$$\forall x \neq y \ \exists \alpha: \ p_\alpha(x-y) \neq 0.$$

The latter assumption is required for the topology to be Hausdorff (otherwise, the limit may be nonunique).

The **base** for the topology of the polynormed space is so-called “standard” balls $U_{\varepsilon,\alpha_1,\dots,\alpha_n}(x_0) = \{y \in X : \ \forall i = 1,\dots,n : \ p_{\alpha_i}(x_0 - y) < \varepsilon\}$, i.e. this is an intersection of the balls of the **prebase**:

$$U_{\varepsilon,\alpha_1,\dots,\alpha_n}(x_0) = \cap_{i=1}^{n} U_{\varepsilon,\alpha_i}(x_0).$$

Now we can consider the following constructions:

1) Let $(X, \{p_\alpha\}_{\alpha=1}^{n})$ be a polynormed space with a finite number of seminorms. We claim that this space is a normed space: $(X, \|\cdot\|)$; for instance, we can choose

$$\|x\| = \sum_{k=1}^{n} p_k(x) \quad \text{or} \quad \|x\| = \max_{1\leqslant k\leqslant n} p_k(x).$$

2) $(X, \{p_k\}_{k=1}^{\infty})$. This space is a metric space (X, ρ), where the metric can be defined, for example, in the following way:

$$\rho(x,y) = \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{p_k(x-y)}{1 + p_k(x-y)}.$$

Banach Spaces

Earlier, we considered complete metric spaces. Any normed space is a metric space. A natural question arises about the restriction of the concept of completeness to normed spaces.

Definition 2.6. *A complete normed space $(X, \|\cdot\|)$ is called a **Banach space**.*

The following spaces considered in the first lecture are Banach spaces: $\mathbb{R}^n$, $\mathbb{C}^n$, c_0, c, $\ell_p(n)$, ℓ_p, $C[a,b]$, $C^n[a,b]$, $L_p(\Omega,\mu)$, $W_p^n[a,b]$ (and, in fact, all Sobolev spaces).

Lemma 2.1. *Let (X,ρ) be a complete metric space, and $M \subset X$. Then*

$$(M,\rho) \text{ is complete} \Leftrightarrow M \text{ is closed.}$$

Proof. $\Leftarrow$. If $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence in (M,ρ), then it is Cauchy in (X,ρ), and $\exists x := \lim x_n$; x is a limit point, so $x \in M$, therefore, M is complete.

$\Rightarrow$. Let x be a limit point of M; then there exists a sequence $x_n \in M$ such that

$$x = \lim_{n\to\infty} x_n.$$

$\{x_n\}$ is Cauchy, thus, $x \in M$; therefore, M is closed. □

Theorem 2.3. *For any metric space (X,ρ), there exists a completion, and it is unique up to isometry.*

Proof.

1) Consider a space $B(X)$ of bounded functions on X with norm

$$\|f\| := \sup_{x\in X}|f(x)|. \quad (f : X \to \mathbb{R}.)$$

2) $B(X)$ is complete (i.e., it is a Banach space): Let $\{f_n\}_{n=1}^{\infty}$ be a Cauchy sequence; then

$$\forall \varepsilon > 0 \ \exists N = N(\varepsilon) : \ \forall n, m \geqslant N :$$

$$\sup_x |f_n(x) - f_m(x)| < \varepsilon. \tag{2.1}$$

Then we immediately obtain that the sequence of values is Cauchy (at any x):

$$\forall x \in X : |f_n(x) - f_m(x)| < \varepsilon \ \Rightarrow \ \{f_n(x)\}_{n=1}^{\infty} \text{ is Cauchy,}$$

therefore,

$$\forall x \ \exists \lim_{n\to\infty} f_n(x) =: f(x) \quad \text{(a pointwise limit).}$$

Then we have to demonstrate that this function is bounded, and we must show that this is a limit in the supremum sense. In order to do so, we use (2.1). This inequality is uniform with respect to $x \in X$ and $n, m \geqslant N$. Take

$$\lim_{m\to\infty} \cdots =: \sup_x |f_n(x) - f(x)| \leqslant \varepsilon$$

(we can take it under the supremum due to the uniformity). Thus, f is the limit function in $B(X)$. Then

$$\|f\| \leqslant \|f - f_n\| + \|f_n\| \leqslant \varepsilon + \|f_n\|;$$

the second term is finite, therefore, f is bounded.

3) Construct an isometric embedding $X \hookrightarrow B(X)$. For any $x \in X$, we put in correspondence a bouded function $f_x(t) = \rho(x,t) - \rho(x_0,t)$, where x_0 is some fixed point. For $y \in X$, it is $f_y(t) = \rho(y,t) - \rho(x_0,t)$.

   a) f_x is bounded:

   $$|f_x(t)| = |\rho(x,t) - \rho(x_0,t)| \leqslant |\rho(x,x_0) + \rho(x_0,t) - \rho(x_0,t)| \leqslant \rho(x,x_0) \ (\forall t).$$

   b) f_x is an isometry:

   $$\|f_x - f_y\| = \sup_t |\rho(x,t) - \rho(y,t)| \leqslant \rho(x,y),$$

   and for $t = x$ or $t = y$, we have an equality.

Let the image of X in $B(X)$ under the embedding described be denoted by Y_0. Take a closure: $Y = \overline{Y_0}$. It is a closed subset of $B(X)$, thus, according to the lemma above, Y is complete. Therefore, Y is a completion of X. The uniqueness will be discussed in the next lecture.

Self-Study Exercises

The following exercises are for self-study.

Exercise 2.1. *1) Prove that* c_0 *is complete.*

2) Prove that $B[a,b]$ *(bounded functions on* $[a,b]$*) with norm* $\|f\| = \sup\limits_{x\in[a,b]} |f(x)|$ *is not separable.*

3) Using the fixed-point theorem, find the limit of the sequence

$$2,\ 2+\frac{1}{2},\ 2+\frac{1}{2+\frac{1}{2}},\ldots.$$

4) Give an example of a complete metric space (X,ρ) *with system of closed nested balls* $B_n = B[x_n,r_n]$ *such that* $r_n \to r > 0$ *and* $\cap_{n=1}^{\infty} B_n = \varnothing$*.*

5) 2-adic metric: let $x,y\in\mathbb{Q}$*,* $x\neq y$*. Then there exists a representation* $x-y=\frac{1}{2^n}\frac{a}{b}$*,* $n\in\mathbb{Z}$*,* a *and* b *are odd. Prove the following:*

a)

$$\rho(x,y) = \begin{cases} \dfrac{1}{2^n}, & x\neq y, \\ 0, & x=y \end{cases}$$

is a metric, and $\rho(x,y)\leqslant \max(\rho(x,z),\rho(z,y))$*;*

b) if $(B_1 = B(x_1,r_1))\cap(B_2=B(x_2,r_2))\neq\varnothing$*, then either* $B_1\subset B_2$*, or* $B_2\subset B_1$*;*

c) let $a,b,c\in\mathbb{Q}$*, then at least two of*

$$\rho(a,b),\quad \rho(b,c),\quad \rho(a,c)$$

coincide.

Lecture 3. Euclidean and Hilbert Spaces.

Proof of Uniqueness of the Completion

In the previous lecture, we proved only the existence of the completion of the metric space. Now, we prove the uniqueness.

Let (X,ρ) be a metric space and (Y,d), (Z,w) be two completions. By definition of completeness,

$$\exists Y_0 \subset Y \quad \text{and} \quad Z_0 \subset Z: \quad Y_0 \cong X \cong Z_0, \quad \overline{Y_0} = Y, \quad \overline{Z_0} = Z.$$

Then there exists a bijection $\varphi : Y_0 \to Z_0$. So we can just extend it to the limit points. If y is a limit point of Y, and $y \notin Y_0$, then

$$\exists \{y_{n,0}\}_{n=1}^{\infty} \in Y_0 \quad \text{s.t.} \quad y_{n,0} \to y.$$

We have the bijection φ between our spaces, so we map into a sequence $z_{n,0} := \varphi(y_{n,0})$. Since φ is isometry,

$$w(z_{n,0}, z_{m,0}) = d(y_{n,0}, y_{m,0}) \to 0 \quad \text{as} \quad n,m \to \infty,$$

therefore, $\{z_{n,0}\}_{n=1}^{\infty}$ is Cauchy, and we define

$$\varphi(y) = \lim_{n\to\infty} \varphi(z_{n,0}) = z.$$

This construction is well-defined: consider another sequence $\{y'_{n,0}\}_{n=1}^{\infty}$, $y'_{n,0} \to y$, and combine both of them

$$y_{1,0}, y'_{1,0}, \dots, y_{n,0}, y'_{n,0} \dots \to y,$$

therefore,

$$\varphi(y_{1,0}), \varphi(y'_{1,0}), \varphi(\dots, y_{n,0}), \varphi(y'_{n,0}), \dots$$

converges, so the construction of z is well-defined. □

Note that for the normed spaces this construction based on the embedding into the bounded functions does not preserve the linear structure. Even though, for normed spaces, there always exists a completion preserving the linear structure.

Why Banach Spaces Are not Good Enough

Recall that we call a complete normed space $(X, \|\cdot\|)$ a Banach space. Sometimes the property of being complete is not sufficient for further constructions and applications. There are two historical questions:

1) The existence of the closed complement.

   Let X be a Banach space, and X_0 be a closed subspace $X_0 \subset X$; one can easily prove that as it is closed, the space $(X_0, \|\cdot\|)$ is Banach itself.

   Question: Is there a closed subspace X_1 such that

$$X = X_0 \oplus X_1?$$

   The common answer is, unfortunately, **no**. Example can be provided by $c_0 \subset \ell_\infty$, which is closed, but does not have a closed complement.

2) Approximation. More precisely, existence of a basis.

   For infinite-dimensional spaces, there are two commonly used different definitions of a basis:

   Definition 3.1 (of algebraic (or Hamel) basis)**.** *Let X be a linear space,* $\dim X = \infty$. *A system $\{e_\nu\}_{\nu\in\Lambda}$ (Λ may be uncountable) is called a* ***Hamel basis*** *if*

   - *it is linear independent, i.e., any finite subsystem of $\{e_\nu\}_{\nu\in\Lambda}$ is linear independent,*
   - $\forall x \in X$*:* $x = \sum_{k=1}^{n} c_k e_{\nu_k}$.

   There is a theorem valid for any linear space claiming that there exists a Hamel basis; this theorem is not constructive. A rare exception is c_{00}, where the Hamel basis can be explicitly constructed.

   Definition 3.2. *Let X be a separable normed space,* $\dim X = \infty$. $\{e_k\}_{k=1}^{\infty}$ *is called a* ***Schauder basis*** *if*

   - *it is linear independent, i.e., any finite subsystem of $\{e_\nu\}_{\nu\in\Lambda}$ is linear independent,*
   - $\forall x \in X$*:* $\exists!$ *representation*

$$x = \sum_{k=1}^{\infty} c_k e_k, \quad c_k \in \mathbb{K}\ (\mathbb{R}\ \textit{or}\ \mathbb{C})$$

     such that

$$\|x - \sum_{k=1}^{n} c_k e_k\| \to 0 \quad \textit{as} \quad n \to \infty.$$

So we can approximate any vector with a finite sum.

If Schauder basis exists, our space is forced to be separable since we have a countable set of functions e_k, and if we replace the coefficients c_k with $\widetilde{c}_k \in \mathbb{Q}$, we obtain a countable dense subset

$$\left\{ \sum_{k=1}^{n} \widetilde{c}_k e_k,\ \widetilde{c}_k \in \mathbb{Q} \right\}.$$

Question: Is it true that for any separable normed space there exists a Schauder basis?

The answer is **no** again.

First example was given in 1972 by Enflo; he constructed an example of separable Banach space without a Schauder basis.

Euclidean and Hilbert Spaces

For Hilbert spaces, one can construct both the closed complement and the basis. These spaces are also commonly used in applications, i.e., in Quantum Mechanics.

Definition 3.3. *Let H be a linear space over a field $\mathbb{K}$ ($\mathbb{R}$ or $\mathbb{C}$). A function $(\cdot,\cdot) : H \times H \to \mathbb{K}$ is called a* ***inner product*** *if*

1) $\forall x \in H\colon (x,x) \geqslant 0$ *and* $(x,x) = 0 \Leftrightarrow x = 0$,

2) $\forall \alpha, \beta \in \mathbb{K},\ \forall x,y,z \in H\colon (\alpha x + \beta y, z) = \alpha(x,z) + \beta(y,z)$,

3) $\forall x,y \in H\colon (x,y) = \overline{(y,x)}$.

The space $\big(H, (\cdot,\cdot)\big)$ is called a ***Euclidean space****, furthermore, if H is complete w.r.t. the Euclidean norm $\|x\| = \sqrt{(x,x)}$, then H is called a* ***Hilbert space****.*

Properties of Inner Product

1) Cauchy–Bunyakovsky (Cauchy–Schwarz) inequality:

$$\forall x,y \in H: \quad |(x,y)| \leqslant \sqrt{(x,x)} \cdot \sqrt{(y,y)}.$$

2) $\sqrt{(x,x)}$ is the Euclidean norm in H: $\|x\| = \sqrt{(x,x)}$, so

$$|(x,y)| \leqslant \|x\| \cdot \|y\|.$$

3) $x \perp y$ if $(x,y) = 0$.

   Then we can define an **orthogonal complement** to $M \subset H$ by $M^{\perp} = \{y \in H : \ \forall x \in M \ (x,y) = 0\}$.

   In real spaces, we can also define an angle between vectors.

   There is a simple statement:

   Statement 3.1. *$M^{\perp}$ is a closed linear subspace.*

   It follows from the linearity of the inner product and the fact that $(\cdot,\cdot)$ is a continuous function (by Cauchy–Bunyakovsky inequality).

4) The Pythagorean Theorem. If $x \perp y$, then

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

5) The Parallelogram law (identity):

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Example 3.1. *Show that $C[0,1]$ with norm $\|f\| = \max\limits_{x\in[0,1]} |f(x)|$ is not Euclidean. How do we show it? We can simply prove that for this kind of norm, the parallelogram law does not hold. So we have to find a pair of functions for which it is not valid. Take, for example,*

$$f(x) \equiv 1, \quad g(x) = x.$$

Then

$$\|f\| = 1, \quad \|g\| = 1, \quad \|f+g\| = 2, \quad \|f-g\| = 1,$$

so, according to the parallelogram law, $4+1 = 2+2$, which is incorrect.

Theorem 3.1. *Let H be a Hilbert space, and $H_0 \subset H$ be a nontrivial closed subspace ($H_0 \neq \{0\}$, $H_0 \neq H$). Suppose $x \notin H_0$. Then*

$$\exists! x_0 \in H_0 : \quad \|x - x_0\| = \operatorname{dist}(x, H_0), \quad \text{and} \quad x - x_0 \perp H_0.$$

Definition 3.4. *Let (X,ρ) be a metric space, $M \subset X$, and $x \in X$. Then*

$$\operatorname{dist}(x,M) = \inf_{y\in M} \rho(x,y).$$

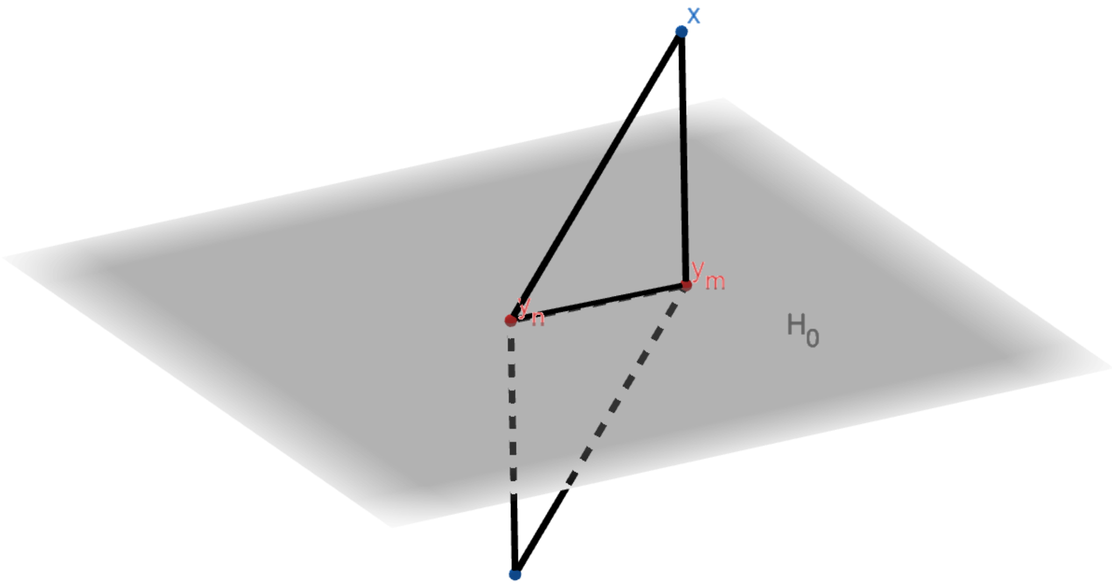

Рис. 3.1. x, y_n, y_m, and the parallelogram

Proof. $x \notin H_0$, H_0 is closed $\Rightarrow \operatorname{dist}(x, H_0) =: d > 0$ (or else x is forced to be a limit point of H_0). By definition of **inf**,

$$\exists \{y_n\}_{n=1}^{\infty} : \quad y_n \in H_0, \ \|x - y_n\| \to d,$$

so,

$$\forall \varepsilon > 0 \ \exists N = N(\varepsilon) \text{ s.t. } \forall n > N : \ d \leqslant \|x - y_n\| < d + \varepsilon.$$

Take $n, m \geqslant N$, and consider the geometric interpretation (see figure 3.1 below).

Write down the parallelogram law:

$$\|y_n - y_m\|^2 + \|2x - y_n - y_m\|^2 = 2\|x - y_n\|^2 + 2\|x - y_m\|^2.$$

Then we rewrite it as

$$\|y_n - y_m\|^2 = 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\|x - \frac{y_n + y_m}{2}\|^2,$$

so

$$\|y_n - y_m\|^2 < 2(d + \varepsilon)^2 + 2(d + \varepsilon)^2 - 4d^2 = 8d\varepsilon + 4\varepsilon^2 \to 0 \text{ as } \varepsilon \to 0;$$

this means that $\{y_n\}_{n=1}^{\infty}$ is Cauchy, and, therefore, there exists a limit, which we denote by x_0.

The sequence $\{y_n\}_{n=1}^{\infty}$ such that $\|x - y_n\| \to d$ is not unique. However, if we take another sequence, $\{z_n\}_{n=1}^{\infty} \subset H_0$, so that $\|x - z_n\| \to d$, and write down the parallelogram law for y_n

and z_n, then

$$\|y_n - z_n\|^2 = 2\|x - y_n\|^2 + 2\|x - z_n\|^2 - 4\|x - \frac{y_n + z_n}{2}\|^2,$$

so the same bound holds:

$$\|y_n - z_n\|^2 < 8d\varepsilon + 4\varepsilon^2,$$

therefore, the limit is unique: $\lim y_n = \lim z_n$.

Why is $x - x_0$ orthogonal to H_0? Consider a vector

$$x(t) = x - x_0 + tz$$

for an arbitrary $z \in H_0$ and $t \in \mathbb{R}$, and a function

$$f(t) = \|x - x_0 + tz\|^2.$$

We know that $t = 0$ is a minimum of $f(t)$. Rewrite the formula for $f(t)$:

$$f(t) = (x - x_0 + tz, x - x_0 + tz) = \|x - x_0\|^2 + 2\,\mathrm{Re}(x - x_0, z)t + \|z\|^2 t^2.$$

Since $t = 0$ is the minimum,

$$f'(t)\big|_{t=0} = 0 \quad \Rightarrow \quad \mathrm{Re}(x - x_0, z) = 0;$$

in real space, it means that $x - x_0 \perp z$ $(\forall z \in H_0)$. In complex space, we can replace z with iz, and then we obtain $\mathrm{Im}(x - x_0, z) = 0$. Therefore, $x - x_0 \perp z$. □

Corollary 3.1. *Let H be a Hilbert space and $H_0 \subset H$ be a closed nontrivial subspace. Then there exists a closed subspace H_1 such that $H = H_0 \oplus H_1$ $(H_1 := H_0^\perp)$.*

Proof. If $x \in H_0$, then $x = x + 0$, where $x \in H_0$ and $0 \in H_0^\perp$. If $x \notin H_0$, due to the theorem above,

$$\exists x_0 \in H_0 : \|x - x_0\| = \mathrm{dist}(x, H_0),$$

and $x - x_0 \perp H_0$. So we take $x_1 := x - x_0$, and $x = x_0 + x_1$, where $x_0 \in H_0$ and $x_1 \in H_1$; this is an orthogonal sum, and, therefore, it is a direct sum. □

Orthogonal Systems in Euclidean and Hilbert Spaces

We will consider only separable Euclidean spaces H, $\dim H = \infty$.

Definition 3.5. *A system $\{e_n\}_{n=1}^\infty$ is **orthonormal** (ONS) if $(e_i, e_j) = \delta_{ij}$.*

Given an orthonormal system, for any $x \in H$, we can obtain $x_n := (x, e_n)$ (the **Fourier coefficients**); the series $\sum_{k=1}^\infty (x, e_n)e_n$ is called a **Fourier series**.

Theorem 3.2 (Bessel inequality)**.** *Let H be a separable Euclidean space,* $\dim H = \infty$*, and* $\{e_n\}_{n=1}^{\infty}$ *be an ONS in H. Then for any $x \in H$:*

$$\sum_{k=1}^{\infty} |x_n|^2 \leqslant \|x\|^2.$$

To prove this, we begin with

Lemma 3.1. *Define*

$$x^n = \sum_{k=1}^{n} x_k e_k.$$

Then $x - x^n \perp x^n$.

Proof of the Lemma. Write down the inner product:

$$(x - x^n, x^n) = (x - \sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} x_j e_j) = \sum_{j=1}^{n} \overline{x_j}(x, e_j) - \sum_{i,j=1}^{n} x_i \overline{x_j}(e_i, e_j),$$

where $(e_i, e_j) = \delta_{ij}$, so

$$(x - x^n, x^n) = \sum_{j=1}^{n} |x_j|^2 - \sum_{j=1}^{n} |x_j|^2. \qquad \square$$

Proof of the Theorem. For

$$\|x\|^2 = \|x - x^n + x^n\|^2,$$

we use the Pythagorean theorem:

$$\|x\|^2 = \|x - x^n\|^2 + \|x_n\|^2 \geqslant \|x^n\|^2 = \sum_{j=1}^{n} |x_j|^2$$

for any positive integer n. Then, we take a limit

$$\lim_{n\to\infty} : \quad \|x\|^2 \geqslant \sum_{j=1}^{\infty} |x_j|^2. \quad \square$$

Remark 3.1. *The Bessel inequality implies that $\{x_j\}_{j=1}^{i} nfty \in \ell_2$.*

Theorem 3.3 (Riesz, Fisher)**.** *Let H be a Hilbert space, $\{e_n\}_{n=1}^{\infty}$ be an ONS in H, and $\{x_n\}_{n=1}^{\infty} \in \ell_2$. Then there exists $x \in H$: $x_k = (x, e_k)$.*

Proof. Consider the partial sum

$$x^n = \sum_{k=1}^{n} x_k e_k.$$

Let $n > m$:

$$\|x^n - x^m\| = \sum_{j=m+1}^{n} |x_j|^2 \to 0 \text{ as } n, m \to \infty,$$

therefore, $\{x_n\}_{n=1}^{\infty}$ is Cauchy, and so there is a limit $x^n \to x$. It is clear that $x_k = (x, e_k)$. □

Now we have to introduce some additional notions.

Definition 3.6. *Let $\big(X, \|\cdot\|\big)$ be a normed space. A system $\{e_k\}_{k=1}^{\infty}$ is called* ***closed*** *if the closure of its linear span is X: $\overline{\langle\{e_k\}_{k=1}^{\infty}\rangle} = X$.*

By default, if we say *basis*, we mean a Schauder basis.

Remark 3.2. *What is the difference between a closed ONS and a basis? If $\{e_k\}_{k=1}^{\infty}$ is a basis, then $\{e_k\}_{k=1}^{\infty}$ is closed, since, by definition of basis,*

$$\|x - \sum_{k=1}^{n} c_k e_k\| \to 0 \text{ as } n \to \infty.$$

The converse is false, see an example below.

Example 3.2 (The Weierstrass approximation theorem)**.** $e_k(x) = x^k$, $k \in \mathbb{N} \cup \{0\}$, *in $C[0,1]$. According to the Weierstrass approximation theorem, this system is closed; but this is not a basis.*

For basis, we have a priori representation

$$x = \sum_{k=1}^{\infty} c_k e_k.$$

So if $\|x - \sum_{k=1}^{n} c_k e_k\| < \varepsilon$ and we want to increase the accuracy, say, make it $\|x - \sum_{k=1}^{n_1} c_k e_k\| < \varepsilon/2$, we just have to take $n_1 > n$; this is not true for the closed systems: we have no representation for x as a sum.

Definition 3.7. *Let H be a Euclidean space. A system $\{e_k\}_{k=1}^{\infty}$ is called* ***complete*** *if*

$$\forall x \in H: \ \Big((x, e_k) = 0 \ \forall k\Big) \Rightarrow (x = 0).$$

Theorem 3.4. *Let H, $\dim H = \infty$, be a separable Hilbert space and $\{e_k\}_{k=1}^{\infty}$ be an ONS in H. Then the following statements are equivalent:*

1) $\{e_k\}_{k=1}^{\infty}$ *is closed,*

2) $\{e_k\}_{k=1}^{\infty}$ *is complete,*

3) $\{e_k\}_{k=1}^{\infty}$ *is basis,*

4) $\forall x \in H$*:* $\sum_{k=1}^{\infty} = \|x\|^2$ *(Parseval's identity).*

Proof. The idea is to show that $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 1$.

$1 \Rightarrow 2$) Assume that $\{e_k\}_{k=1}^{\infty}$ is a closed system, and $x \perp e_k$ $(\forall k)$, $x \neq 0$. By the definition of a closed system, there exist sequences of linear combinations

$$\sum_{k=1}^{n} c_k e_{n_k} \to x$$

(here we vary n, c_k, and n_k):

$$\|x\|^2 = \lim_{n,c_k,n_k} \Big(\sum_{k=1}^{n} c_k e_{n_k}, x \Big) = 0,$$

since under the limit we have $\sum c_k (e_{n,k}, x)$, and this inner product vanishes for any n_k; therefore, $x = 0$.

$2 \Rightarrow 3$) Take x, then take the Fourier coefficients $x_k = (x, e_k)$, and consider the sum

$$\sum_{k=1}^{\infty} x_k e_k.$$

If $\sum_{k=1}^{\infty} x_k e_k \neq x$, we define another vector

$$y := \sum_{k=1}^{\infty} x_k e_k.$$

Consider the inner product

$$(x - y, e_k) = (x, e_k) - (y, e_k) = x_k - x_k = 0,$$

where $(x, e_k) = x_k$ by the definition of x_k, and $(y, e_k) = x_k$ by the construction of y. Thus, due to the completeness of the system, $x - y = 0$, therefore, $x = y$.

$3 \Rightarrow 4$) By definition of the basis, $\forall x$:

$$x = \sum_{k=1}^{\infty} x_k e_k, \text{ and } x = \lim_{n\to\infty} \sum_{k=1}^{n} x_k e_k.$$

Then we obtain that

$$\|x\|^2 = (x, x) = \lim_{n\to\infty} \Big(\sum_{k=1}^{n} x_k e_k, x\Big) = \lim_{n\to\infty} \sum_{k=1}^{n} x_k (e_k, x),$$

where $(e_k, x) = \overline{x_k}$, so

$$\|x\|^2 = \lim_{n\to\infty} \sum_{k=1}^{n} |x_k|^2 = \sum_{k=1}^{\infty} |x_k|^2.$$

All we have to prove by now is $4 \Rightarrow 1$), or, more precisely, $4 \Rightarrow 3 \Rightarrow 1$), where $3 \Rightarrow 1$) follows from definition of the basis and the closed system.

To prove $4 \Rightarrow 3$, take $x \in H$: $x_k := (x, e_k)$. Assume that

$$\sum_{k=1}^{\infty} x_k e_k = y \neq x.$$

(This series converges due to the Bessel inequality.) We know that

$$\|x-y\|^2 = \sum_{k=1}^{\infty} |(x-y, e_k)|^2 = \sum_{k=1}^{\infty} |(x, e_k) - (y, e_k)|^2,$$

where $(x, e_k) = x_k$ by the definition of x_k, and $(y, e_k) = x_k$ by the construction of y, so all the terms cross out, i.e. $x = y$. □

Lecture 4. Separable Hilbert Spaces. Bases in Hilbert Spaces.

Further Development of the Previous Lecture: Existence of an Orthonormal Basis in Separable Hilbert Spaces

We continue discussing Hilbert spaces and their properties.

Theorem 4.1. *Let H be a separable Hilbert space,* $\dim H = \infty$. *Then there exists an orthonormal basis (ONB)* $\{e_k\}_{k=1}^{\infty}$

Proof.

1) Since H is separable, there exists a dense system $\{h_k\}_{k=1}^{\infty}$:

$$\overline{\{h_k\}_{k=1}^{\infty}} = H.$$

This system may be quite excessive. We would like to build a system of linearly independent vectors that would have a dense linear span. So, our next step is the following.

2) Without loss of generality, one can assume that $h_1 \neq 0$; then we take $f_1 := h_1$, and $f_2 = h_k$, $k \geqslant 2$, where k is the first number such that f_1 and f_2 are linearly independent.

If we construct $f_1, f_2, \ldots, f_m$ to be linearly independent, then we can take $f_{m+1} = h_j$, where $j = \min\{i : \ h_i \ni \langle f_1, f_2, \ldots, f_m \rangle\}$ (i.e., we require that f_{m+1} does not belong to the linear span of $f_1, \ldots, f_m$).

Thus, we obtain a system $\{f_j\}_{j=1}^{\infty}$ of linearly independent vectors such that

$$\overline{\langle f_1, f_2, \ldots \rangle} = \overline{\{h_k\}_{k=1}^{\infty}} = H.$$

3) Finally, we use the Gram–Schmidt process to generate an orthogonal (moreover, orthonormal) system from $\{f_j\}_{j=1}^{\infty}$:

$$e_1 = \frac{f_1}{\|f_1\|}, \quad \langle e_1 \rangle = \langle f_1 \rangle \quad \tilde{e}_2 = f_2 - c e_1, \quad c \in \mathbb{K},$$

where $c = (f_2, e_1)$, as follows from the relation $\tilde{e}_2 \perp e_1$ that we desire, and then

$$e_2 = \frac{\tilde{e}_2}{\|\tilde{e}_2\|}, \quad \langle e_1, e_2 \rangle = \langle f_1, f_2 \rangle$$

If we construct an orthonormal system $e_1, e_2, \ldots, e_m$ such that

$$\langle e_1, e_2, \ldots, e_k \rangle = \langle f_1, f_2, \ldots, f_k \rangle \quad (\forall k \leqslant m),$$

then, by induction,

$$\tilde{e}_{m+1} = f_{m+1} - \sum_{j=1}^{m} c_j e_j,$$

where c_j, as before, can be found from $(\tilde{e}_{m+1}, e_j)$, $j = 1, 2, \ldots, m$, i.e., $c_j = (f_{m+1}, e_j)$, and

$$e_{m+1} = \frac{\tilde{e}_{m+1}}{\|\tilde{e}_{m+1}\|}.$$

Following this way, we obtain an ONS $\{e_m\}_{m=1}^{\infty}$ that is closed: $\overline{\{e_m\}_{m=1}^{\infty}} = H$. Then, by the last theorem from the previous lecture, $\{e_m\}_{m=1}^{\infty}$ is an orthonormal basis. □

This is one of two main approaches to find an orthonormal basis – find a closed system and make it orthogonal by Gram–Schmidt process. Later, when the time comes to prove the Hilbert–Schmidt theorem, we will discuss the other important way to obtain such a basis.

Applications to Quantum Mechanics and Isometric Isomorphisms of Separable Hilbert Spaces

In Quantum Mechanics, there are different models for describing the states of systems:

- Heisenberg's model, or so-called matrix model (a.k.a. *matrix mechanics*), where observables are operators (infinite matrices) acting on ℓ_2, and the states are vectors from ℓ_2.
- The Schrödinger model, or the model of wave mechanics. In this model, observables are symmetric operators on $L_2(\mathbb{R}^3)$, and the states are wavefunctions $f \in L_2(\mathbb{R}^3)$.

Physicists argued a lot about whose model was more precise. In fact, both are correct, since there is an isometric isomorphism between ℓ_2 and L_2:

Theorem 4.2. *All infinite-dimensional separable Hilbert spaces over the same field are isometrically isomorphic.*

Proof. Let H_1 and H_2 be infinite-dimensional separable Hilbert spaces over $\mathbb{K}$. Let $\{e_k\}_{k=1}^{\infty}$ and $\{f_k\}_{k=1}^{\infty}$ be ONBs in H_1 and H_2 respectively.

We can construct an isomorphism

$$\varphi : H_1 \to H_2$$

in the following way:

$$\varphi(e_k) = f_k.$$

Then

$$\forall x \in H_1 : x = \sum_{k=1}^{\infty} x_k e_k$$

maps to

$$y = \varphi(x) := \sum_{k=1}^{\infty} x_k f_k.$$

One can easily check that the inner product is preserved by this mapping; indeed, take $x' = \sum_{k=1}^{\infty} x'_k e_k$ and $y' = \varphi(x') := \sum_{k=1}^{\infty} x'_k f_k$, then

$$(x, x')_{H_1} = \sum_{k=1}^{\infty} x_k x'_k = (\varphi(x), \varphi(x'))_{H_2},$$

where the formula in the middle is in fact the inner product of the sequences $\{x_k\}_{k=1}^{\infty}$ and $\{x'_k\}_{k=1}^{\infty}$ in ℓ_2. In other words, the theorem can be reformulated as *all separable infinite-dimensional Hilbert spaces are isometrically isomorphic to ℓ_2.* □

Discussion of Self-Study Problems

Now we will discuss some self-study problems from previous lectures.

Problem no. 2 from Lecture 2: $B[a,b]$ (the space of bounded functions) with norm $\|f\| = \sup\limits_{x \in [a,b]} |f(x)|$ is not separable.

We will use the lemma from the first lecture: if there is an uncountable set $M \subset X$ such that $\exists d > 0\ \forall x, y \in M$: $\rho(x,y) \geqslant d$, then X is not separable, which we reformulate as

$$\exists d > 0\ \|f - g\| \geqslant d\ (\forall f, g \in M).$$

Take the following set: $M = \{f_t(x) = \chi_{[a,t)}(x),\ t \in (a,b]\}$, where $\chi_{[a,t)}$ is the characteristic function of $[a,t)$:

$$\chi_W(x) = \begin{cases} 1, & x \in W, \\ 0, & x \ni W. \end{cases}$$

This set is uncountable: we can parametrize M by the parameter $t \in (a,b]$, and $(a,b]$ is uncountable. One can also see that

$$\|f_{t_1} - f_{t_2}\| = 1, \quad t_1 \neq t_2,$$

so we have found an uncountable set with unit distance between any elements, therefore, by the lemma above, $B[a,b]$ is not separable.

Problem no. 4 from Lecture 2: give an example of (X,ρ), a complete space, with the system of closed nested balls $B_n = B[x_n, r_n]$ such that $r_n \to r > 0$ and $\cap_{n=1}^{\infty} B_n = \varnothing$.

An example is a little tricky. One can take $X = \mathbb{N}$ with metric

$$\rho(m,n) = \begin{cases} 0, & m = n, \\ 1 + \dfrac{1}{m} + \dfrac{1}{n}, & m \neq n. \end{cases}$$

The triangle equality for this metric can be verified in a straightforward way:

$$\rho(m,n) \overset{?}{\leqslant} \rho(m,k) + \rho(k,n), \quad n \neq k \neq m,$$

where the left-hand side is at most $1 + \frac{1}{2} + \frac{1}{3}$ and the right-hand side is at least $2 + \ldots$.

Convergence in this space is similar to one in the discrete metric space, i.e. all converging sequences stabilize:

$$x_n \to x \quad \Rightarrow \quad x_1, x_2, \ldots, x_k, x, x, x, \ldots,$$

so X is complete.

Now we take balls $B_n = B[n, 1 + \frac{2}{n}] = \{m \in \mathbb{N} : \ 1 + \frac{1}{m} + \frac{1}{n} \leqslant 1 + \frac{2}{n}\}$, which is the same as $\frac{1}{m} \leqslant \frac{1}{n}$ since $m \geqslant n$. Thus,

$$B_n = [n, n+1, n+2, \ldots),$$

and, therefore, $\cap_{n=1}^{\infty} B_n = \varnothing$.

Typical Examples of Hilbert Spaces

1) $\mathbb{C}^n$ with inner product $(x,y) = \sum_{i=1}^{n} x_i \overline{y}_i$ is a (finite-dimensional) Hilbert space.

2) ℓ_2, which consists of infinite sequences $x = (x_1, \ldots, x_n, \ldots)$ such that $\sum_{i=1}^{\infty} |x_i|^2 < \infty$, with inner product
$$(x,y) = \sum_{i=1}^{n} x_i \overline{y}_i$$
is a Hilbert space.

3) $L_2(\Omega, \mu)$, the space of square-integrable functions on Ω with respect to the measure μ, with inner product
$$(f,g) = \int_{\Omega} f(x) \overline{g(x)} \, d\mu.$$

4) Sobolev spaces $W_2^n[a,b] = \{f : \forall j = 0,1,\dots,n-1\ f^{(j)} \in AC[a,b],\ f^{(n)} \in L_2[a,b]\}$ with inner product

$$(f,g) = \sum_{j=0}^{n} (f^{(j)}, g^{(j)})_{L_2} \equiv \sum_{j=0}^{n} \int_{[a,b]} f^{(j)}(x)\overline{g^{(j)}(x)}\, d\mu.$$

Exercises

Now we will discuss and solve some problems:

1) Consider ℓ_2, and its subspace $H_n = \{x \in \ell_2 :\ \sum_{j=0}^{n} x_j = 0\}$. What is the distance between $e_1 = (1,0,0,\dots)$ and H_n?

   As H_n is a nontrivial closed subspace, by the theorem from the previous lecture, there exists a unique $x^* \in H_n$

$$\|e_1 - x^*\| = \inf_{y \in H_n} \|e_1 - y\|,$$

   and $e_1 - x^* \perp H_n$.

   Here is a way to find such x^*. Consider $x^* = (x_1, x_2, \dots, x_n, x_{n+1}, \dots)$, and minimize the norm of the difference

$$e_1 - x^* = (1 - x_1, -x_2, \dots, -x_n, \dots).$$

   In H_n, we have information only about the coordinates with indices $< n$, and our aim is to minimize the norm. To do so, we should set all the “tail” coordinates to zero:

$$x_{n+1} = x_{n+2} = \dots = 0,$$

   so $x^* \in \ell_2(n)$, i.e., now we consider $H_n\big|_{\ell_2(n)}$. Now, x^* can be simply found, as it is now required that $e_1 - x^*$ is orthogonal to a finite-dimensional subspace $H_n\big|_{\ell_2(n)}$, $\dim H_n\big|_{\ell_2(n)} = n-1$. Take some basis in this space, for instance,

$$f_1 = (1, -1, 0, \dots, 0),$$
$$f_2 = (1, 0, -1, \dots, 0),$$
$$\dots$$
$$f_{n-1} = (1, 0, \dots, 0, -1).$$

   In $\ell_2(n)$, for $x^* = (x_1, x_2, \dots, x_n)$, from the condition

$$e_1 - x^* \perp f_k, \quad k = 1, 2, \dots, n-1,$$

we obtain the system of equations

$$\begin{aligned} 1 - x_1 + x_2 &= 0, \\ 1 - x_1 + x_3 &= 0, \\ &\dots \\ 1 - x_1 + x_n &= 0, \end{aligned}$$

therefore, $x_2 = x_3 = \dots = x_n = a$, where $1 - x_1 + a = 0$, which gives $x_1 = 1 + a$; the value a can be found from the condition $x^* \in H_n$: for

$$x^* = (1 + a, a, \dots, a),$$

we have $1 + na = 0$, or $a = -\frac{1}{n}$, which gives

$$\operatorname{dist}(e_1, H_n) = \|e_1 - x^*\| = \frac{1}{\sqrt{n}}.$$

Exercises: Typical Examples of Bases in Hilbert Spaces

Let us consider the following exercises.

1) Prove that the system $\{e_k\}_{k=1}^{\infty}$, $e_k = (0, 0, \dots, \underset{k\text{-th place}}{1}, 0, \dots)$, is a basis in c_0 and not a basis in c (recall that c_0 is the space of zero-limit sequences with norm $\|x\| = \max\limits_{k \geqslant 1} |x_k|$, and c is the space of converging sequences with norm $\|x\| = \sup\limits_{k \geqslant 1} |x_k|$).

   It is clear that $\{e_k\}_{k=1}^{\infty}$ is a system of linearly independent vectors. For any $x \in c_0$,

$$x = \sum_{k=1}^{\infty} x_k e_k, \quad \text{and} \quad \left\| x - \sum_{k=1}^{n} x_k e_k \right\| = \max_{k \geqslant n+1} |x_k| \to 0 \text{ as } n \to \infty,$$

   since $x \in c_0$.

   What becomes wrong, if we consider this system in the space of converging sequences? In fact, we cannot represent some elements of this space by the sum $\sum_{k=1}^{\infty} x_k e_k$. For example, take

$$e_0 = (1, 1, 1, \dots, 1, \dots);4$$

   if we set $x_k = 1$, then, for any n,

$$\left\| e_0 - \sum_{k=1}^{n} e_k \right\| = \sup_{k \geqslant 1} |x_k| = 1.$$

Nevertheless, if we add this element to the system, i.e., consider $\{e_k\}_{k=0}^{\infty}$ (from $k=0$ instead of $k=1$), then we obtain a basis in c: take

$$x \in c: \quad \lim_{k\to\infty} x_k = a.$$

Consider $\tilde{x} = x - a\cdot e_0$; it is obvious that this element belongs to c_0, and, therefore,

$$x - a\cdot e_0 = \sum_{k=1}^{\infty}(x_k - a)e_k,$$

or simply

$$x = ae_0 \sum_{k=1}^{\infty}(x_k - a)e_k.$$

2) What is a basis in $L_2[a,b]$? Consider, for simplicity, $L_2[0,1]$, $L_2[0,2\pi]$, or $L_2[-\pi,\pi]$. Classical construction of bases in these spaces is given by either exponential function with complex exponents or by sine and cosine, depending on what functions we consider: complex- or real-valued. In $L_2[0,2\pi]$ or $L_2[-\pi,\pi]$, one can take

$$\frac{1}{\sqrt{2\pi}}e^{inx},\ n\in\mathbb{Z},\ \text{or}\ \frac{1}{\sqrt{2\pi}},\frac{1}{\sqrt{\pi}}\cos nx,\frac{1}{\sqrt{\pi}}\sin nx,\ n\in\mathbb{N}.$$

It can be extended to $L_2[a,b]$:

$$\frac{1}{\sqrt{b-a}}e^{\frac{2\pi inx}{b-a}},\quad n\in\mathbb{Z};$$

for $L_2[0,1]$, the normalizing factor is simply equal to 1.

For real-valued functions on a half-interval, i.e., $L_2[0,\pi]$, one can take only sines or cosines (with a constant included, for $n=0$) as a basis, since these functions can be extended in either odd or even way to the complete interval $[-\pi,\pi]$, so there is a basis $\{\sin nx\}_{n=1}^{\infty}$ or $\{\cos nx\}_{n=0}^{\infty}$ respectively (with normalizing factor omitted): if we extend $f\in L_2[0,\pi]$ to $L_2[-\pi,\pi]$ as an odd function, then

$$\int_{-\pi}^{\pi} f(x)\cos nx\,dx = 0,$$

or, for the even extension,

$$\int_{-\pi}^{\pi} f(x)\sin nx\,dx = 0.$$

Bases have a lot of applications. For instance, it allows one to reduce differential or integral equations to finite-dimensional matrix problems, if we consider partial sums.

Basis is also a powerful instrument to compute the sums of series. Consider, for example, the series

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

To compute this sum, one can use Parseval's identity; in order to do so, we have to choose the space, take a basis, and find an appropriate element. Take $L_2[-\pi,\pi]$, the sine-cosine basis $\frac{1}{\sqrt{2\pi}}$, $\frac{1}{\sqrt{\pi}}\cos nx$, $\frac{1}{\sqrt{\pi}}\sin nx$, and the identity function $f(x) = x$. It is an odd function, so the Fourier coefficients in the cosine series of f are equal to 0. Thus, we have to find only the coefficients in sines:

$$\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} x\sin nx\,dx = -\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} x\,d(\frac{\cos nx}{n}) = -\frac{1}{\sqrt{\pi}}\frac{\cos nx}{n}\Big|_{-\pi}^{\pi} + \frac{1}{\sqrt{\pi}n}\int_{-\pi}^{\pi}\cos nx\,dx,$$

where the integral term vanishes, since cosine is 2π-periodic function. Therefore,

$$\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} x\sin nx\,dx = \frac{2\pi}{\sqrt{\pi}n}(-1)^{n+1} = \frac{2\sqrt{\pi}}{n}(-1)^{n+1}.$$

Then,

$$\|f\|^2 = \int_{-\pi}^{\pi} x^2\,dx = \frac{2\pi^3}{3},$$

and, according to Parseval's identity,

$$\frac{2\pi^3}{3} = 4\pi\sum_{n=1}^{\infty}\frac{1}{n^2},$$

which gives the required.

3) Consider $W_2^1[-\pi,\pi]$. Prove that the system $\{e^{inx}\}_{n\in\mathbb{Z}}$ is orthogonal but is not a basis.

To prove the orthogonality, we just calculate the inner product in a straightforward way:

$$(e^{inx}, e^{ikx})_{W_2^1} = \int_{-\pi}^{\pi} e^{i(n-k)x}\,dx + nk\int_{-\pi}^{\pi} e^{i(n-k)x}\,dx = \begin{cases} 0, \ n\neq k, \\ 2\pi(1+n^2), \ n = k, \end{cases}$$

so this system is orthogonal (we can even make this system an ONS by multiplying it by normalizing factor: $\frac{1}{\sqrt{2\pi}\sqrt{n^2+1}}e^{inx}$).

To prove that this is not a basis, it is sufficient to show that either this system is incomplete (so it is necessary to find a nonzero element which is orthogonal to this system) or that Parseval's identity for this system is violated (in order to do so, one can find an element of the space for which it does not hold).

Thus, our options are

1. to find $f \in W_2^1$ such that $f \perp e^{inx}$, $f \neq 0$,

2. to find $f \in W_2^1$ such that $\|f\|^2 \neq \sum_k |c_k|^2$, where c_k is a k-th Fourier coefficient of f.

We will follow the first way. The idea is to find a function that has more than 1 derivative, and take

$$(f, e^{inx})_{W_2^1} = \int_{-\pi}^{\pi} f(x)e^{-inx}\,dx + \int_{-\pi}^{\pi} f'(x)(-in)e^{-inx}\,dx,$$

then, using the integration by parts, we obtain

$$(f, e^{inx})_{W_2^1} = \int_{-\pi}^{\pi} f(x)e^{-inx}\,dx + f'(x)e^{-inx}\Big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} f''(x)e^{-inx}dx,$$

and equate it to zero:

$$\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx + f'(x)e^{-inx}\Big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} f''(x)e^{-inx}dx = 0.$$

This leads to

$$\int_{-\pi}^{\pi} (f(x) - f''(x))e^{-inx}\,dx + \big(f'(\pi) - f'(-\pi)\big)(-1)^n = 0.$$

If we assume that f satisfies the differential equation

$$f - f'' = 0$$

with boundary condition

$$f'(\pi) = f'(-\pi),$$

then f is orthogonal to our system. The solution of this boundary value problem is the hyperbolic sine: $f(x) = a \cdot \sinh x$, and $f(x) \perp e^{inx}$ for any n.

Self-Study Exercises

1) Prove that $\{e_k\}_{k=1}^{\infty}$, $e_k = (0, 0, \ldots, 0, \underset{k\text{-th place}}{1}, 0, \ldots)$, is a basis in ℓ_p, $1 \leqslant p < \infty$, but not a basis in ℓ_∞.

2) Let H be a Hilbert space, and $M \subset H$ be an arbitrary subspace. Prove that

$$(M^\perp)^\perp = \overline{\langle M \rangle}.$$

(Obviously, by the duality property, the double orthogonal complement contains M, and orthogonal complement of any subspace is closed.)

3) Find an example of a closed Euclidean H such that $H \neq H_0 \oplus H_0^{\perp}$ (for Hilbert space, this property holds, so this example must be an incomplete space).

4) Compute $\sum_{k=1}^{\infty} \frac{1}{k^4}$.

5) Compute $\sum_{k=1}^{\infty} \frac{1}{k^2+1}$.

6) Let $H = W_2^1[-\pi, \pi]$, $H_0 = \{f \in W_2^1[-\pi, \pi] : f(x) = 0 \text{ for } x \leqslant 0\}$. Find $H_0^{\perp}$.

Lecture 5. Compact and Precompact Sets in Metric Spaces

Compact Sets. Precompact Sets. Compactness Criteria

We begin by defining the notion of a compact set in a metric space, which plays a fundamental role in functional analysis.

Definition 5.1. *Set $M \subset (X,\rho)$ is* ***compact*** *if for any sequence $\{x_n\}_{n=1}^{\infty} \subset M$ there exists a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ such that*

$$x_{n_k} \xrightarrow{\rho} x \in M.$$

Remark 5.1. *In topological spaces, this kind of compactness is called a* ***sequential compactness****. In metric spaces, these two notions coincide, so we will use it as equivalent definitions, while we do not intend to prove that they are indeed equivalent in this course.*

To recall, in the general topological sense, compactness means the following: for any open covering $\{U_\alpha\}$ of M, $M \subset \cup_\alpha U_\alpha$, there exists a finite subcovering, that is $\exists \alpha_1, \dots, \alpha_n$ such that

$$M \subset \cup_{i=1}^{n} U_{\alpha_i}.$$

Let us emphasize the important difference of compactness in finite-dimensional and infinite-dimensional spaces. Recall that in finite-dimensional spaces, compact sets are just bounded and closed. This result simplifies the verification of compactness significantly. However, in infinite-dimensional spaces, which are of interest in functional analysis, this equivalence does not hold. Therefore, we must develop and rely on alternative criteria to determine compactness in metric and normed spaces.

To introduce related concepts, we now provide a useful definition of ε-nets, which form the foundation of other compactness-related notions.

Definition 5.2. *Let (X,ρ) be a metric space, and $Y,M \subset X$. We say that Y is an* ***ε-net*** *for M if for any $x \in M$ there exists $y \in Y$ such that $x \in B(y,\varepsilon)$.*

In other words, M can be covered by balls of radius ε with centers $y \in Y$:

$$M \subset \cup_{y \in Y} B(y,\varepsilon).$$

A notion, which is closely related to the previous one, is following:

Definition 5.3. *A set $M \subset (X,\rho)$ is called* ***totally bounded*** *if for any ε there exists a finite ε-net for M.*

There is a generalization of compactness for circumstances in which a set is not necessarily closed; the compactness itself is a very strong notion, so a slightly weaker one is useful in functional analysis, as it is preserved for subsets:

Definition 5.4. *A set $M \subset (X, \rho)$ is called **precompact** if its closure $\overline{M}$ is a compact set.*

Notice the subtle difference: precompact sets may not be closed themselves, but their closures must satisfy the compactness criteria. This makes precompactness a slightly weaker property than compactness, yet a highly useful one in many areas of analysis.

Remark 5.2. *Note that the definition of the precompact set is not based on sequences. But the sequences represent a powerful tool, considered in metric spaces.*

Indeed, let us apply it to the notion of precompact set. In metric space (X, ρ), for the set M to be compact, it is necessary that for any $\{x_n\}_{n=1}^{\infty} \subset M$ there exists a Cauchy subsequence. In a complete metric space (X, ρ), it is also sufficient.

Example: Closed Unit Ball in ℓ_2 is Not Compact

Example 5.1. *Consider the closed unit ball in ℓ_2. This set is obviously bounded and closed. At first glance, these properties might suggest compactness, but we will show this is not the case.*

Now consider the standard basis elements of ℓ_2: $\{e_k\}_{k=1}^{\infty}$, where $e_k = (0, \ldots, 0, \overset{k}{1}, 0, \ldots)$ (the 1 is at the k-th position). It is clear that

$$\rho(e_k, e_j) = \|e_k - e_j\| = \sqrt{2}, \quad k \neq j,$$

so there is no Cauchy subsequence.

This example can be generalized, that is, a unit ball in an infinite-dimensional case is a typical example of a noncompact set. We will prove it a little later.

Riesz's Lemma Corollary: Unit Closed Ball is Not Compact in Infinite-Dimensional Space

Theorem 5.1 (Riesz's Lemma)**.** *Let X be a normed space, and X_0 be a nontrivial closed subspace of X. Then, for any $\varepsilon \in (0,1)$, there exists x_ε, $x_\varepsilon \notin X_0$, such that $\|x_\varepsilon\| = 1$ and*

$$\operatorname{dist}(x_\varepsilon, X_0) \equiv \inf_{x_0 \in X_0} \|x_\varepsilon - x_0\| \geqslant 1 - \varepsilon.$$

Proof. First, take some element $x \notin X_0$ (such x exists since X_0 is nontrivial closed subspace, therefore, it does not coincide with X). Define

$$\mathrm{dist}(x, X_0) =: d > 0$$

(it is positive since $x \notin X_0$ and X_0 is closed). Then, there exists $y \in X_0$ such that

$$d \leqslant \|x-y\| < \frac{d}{1-\varepsilon} \tag{5.1}$$

(by the definition of **inf**). Let us define x_ε by

$$x_\varepsilon = \frac{x-y}{\|x-y\|}.$$

Then $\|x_\varepsilon\| = 1$. Now let us see what happens to the distance: for any $x_0 \in X_0$, find the distance between x_ε and x_0:

$$\|x_\varepsilon - x_0\| = \left\| \frac{x-y}{\|x-y\|} - x_0 \right\| = \frac{1}{\|x-y\|} \Big\| x - y - x_0\|x-y\| \Big\|,$$

where $y + x_0\|x-y\| \in X_0$. Now find the bound for the expression above. The factor $1/\|x-y\|$ is bounded from below:

$$\frac{1}{\|x-y\|} \geqslant \frac{1-\varepsilon}{d},$$

see (5.1). The norm is also bounded from below:

$$\Big\| x - y - x_0\|x-y\| \Big\| \geqslant d.$$

Therefore,

$$\|x_\varepsilon - x_0\| > \frac{1-\varepsilon}{d} \cdot d = 1 - \varepsilon,$$

and this bound is valid for an arbitrary $x_0 \in X_0$, thus, the same bound holds for the infimum, which completes the proof. □

Corollary 5.1. *Let X be a normed space,* $\dim X = \infty$*. Then the closed unit ball $B[0,1]$ is not compact in X.*

Proof. First, take some element $x_1 \in X$ such that $\|x_1\| = 1$. Construct a linear span $X_1 := \langle x_1 \rangle$ (it is a one-dimensional subspace). X_1 closed since it is finite-dimensional. By Riesz's Lemma, there exists x_2, $\|x_2\| = 1$, such that

$$\mathrm{dist}(x_2, X_1) \geqslant 1 - \varepsilon.$$

Now define $X_2 = \langle x_1, x_2 \rangle$, where $\|x_1 - x_2\| \geqslant 1 - \varepsilon$, and so on: we find $x_1, x_2, \ldots, x_n$ such that $\|x_j\| = 1$ and $\|x_i - x_j\| \geqslant 1 - \varepsilon$, $i \neq j$, and then construct a finite-dimensional (and, therefore, closed) space $X_n = \langle x_1, x_2, \ldots, x_n \rangle$. By the same reasoning, there exists x_{n+1} such that

$$\mathrm{dist}(x_{n+1}, X_n) \geqslant 1 - \varepsilon;$$

this inequality implies that

$$\|x_{n+1} - x_k\| \geqslant 1 - \varepsilon, \quad k = 1, 2, \ldots, n.$$

By induction, we construct an infinite sequence $\{x_k\}_{k=1}^{\infty} \subset B[0,1]$ such that

$$\|x_i - x_j\| \geqslant 1 - \varepsilon, \quad i \neq j,$$

so there is no Cauchy subsequence, which completes the proof. □

Now we proceed to criteria that allow one to establish whether a set is precompact or not.

Hausdorff Criterion for Precompactness

Theorem 5.2 (Hausdorff criterion)**.** *Let (X, ρ) be a complete metric space. A set $M \subset X$ is precompact if and only if M is totally bounded.*

Remark 5.3. *It can also be shown that in an incomplete space, this condition is a necessary but not sufficient criterion for precompactness.*

Proof.

1) $\Rightarrow$. We will prove the statement by contradiction. Suppose that M is precompact and is not totally bounded. This means that there exists $\varepsilon > 0$ for which there does not exist a finite ε-net.

   Let us begin by taking an arbitrary point $x_1 \in M$; it does not form an ε-net, therefore, there exists $x_2 \in M$: $\rho(x_1, x_2) \geqslant \varepsilon$. The set $\{x_1, x_2\}$ is not an ε-net as well, therefore, there exists $x_3 \in M$ with the same property: $\rho(x_3, x_i) \geqslant \varepsilon$, $i = 1, 2$.

   Now, suppose we have already chosen the points $x_1, x_2, \ldots, x_n \in M$ such that $\rho(x_i, x_j) \geqslant \varepsilon$, $i \neq j$. The set $\{x_i\}_{i=1}^{n}$ still cannot be an ε-net, and therefore, there exists $x_{n+1} \in M$ such that $\rho(x_{n+1}, x_i)$, $i = 1, \ldots, n$.

   By induction, we construct a sequence $\{x_k\}_{k=1}^{\infty}$ with the property $\rho(x_i, x_j) \geqslant \varepsilon$, $i \neq j$, leading to the conclusion that M is not precompact, which gives a contradiction.

2) $\Leftarrow$. Now, assume that M is totally bounded. This part of the proof is also based on the mathematical induction.

   We begin with an arbitrary sequence $\{x_k\}_{k=1}^{\infty} \subset M$. We would like to prove that the set is precompact, so we must show that there exists a Cauchy subsequence of $\{x_k\}_{k=1}^{\infty}$. Take $\varepsilon_1 = 1/2$. For M, there exists an ε_1-net $\{y_1^1, \ldots, y_{n_1}^1\}$ (here, the superscript numerates the step of induction and the subscript numerates the elements of the corresponding net).

Thus,

$$\{x_k\}_{k=1}^{\infty} \subset M \subset \cup_{i=1}^{n_1} B(y_i^1, \varepsilon_1),$$

where we have a countable sequence on the left-hand side and a finite covering on the right-hand side. Therefore, there exists a ball $B(y_{i_1}^1, \varepsilon_1)$ containing an infinite subsequence of $\{x_k\}_{k=1}^{\infty}$; Denote this sequence by $\{x_k^1\}_{k=1}^{\infty}$.

At the second step, take $\varepsilon_2 = 1/4$. For M, there exists an ε_2-net

$$\{y_1^2, y_2^2, \dots, y_{n_2}^2\}.$$

The sequence $\{x_k^1\}_{k=1}^{\infty}$ belongs to a finite union $\cup_{i=1}^{n_2} B(y_i^2, \varepsilon_2)$. Therefore, there exists a ball $B(y_{i_2}^2, \varepsilon_2)$ containing an infinite subsequence of $\{x_k^1\}_{k=1}^{\infty}$; denote this sequence by $\{x_k^2\}_{k=1}^{\infty}$.

By induction, one can construct a countable set of subsequences

$$\{x_k\}_{k=1}^{\infty} \supset \{x_k^1\}_{k=1}^{\infty} \supset \{x_k^2\}_{k=1}^{\infty} \supset \dots \supset \{x_k^m\}_{k=1}^{\infty} \supset \dots$$

such that

$$\rho(x_k^m, x_j^m) < \frac{1}{2^{m-2}},$$

since the entire subsequence $\{x_k^m\}_{k=1}^{\infty}$ lies in the ball

$$B\left(y_{i_{m-1}}^{m-1}, \frac{1}{2^{m-1}}\right).$$

Then, we take the *diagonal* subsequence, that is, $\{x_m^m\}_{m=1}^{\infty}$; it is a Cauchy subsequence, therefore, M is precompact.

Note that in this part of the proof we used the fact that our space is complete. If the space is incomplete, the property of precompactness is not equivalent to the possibility to choose a Cauchy subsequence of any sequence. □

Criteria for Precompactness in Specific Normed Spaces

Building on the Hausdorff criterion, we are to provide criteria for precompactness in specific spaces.

Now, we see that, in a complete space, precompactness and total boundedness, which is close to a topological property (while it is not exactly topological), play a significant role.

We will need an additional tool:

Theorem 5.3 (Dini's Lemma)**.** *Let K be a compact set, $\{f_n\}_{n=1}^{\infty}$ be a continuous function on K, and, for any $x \in K$, $f_n(x) \searrow f(x)$ be a continuous function as well. Then $f_n \underset{K}{\rightrightarrows} f$.*

Remark 5.4. *$f_n(x) \searrow f(x)$ means that $f_n(x)$ approaches $f(x)$ nonincreasingly: $f_n(x) \geqslant f_{n+1}(x)$ for all $n \in \mathbb{N}$ and $x \in K$.*

In calculus, this lemma is usually used for proving that a pointwise limit of a function series is uniform.

The direction of monotonicity is not important: one could multiply the sequence by (-1) to change it.

Proof. Take $\varepsilon > 0$. For any $x \in K$, there exists $N = N(x, \varepsilon)$ such that $\forall n \geqslant N$: $0 \leqslant f_n(x) - f(x) < \varepsilon$.

The function $f_n - f$ is continuous; therefore, there exists a neighborhood U_x of x such that for any $x' \in U_x$:

$$0 \leqslant f_n(x') - f(x') < \varepsilon.$$

$K = \cup_{x \in K} U_x$ is a covering of K. By assumption of the lemma, it is compact, whence, there exist x_i, $i = 1, \ldots, m$, such that $K = \cup_{i=1}^{m} U_{x_i}$.

Take $M = \max_i N(x_i, \varepsilon)$. Then, for any $n \geqslant M$ and $x \in K$: $0 \leqslant f_n(x) - f(x) < \varepsilon$. □

Now we are ready to formulate and prove the criteria for precompactness.

Theorem 5.4. *Let $1 \leqslant p < \infty$. Set $M \subset \ell_p$ is precompact $\Leftrightarrow$ M satisfies the following conditions:*

a) M is bounded,

b) $\forall \varepsilon\ \exists n = n(\varepsilon)$: $\forall x \in M$

$$\left(\sum_{i=n+1}^{\infty} |x_i|^p \right)^{1/p} < \varepsilon$$

The second condition means that *the tails are uniformly small*, or, in other words, the principal parts of our series lie in a finite-dimensional subspace.

Proof. We will use Dini's Lemma to prove the statement in one direction and the Hausdorff criterion for the other one.

1) $\Rightarrow$. Consider the closure of M: $\overline{M}$ is a compact set. The norm $\|\cdot\| : \overline{M} \to \mathbb{R}_0^+$ is a continuous function, therefore, there exists

$$\max_{x \in \overline{M}} \|x\| =: C \geqslant 0,$$

that is, for any $x \in M$:

$$\|x\| \leqslant C,$$

which is exactly the item a).

Now consider the functions f_n on $\overline{M}$:

$$f_n(x) = \Big(\sum_{i=n+1}^{\infty} |x_i|^p \Big)^{1/p};$$

it is clear that $f_n(x) \searrow 0$ as $n \to \infty$ since it is tail of a converging series, and f_n is continuous since

$$f_n(x) = \|(0,0,\dots,0,x_{n+1},x_{n+2},\dots)\|$$

and $\|\cdot\|$ is continuous. By Dini's Lemma, we conclude that $f_n \underset{\overline{M}}{\rightrightarrows} 0$, and, therefore, $f_n \underset{M}{\rightrightarrows} 0$, which is the item b).

2) $\Leftarrow$. By b), there exists $n = n(\varepsilon)$ such that for any $x \in M$:

$$\Big(\sum_{i=n+1}^{\infty} |x_i|^p \Big)^{1/p} < \varepsilon.$$

Define

$$x^n = (x_1, x_2, \dots, x_n, 0, 0, \dots) \in \ell_p(n)$$

and

$$z^n = (0,0,\dots,0,x_{n+1},\dots), \quad \|z^n\| < \varepsilon.$$

We can say that $x^n \in M \cap \ell_p(n)$: it is bounded by a) and lies in a finite-dimensional subspace, so $\{x^n\}_{x \in M}$ is a precompact set. Thus, there exists a finite ε-net $y^1, \dots, y^m \in \ell_p(n)$ of the form

$$y^k = (y_1^k, y_2^k, \dots, y_n^k), \quad k = 1, \dots, m.$$

Any y^k can be embedded into ℓ_p: $y_k \to \widetilde{y}^k$ such that

$$\widetilde{y}^k = (y_1^k, y_2^k, \dots, y_n^k, 0, 0, \dots) \in \ell_p.$$

Let us take an arbitrary $x \in M$. How can we prove that the norm $\|x - \widetilde{y}^k\|$ is small? Decompose x into $x^n + z^n$ and use the triangle inequality:

$$\|x - \widetilde{y}^k\| = \|x^n - \widetilde{y}^k + z^n\| \leqslant \|x^n - y^k\| + \|z^n\|.$$

We can make the second term small, i.e., $\|z^n\| < \varepsilon$, by choosing an appropriate n; the first one is small for an appropriate k: $\exists k$ such that $\|x^n - y^k\| < \varepsilon$. Thus, $\{\widetilde{y}^k\}_{k=1}^m$ is a finite 2ε-net of M, therefore, M is precompact by the Hausdorff criterion. □

To formulate the theorem on precompact sets in $C[a,b]$, we will need the following definition.

Definition 5.5. *A set $M \subset C[a,b]$ is called an* ***equicontinuous family*** *of functions if for any $\varepsilon > 0$ there exists $\delta > 0$: $\forall x, y \in [a,b]$ such that $|x-y| < \delta$ and for all $f \in M$: $|f(x) - f(y)| < \varepsilon$.*

Example 5.2. *Suppose the set consists of a single function: $M = \{f\}$, $f \in C[a,b]$. It is equicontinuous, since, in this case, the property of equicontinuity is equivalent to the uniform continuity.*

The same is true if M contains a finite number of functions: $M = \{f_i\}_{i=1}^n$, so it is more interesting to consider infinite sets of functions.

Remark 5.5. *One can define an equicontinuous family $M \subset C(K)$ for a compact metric space (K, ρ) with replacing $|x-y|$ by $\rho(x,y)$.*

Now, we formulate the Arzelà–Ascoli theorem on precompact sets in $C[a,b]$ and prove it on the next lecture.

Theorem 5.5 (Arzelà–Ascoli)**.** *A set $M \subset C[a,b]$ is precompact $\Leftrightarrow$ the following conditions hold:*

a) M is bounded,

b) M is an equicontinuous family.

Lecture 6. Compact and Precompact Sets in Metric Spaces: Exercises

Proof of the Arzelà–Ascoli Theorem

1) $\Rightarrow$. Suppose $M \subset C[a,b]$ is precompact; let us try to prove that M is bounded and forms an equicontinuous family.

   As before, the proof in this direction will be based on Dini's lemma.

   First, to prove a), consider the closure of M: $\overline{M}$ is compact; the norm is a continuous function on $\overline{M}$, so there exists $\max\limits_{f\in\overline{M}} = C$, therefore, $\forall f \in M \Rightarrow \|f\| \leqslant C$.

   To prove b), consider a function F_n on $\overline{M}$:

$$F_n(f) := \sup_{|x-y|<\frac{1}{n}} |f(x) - f(y)|.$$

   It is clear that we just replaced a continuous parameter δ in the definition of equicontinuity with a discrete parameter $1/n$.

   One can see that the sequence of functions $F_n(f)$ approaches 0 from above as $n \to \infty$ since $f \in C[a,b]$.

   Consider also the functions F_n for different functions, say, $f, g \in C[a,b]$:

$$\Big|F_n(f) - F_n(g)\Big| = \sup_{|x-y|<\frac{1}{n}} \Big|f(x) - f(y)\Big| - \sup_{|x-y|<\frac{1}{n}} \Big|g(x) - g(y)\Big|.$$

   Now, add $-g(x) + g(x) - g(y) + g(y)$ to the first supremum and use the triangle inequality:

$$\sup_{|x-y|<\frac{1}{n}} \Big|f(x) - f(y) - g(x) + g(x) - g(y) + g(y)\Big| - \sup_{|x-y|<\frac{1}{n}} \Big|g(x) - g(y)\Big| \leqslant$$
$$\leqslant \Big| \sup_{|x-y|<\frac{1}{n}} \Big|f(x) - g(x)\Big| + \sup_{|x-y|<\frac{1}{n}} \Big|g(x) - g(y)\Big| + \sup_{|x-y|<\frac{1}{n}} \Big|g(y) - f(y)\Big| - \sup_{|x-y|<\frac{1}{n}} \Big|g(x) - g(y)\Big|\Big|.$$

   The second and the fourth terms here are equal, so they cancel out. The first and the third ones are equal up to the replacement $x \leftrightarrow y$, which is legal since the expression

$$\sup_{|x-y|<\frac{1}{n}} \Big|g(y) - f(y)\Big|$$

   is symmetric with respect to this replacement. Hence, we obtain

$$\Big|F_n(f) - F_n(g)\Big| \leqslant 2 \max_{x\in[a,b]} |f(x) - g(x)| = 2\|f - g\|_{C[a,b]},$$

and, recalling that the norm is a continuous function, we conclude that F_n are continuous.

Then, by Dini's lemma, $F_n \underset{M}{\rightrightarrows} 0$, therefore, $F_n \underset{M}{\rightrightarrows} 0$, which is the very condition b) with parameter δ being replaced by $1/n$.

2) $\Leftarrow$. Suppose that M is bounded and forms an equicontinuous family, and prove that M is precompact. The idea is to construct a finite ε-net for an arbitrary ε, and then use the Hausdorff criterion.

Without loss of generality, we consider only real-valued functions. To generalize our proof, one can use the decomposition $f(x) = u(x) + iv(x)$ and apply our proof for $u(x)$ and $v(x)$.

By a), there exists $C > 0$ such that $\forall f \in M$: $\max\limits_{[a,b]}|f(x)| \leqslant C$. By b),

$$\forall \varepsilon > 0 \ \exists \delta > 0 : \ \forall x, y \in [a,b], \ |x-y| < \delta \Rightarrow \forall f \in M : \ |f(x) - f(y)| < \frac{\varepsilon}{3}.$$

Take a subdivision of $[a,b]$:

$$T = \{t_i\}_{i=0}^n, \quad a = t_0 < t_1 < t_2 < \dots < t_{n-1} < t_n = b,$$

such that

$$\forall i : \ |t_i - t_{i-1}| < \delta, \quad i = 1, 2, \dots, n.$$

Construct a lattice with t_i, $i = 1, \dots, n$, in x-axis and the distance $\varepsilon/3$ from $-C$ to C in y-axis, see Fig. 6.1.

So, we have a set with a finite number of nodes. Consider the set $Y = \{g(x)$ piecewise linear functions passing through the nodes$\}$, see an example in Fig. 6.2. The set Y is finite.

Let us take $t \in [t_i, t_{i+1}]$, $g \in Y$, and $f \in C[a,b]$. Then

$$\left|f(t) - g(t)\right| \leqslant \left|f(t) - f(t_i)\right| + \left|f(t_i) - g(t_i)\right| + \left|g(t_i) - g(t)\right|.$$

The first summand here is $< \varepsilon/3$ by equicontinuity; the second one is $< \varepsilon/3$ by choosing the function g, and the third one is $< \varepsilon/3$ by the property of the set Y. Thus,

$$\left|f(t) - g(t)\right| < \varepsilon,$$

therefore, Y is a finite ε-net for M. Hence, by the Hausdorff criterion, M is precompact. □

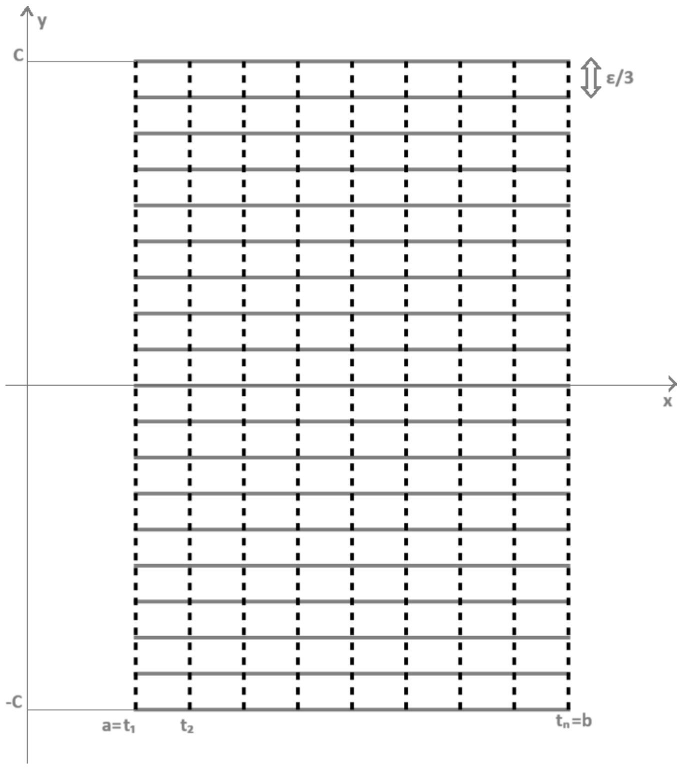

Рис. 6.1. The lattice

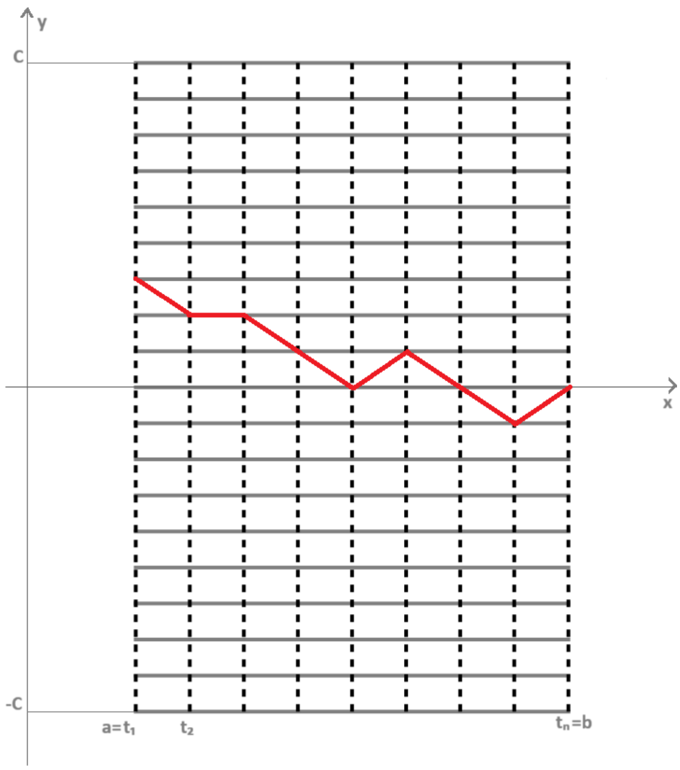

Рис. 6.2. An example of piecewise linear function on the lattice

Theorem on Precompact Sets in L_p

In this section, we formulate a theorem on criteria of precompactness in $L_p[a,b]$ without a proof.

Theorem 6.1. *A set $M \subset L_p[a,b]$, $1 \leqslant p < \infty$, is precompact $\Leftarrow$ the following conditions hold:*

a) M is bounded,

b) $\forall \varepsilon > 0\ \exists \delta > 0$: $\forall h$, $|h| < \delta \Rightarrow \forall f \in M$:

$$\left(\int_a^b \left| f(x+h) - f(x) \right|^p d\mu \right)^{1/p} < \varepsilon.$$

Remark 6.1. *The second condition is called* ***equicontinuity in mean****. Note also that if $x+h \notin [a,b]$, then $f(x+h) := 0$.*

Discussion of Self-Study Exercises from the Previous Lecture

Now we discuss the homework from Lecture 4.

1) Show that $\{e_k\}_{k=1}^{\infty}$, $e_k = (0,\ldots,0,\overset{k}{1},0,\ldots)$ is a basis in ℓ_p, $1 \leqslant p < \infty$ and is not a basis in ℓ_∞.

   The second part is quite simple: ℓ_∞ is not separable, so it cannot have a countable basis.

   However, ℓ_p, with p being finite, can have one. Take $x \in \ell_p$, $x = (x_1, x_2, \ldots, x_n, x_{n+1}, \ldots)$, and consider the representation

   $$x = \sum_{k=1}^{\infty} x_k e_k.$$

   One can see that this representation is unique since we have fixed coordinates.

   Consider the remainder for an approximation with a finite number of e_j:

   $$\left\| x - \sum_{k=1}^{n} x_k e_k \right\| = \left(\sum_{k=n+1}^{\infty} |x_k|^p \right)^{1/p} \to 0 \quad \text{as} \quad n \to \infty$$

   by definition of $x \in \ell_p$. Therefore, $\{e_k\}_{k=1}^{\infty}$ is a basis in ℓ_p.

2) Let $M \subset H$, where H is a Hilbert space. Prove that $\left(M^{\perp}\right)^{\perp} = \overline{\langle M \rangle}$.

   We know that $M^{\perp}$ is a closed linear subspace. By duality, it is clear that $\left(M^{\perp}\right)^{\perp} \supset \overline{\langle M \rangle}$, so we now have to prove the inverse inclusion. Let us try to obtain two different representations for H:

$$H = \overline{\langle M \rangle} \oplus \left(\overline{\langle M \rangle}\right)^{\perp} \quad \text{and} \quad H = \left(M^{\perp}\right)^{\perp} \oplus M^{\perp}. \tag{6.1}$$

   Here, $M^{\perp} = \overline{\langle M \rangle}^{\perp}$; let us prove it. $M \subset \overline{\langle M \rangle}$, and therefore, $M^{\perp} \supset \overline{\langle M \rangle}^{\perp}$; if $x \in M^{\perp}$, which means that $(x,y) = 0\ \forall y \in M$, then $(x, \alpha y_1 + \beta y_2) = \alpha(x,y_1) + \beta(x,y_2) = 0$ $\forall y_1, y_2 \in M \Rightarrow x \in \overline{\langle M \rangle}^{\perp}$. We also know that the orthogonal complement is closed, so $x \in \overline{\langle M \rangle}^{\perp}$.

   Therefore, the second terms of decompositions (6.1) coincide. Since this decomposition is unique, we immediately obtain that the first terms coincide as well, that is, $\overline{\langle M \rangle} = \left(M^{\perp}\right)^{\perp}$.

3) Find an example of a closed Euclidean space H such that $H \neq H_0 \oplus H_0^{\perp}$.

   Consider the space $H = C_2[-1,1]$ (a real-valued one) with

$$(f,g) = \int_{-1}^{1} f(x)g(x)\,dx.$$

   The norm here is given by

$$\|f-g\|_2 = \left(\int_{-1}^{1} |f(x)-g(x)|^2\,dx\right)^{1/2}.$$

   The incompleteness of $C_1[0,1]$ was discussed on the first lecture. $C_2[-1,1]$ is incomplete as well.

   Take

$$H_0 = \{f \in C_2[-1,1]:\ f(x) = 0 \text{ for } x \in [-1,0)\}.$$

   In $C_2[-1,1]$, it is a closed subspace. One can see that

$$H_0^{\perp} = \{f \in C_2[-1,1]:\ f(x) = 0 \text{ for } x \in [0,1]\}.$$

   Next, considering the sum of these spaces, we see that

$$H_0 \oplus H_0^{\perp} = \{f \in C_2[-1,1] : f(0) = 0\} \neq H,$$

   since the sum consists only of functions vanishing at $x = 0$.

4) Calculate $\sum_{k=1}^{\infty}\frac{1}{k^4}$.

Take $L_2[-\pi,\pi]$ along with a basis

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{1}{\sqrt{\pi}}\cos nx, \quad \frac{1}{\sqrt{\pi}}\sin nx, \quad n\in\mathbb{N},$$

and $f(x)=x^2$. This function is even, so its Fourier series consists only of cosines. It is clear that

$$f_0=\frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi}x^2\,dx=\frac{1}{\sqrt{2\pi}}\frac{2x^3}{3}\Big|_0^{\pi}=\frac{2\pi^3}{3\sqrt{2\pi}}.$$

Next, compute coefficients in cosines:

$$f_n=\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi}x^2\cos nx\,dx;$$

it can be integrated by parts:

$$\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi}x^2\frac{1}{n}d(\sin nx)=\frac{1}{\sqrt{\pi}}x^2\frac{\sin nx}{n}\Big|_{-\pi}^{\pi}-\frac{2}{\sqrt{\pi}n}\int_{-\pi}^{\pi}x\sin nx\,dx,$$

where the first term vanishes, and we get

$$\frac{2}{\sqrt{\pi}n^2}\int_{-\pi}^{\pi}x\,d(\cos nx)=\frac{2}{\sqrt{\pi}n^2}x\cos nx\Big|_{-\pi}^{\pi}-\frac{2}{\sqrt{\pi}n^2}\int_{-\pi}^{\pi}\cos nx\,dx,$$

where the last term vanishes since it is integration of a periodic function over the period, so we finally obtain

$$f_n=\frac{4\pi(-1)^n}{\sqrt{\pi}n^2}.$$

Let us use Parseval's identity. First, find the squared norm:

$$\|f\|^2=\int_{-\pi}^{\pi}x^4\,dx=\frac{2x^5}{5}\Big|_0^{\pi}=\frac{2\pi^5}{5}.$$

Next, equate this to the sum of squared Fourier coefficients:

$$\frac{2\pi^5}{5}=f_0^2+\sum_{n=1}^{\infty}|f_n|^2\equiv\frac{4\pi^6}{9\cdot 2\pi}+\sum_{n=1}^{\infty}\frac{16\pi}{n^4}.$$

Thus,

$$2\pi^5\Big(\frac{1}{5}-\frac{1}{9}\Big)=16\pi\sum_{n=1}^{\infty}\frac{1}{n^4},$$

and, simplifying it, we get

$$\sum_{n=1}^{\infty}\frac{1}{n^4}=\frac{\pi^4}{90}.$$

One can calculate

$$\sum_{n=1}^{\infty}\frac{1}{n^{2k}}=\frac{\pi^{2k}}{B_{2k}}$$

using the same basis in $L_2[-\pi,\pi]$ and the function $f=x^k$, where B_{2k} is a sequence somehow related to Bernoulli numbers.

5) Calculate $\sum_{k=1}^{\infty}\frac{1}{k^2+1}$.

Take $L_2[-\pi,\pi]$, a basis $\frac{1}{\sqrt{2\pi}}e^{inx}$, $n\in\mathbb{Z}$, and the function $f(x)=e^{-x}$. Now we find

$$\sqrt{2\pi}f_n=(f,e^{inx})=\int_{-\pi}^{\pi}e^{-x}e^{-inx}\,dx=\int_{-\pi}^{\pi}e^{-(1+in)x}\,dx=\frac{-1}{1+in}e^{-(1+in)x}\Big|_{-\pi}^{\pi},$$

or, simplifying it,

$$\sqrt{2\pi}f_n=\frac{(-1)^n(e^{\pi}-e^{-\pi})(1-in)}{1+n^2},$$

i.e.,

$$\sqrt{2\pi}|f_n|=\frac{(e^{\pi}-e^{-\pi})Sqrt1+n^2}{1+n^2}=\frac{2\sinh\pi}{\sqrt{1+n^2}},\quad n\in\mathbb{Z}.$$

Find the norm:

$$\|f\|^2=\int_{-\pi}^{\pi}e^{-2x}\,dx=-\frac{1}{2}e^{-2x}\Big|_{-\pi}^{\pi}=\frac{1}{2}\left(e^{2\pi}-e^{-2\pi}\right)=2\sinh\pi\cosh\pi.$$

Write down Parseval's identity:

$$2\sinh\pi\cosh\pi=f_0^2+2\sum_{n=1}^{\infty}|f_n|^2,$$

where the coefficient 2 for sum is taken since for $n'=-n$ we have the same expression under the sum. Thus,

$$\cosh\pi=\sinh\pi+\frac{2\sinh\pi}{\pi}\sum_{n=1}^{\infty}\frac{1}{n^2+1},$$

or, after simplification,

$$\sum_{n=1}^{\infty}\frac{1}{n^2+1}=\frac{1}{2}\Big(\pi\coth\pi-1\Big).$$

Exercise 6.1. *Try to calculate*

$$\sum_{n=1}^{\infty}\frac{1}{n^2+a^2},\quad a>0.$$

6) $H = W_2^1[-1,1]$,

$$H_0 = \{f \in W_2^1[-1,1] : \ f(x) = 0 \text{ for } x \leqslant 0\}.$$

Find $H_0^{\perp}$.

For $g \in H_0^{\perp}$, $\forall f \in H_0$: $(f,g)_{W_2^1} = 0$. By the definition of the inner product in W_2^1,

$$\int_0^1 f(x)\overline{g(x)}\,dx + \int_0^1 f'(x)\overline{g'(x)}\,dx = 0.$$

The second integral can be rewritten as

$$\int_0^1 f'(x)\overline{g'(x)}\,dx = \int_0^1 \overline{g'(x)}\,df = g'(x)f(x)\Big|_0^1 - \int_0^1 \overline{g''(x)}f(x)\,dx,$$

so we arrive at the equation

$$\int_0^1 f(x)\overline{\Big(g(x) - g''(x)\Big)}\,dx + \overline{g'(1)}f(1) - \overline{g'(0)}f(0) = 0,$$

where $f(0) = 0$. A sufficient condition for g to satisfy this equation, for example, can be given by

$$g(x) - g''(x) = 0,$$
$$g'(1) = 0.$$

We will seek for solutions of the form

$$g(x) = a\sinh(x-1) + b\cosh(x-1),$$

so $g'(x) = a\cosh(x-1) + b\sinh(x-1)$, and $g'(1) \equiv a = 0$. Therefore,

$$g(x) = \begin{cases} b\cosh(x-1) \text{ for } x \in [0,1], \\ \text{an arbitrary function for } x \in [-1,0] \end{cases}$$

with a condition that

$$g(-0) - b\cosh(-1) \equiv b\cosh 1 \quad \Rightarrow \quad b = \frac{g(0)}{\cosh 1}$$

since g must belong to W_2^1. Hence,

$$H_0^{\perp} = \Big\{g \in W_2^1[-1,1] : \ g(x) = \frac{g(0)}{\cosh 1}\cosh(x-1), \ x \geqslant 0, \\ \text{and an arbitrary } \widetilde{g} \in W_2^1[-1,0], \ x \leqslant 0\Big\}. \tag{6.2}$$

The only tricky thing here is that we found the function $g(x)$ as a solution of second-order differential equation, therefore, we assumed that it has 2 derivatives. We have to show that (6.2) is the entire orthogonal complement.

Take $f \in W_2^1[-1,1]$ and decompose it:

$$f = f_0 + f_1, \quad f_0 \in H_0, \quad f_1 \in H_0^{\perp}.$$

One can see that

$$f_1 = \begin{cases} \dfrac{f(0)}{\cosh 1} \cosh(x-1), & x \in [0,1], \\ f(x), \quad x \in [-1,0]. \end{cases}$$

It is also easy to see that this function is continuous at $x = 0$.

For f_1 of this form,

$$f_0 = f - f_1, \quad \text{and} \quad f_0\Big|_{[-1,0]} = 0,$$

so $H_0^{\perp}$ is indeed the entire orthogonal complement.

Exercises on Precompactness

1) Consider a set

$$M = \{x \in \ell_p : \ |x_k| \leqslant a_k\}, \quad 1 \leqslant p < \infty,$$

where $\{a_k\}_{k=1}^{\infty}$ is some certain sequence. Prove that M is precompact $\Leftrightarrow$ $\{a_k\}_{k=1}^{\infty} \in \ell_p$.

a) $\Leftarrow$. In this direction, the proof is simple:

$$\|x\| = \Big(\sum_{k=1}^{\infty} |x_k|^p\Big)^{1/p} \leqslant \Big(\sum_{k=1}^{\infty} a_k^p\Big)^{1/p} < \infty.$$

Also, for $\forall x \in M$, the tail is small: $\forall \varepsilon\ \exists n$ such that

$$\Big(\sum_{k=n+11}^{\infty} |x_k|^p\Big)^{1/p} \leqslant \Big(\sum_{k=n+1}^{\infty} a_k^p\Big)^{1/p} < \varepsilon,$$

since $a_k \in \ell_p$.

b) $\Rightarrow$. Let $\{a_k\}_{k=1}^{\infty} \notin \ell_p$. Note that these numbers are nonnegative: $a_k \geqslant 0$. Therefore,

$$S_n := \Big(\sum_{k=1}^{n} a_k\Big)^{1/p} \to +\infty \quad \text{as} \quad n \to \infty.$$

Consider $x^n \in M$:

$$x = (a_1, a_2, \dots, a_n, 0, 0, \dots) \in \ell_p.$$

This sequence belongs to ℓ_p, but $\|x^n\| \to +\infty$, so the set M is unbounded, which gives us a contradiction.

2) Study the equicontinuity of the system $\{f_n(x) = x^n\}_{n=1}^{\infty}$ in $C[0,1]$.

   It is clear that $\|f_n\| = 1$, so it is a bounded set. To study the precompactness of this set, we have to find out only whether it is equicontinuous or not.

   Let us take $x = 1$ and $y = 1 - \delta/2$, $|x - y| = \delta/2 < \delta$. Calculate

$$\left|f_n(x) - f_n(y)\right| = 1 - \left(1 - \frac{\delta}{2}\right)^n,$$

   where

$$1 - \frac{\delta}{2} < 1 \quad \Rightarrow \quad \exists n: \quad \left(1 - \frac{\delta}{2}\right)^n < \frac{1}{2}.$$

   Whence,

$$\exists n: \quad \left|f_n(x) - f_n(y)\right| > \frac{1}{2},$$

   which gives us a contradiction with the property of equicontinuity.

Self-Study Exercises

1) Consider an ellipsoid in ℓ_2:

$$M = \left\{x \in \ell_2 : \sum_{i=1}^{\infty} \frac{|x_i|^2}{a_i^2} \leqslant 1\right\}.$$

   Prove that M is precompact if and only if $\{a_i\}_{i=1}^{\infty} \in c_0$.

2) Consider $\{\sin nx\}_{n=1}^{\infty}$. Find out whether it is precompact in $C[0,1]$ or not.

3) Consider $\{\sin \alpha x\}_{\alpha \in [1,2]}$. Find out whether it is precompact in $C[0,1]$ or not.

4) Consider

   a) $M_1 = \left\{f \in C^1[a,b] : \ |f(a)| \leqslant c_1 \text{ and } \int_a^b |f'(x)|\, dx \leqslant c_2\right\}$,

   b) $M_2 = \left\{f \in C^1[a,b] : \ |f(a)| \leqslant c_1 \text{ and } \int_a^b |f'(x)|^2\, dx \leqslant c_2\right\}$,

   c) $M_3 = \left\{f \in C^1[a,b] : \ \int_a^b \left(|f(x)|^2 + |f'(x)|^2\right) dx \leqslant c\right\}$,

   where c_1, c_2, and c are some constants. Study the compactness of these sets.

5) Prove that the unit ball $B[0,1] \subset L_2[0,1]$ is not precompact in $L_1[0,1]$ (note that $L_2[0,1] \subset L_1[0,1]$).

6) Show that

   a) a unit ball $B[0,1] \subset C^1[0,1]$ is precompact in $C[0,1]$,

   b) a unit ball $B[0,1] \subset W_2^1[0,1]$ is precompact in $L_2[0,1]$.

Lecture 7. Linear Operators and Functionals in Normed Spaces

Linear Operators in Normed Spaces. Bounded Operators

Let us begin with the following definition:

Definition 7.1. *Let X, Y be linear spaces over one field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. A map $A : X \to Y$ is called a* ***linear operator*** *if $\forall \alpha, \beta \in \mathbb{K}$, $x_1, x_2 \in X$: $A(\alpha x_1 + \beta x_2) = \alpha A x_1 + \beta A x_2$.*

If X and Y are normed spaces, a ***norm of an operator*** *can be also defined:*

$$\|A\|_{X \to Y} := \sup_{X \ni x \neq 0} \frac{\|Ax\|_Y}{\|x\|_X}.$$

It is easy to verify that this expression indeed defines a norm: it is is nonnegative, vanishes only for the zero operator, it is homogeneous with respect to multiplication on the elements of the field (up to an absolute value), and the triangle inequality holds due to the fact that it holds for the norm in Y. Define also some spaces of operators:

Definition 7.2. *$\mathcal{L}(X,Y)$ is the space of all linear operators $X \to Y$; note that linear operations in this space are well-defined: $(A+B)x = Ax + Bx$ and $\forall \alpha \in \mathbb{K}$: $(\alpha A)(x) = \alpha(Ax)$.*

Let $A \in \mathcal{L}(X,Y)$, where X and Y are normed spaces. A is ***bounded*** *if $\|A\| < \infty$ (it is usually denoted as $A \in B(X,Y)$).*

Consider two additional ways to find the norm: taking only the elements from a unit sphere or from a unit ball:

$$\|A\|_1 = \sup_{\|x\|=1} \|Ax\|, \qquad \|A\|_2 = \sup_{\|x\| \leqslant 1} \|Ax\|.$$

Proposition 7.1. $\|A\| = \|A\|_1 = \|A\|_2$.

Proof. Note that $\|A\|_1 \leqslant \|A\|_2$ since $\{\|x\| = 1\} \subset \{\|x\| \leqslant 1\}$, and $\|A\|_1 \leqslant \|A\|$, which follows from $\sup\limits_{x \neq 0} \frac{\|Ax\|}{\|x\|}$ if we put here $\|x\| = 1$.

To prove the statement, we have to show the validity of inverse inequalities. Rewrite:

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|},$$

including $\|x\|$ into the norm in the numerator:

$$\sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{x \neq 0} \left\| A \frac{x}{\|x\|} \right\| \leqslant \|A\|_1.$$

Further,

$$\|A\|_2 = \sup_{\|x\|\leqslant 1}\|Ax\| = \|A\|_2 = \sup_{\|x\|\leqslant 1, x\neq 0}\|Ax\| = \|A\|_2 = \sup_{\|x\|\leqslant 1, x\neq 0}\|x\|\|A\frac{x}{\|x\|}\|,$$

where the norm of $x/\|x\|$ is equal to 1, so $\|A\|_2 \leqslant \|A\|_1$. □

Remark 7.1. *From the definition of the norm, we can obtain the following inequalities:*

$$\|A\| \geqslant \frac{\|Ax\|}{\|x\|}\ (\forall x\neq 0) \quad \Rightarrow \quad \forall x:\ \|Ax\| \leqslant \|A\|\|x\|.$$

Usually, the way to find the norm of an operator is following: begin with $\|Ax\|$, and use some classical inequalities to estimate it with $\|x\|$:

$$\|Ax\| \leqslant \cdots \leqslant C\cdot\|x\|,$$

then the norm of A is bounded from above by C. If the inequalities used on this way are sharp, then C may be exactly the norm of A.

There are two possible ways to show that an upper bound for the norm is sharp:

1) Find x, $\|x\| = 1$, such that $\|Ax\| = C$, or

2) Find a sequence $\{x_n\}_{n=1}^{\infty}$, $\|x_n\| = 1$, such that $\|Ax_n\| \nearrow C$ as $n\to\infty$;

any of these allows one to conclude that $\|A\| = C$.

Examples: Finding Norms of Operators

Take some $\varphi \in C[a,b]$. Consider an operator of multiplication by the function φ:

$$A_\varphi f(x) = \varphi(x)f(x).$$

For instance, A_φ with $\varphi(x) = x$, called an *operator of coordinate*, is one of the important subjects of study in Quantum Mechanics.

Let us find the norm of this operator acting in the following spaces:

a) $A_\varphi : C[a,b]\to C[a,b]$,

b) $A_\varphi : L_2[a,b]\to L_2[a,b]$.

In case a),

$$\|Af\| = \max_{[a,b]}\big|\varphi(x)f(x)\big| \leqslant \max_{[a,b]}\big|\varphi(x)\big|\cdot\max_{[a,b]}\big|f(x)\big| = \|\varphi\|_{C[a,b]}\cdot\|f\|_{C[a,b]},$$

therefore, $\|A\| \leqslant \|\varphi\|_{C[a,b]}$. Take $f_0 \equiv 1$ on $[a,b]$. For this function, $\|f_0\| = 1$ and $\|Af_0\| = \|\varphi\|_{C[a,b]}$, so $\|A\| = \|\varphi\|_{C[a,b]}$.

For example, on $C[a,b]$, the operator A_x that acts as $Af = xf(x)$ has norm $\|A\| = 1$.

In case b),

$$\|Af\|^2 = \int_a^b |\varphi(x)f(x)|^2\,dx \leqslant \max_{[a,b]}|\varphi(x)|^2 \int_a^b |f(x)|^2\,dx = \|\varphi\|^2_{C[a,b]} \cdot \|f\|^2_{L_2}.$$

Thus, $\|A\| \leqslant \max\limits_{[a,b]}|\varphi(x)|$. In fact, this bound is sharp. While so, the proof requires to consider a sequence of functions from L_2, since the norm of a constant here is not equal to the constant itself, but is equal to the length of the interval.

We know that the function φ is continuous; therefore, there exists a point $x_0 \in [a,b]$ such that $|\varphi(x_0)| = \max\limits_{[a,b]}|\varphi(x)|$. Without loss of generality, we can assume that this is an interior point of the interval $[a,b]$; if it is an end of the interval, we can consider a one-sided neighborhood. For an interior point, we consider a usual neighborhood: consider the following functions $\{f_n\}_{n=1}^{\infty}$:

$$f_n(x) = \begin{cases} \sqrt{n}, & x \in (x_0 - \dfrac{1}{2n}, x_0 + \dfrac{1}{2n}), \\ 0, & otherwise. \end{cases}$$

The limit function takes the value of $\varphi(x)$ at the point x_0, so it is the delta function $\delta_{x_0}(x)$, see an example in Fig. 7.1.

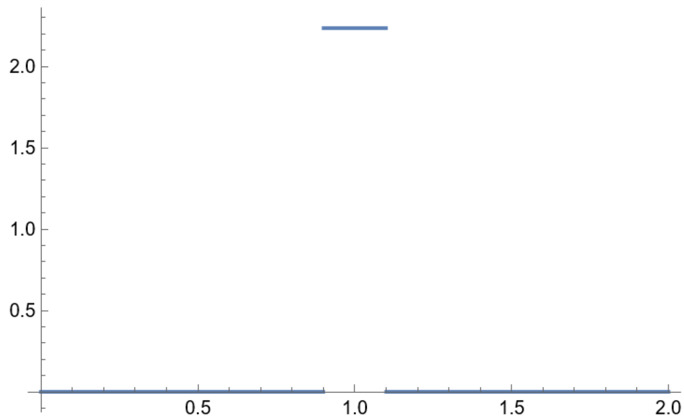

Рис. 7.1. Example: f_5 for $x_0 = 1$.

The norm of these function is equal to 1:

$$\|f_n\|_{L_2} = \left(\int_a^b |f_n(x)|^2\,dx\right)^{1/2} = \left(\int_{x_0-\frac{1}{2n}}^{x_0+\frac{1}{2n}} \sqrt{n}\,dx\right)^{1/2} = 1.$$

Now we find the norm of $\|Af_n\|$:

$$\|Af_n\| = \Big(\int_{x_0-\frac{1}{2n}}^{x_0+\frac{1}{2n}} |\varphi(x)|^2 n\,dx\Big)^{1/2} = \Big(n\int_{x_0-\frac{1}{2n}}^{x_0+\frac{1}{2n}} |\varphi(x)|^2\,dx\Big)^{1/2}; \tag{7.1}$$

since $\varphi(x)$, along with $|\varphi(x)|^2$, is a continuous function, according to the mean value theorem for integrals, there exists at least one point $x_n \in [x-1/(2n), x+1/(2n)]$ such that

$$|\varphi(x_n)|^2 = \frac{1}{n}\int_{x_0-\frac{1}{2n}}^{x_0+\frac{1}{2n}} |\varphi(x)|^2\,dx.$$

Plugging this into (7.1), we finally obtain $\|Af_n\| = |\varphi(x_n)|$. Since φ is continuous, and the length of the interval $(x-1/(2n), x+1/(2n))$ approaches zero as $n\to\infty$,

$$\|Af_n\| \to |\varphi(x_0)| = \|\varphi\|_{C[a,b]}.$$

Continuous Operators. Theorem on Equivalence of Boundedness and Continuity. B(X, Y) is Banach if Y is Banach

Recall the notation:

$\mathcal{L}(X,Y)$ is the space of all linear operators $X\to Y$ and $B(X,Y)$ is the space of all bounded linear operators $X\to Y$. If $X=Y$, we simply write $\mathcal{L}(X)$ and $B(X)$. Now we introduce the following kind of linear operators:

Definition 7.3. *Let $A\in\mathcal{L}(X,Y)$, where X and Y are normed spaces.*

*1) A is **continuous at point** $x_0\in X$, if $(x_n\to x)\Rightarrow(Ax_n\to Ax)$.*

*2) A is **continuous** if A is continuous at any point $x\in X$.*

Theorem 7.1. *Let X and Y be normed spaces, and $A\in\mathcal{L}(X,Y)$. Then the following are equivalent:*

1) A is continuous at a point x_0,

2) A is continuous,

3) A is bounded.

Thus, the continuity is a synonym for the boundedness in the context of linear operators between normed spaces.

Proof. $2 \Rightarrow 1$ is obvious. Let us prove $1 \Rightarrow 2$. Let A be continuous at x_0 and $x_n \to x$. Then $x_n - x + x_0 \to x_0$. Applying A, we get

$$A(x_n - x + x_0) \to Ax_0 \overset{A \in \mathcal{L}(X,Y)}{\Rightarrow} Ax_n - Ax + Ax_0 \to Ax_0,$$

therefore, $Ax_n \to Ax$.

Now we prove $3 \Rightarrow 2$. Let $x_n \to x$;

$$\|Ax_n - Ax\| = \|A(x_n - x)\| \leqslant \|A\| \cdot \|x_n - x\|,$$

where the first term is finite since A is bounded, and the second one tends to zero. Therefore, $\|Ax_n - Ax\| \to 0$, so A is continuous.

The last step of our proof is $2 \Rightarrow 3$. We will prove it by contradiction. Let A be unbounded. Then

$$\exists x_n : \ \|x_n\| = 1 \text{ such that } \|Ax_n\| \geqslant n.$$

Define

$$y_n := \frac{x_n}{n}, \quad \|y_n\| = \frac{1}{n} \to 0,$$

so $y_n \to 0$, but $\|Ay_n\| \geqslant 1$, which is contradiction to the continuity at 0. □

One can pose the question: when is the space of bounded operators complete? The answer to this question is provided by the following theorem:

Theorem 7.2. *Let X and Y be normed spaces, and Y be Banach. Then $B(X,Y)$ is Banach.*

Proof. Let us consider a Cauchy sequence $\{A_n\}_{n=1}^{\infty}$ in $B(X,Y)$. By definition, this means that

$$\forall \varepsilon \ \exists N = N(\varepsilon) : \ \forall n, m \geqslant N \ \|A_n - A_m\| < \varepsilon,$$

and since the norm in the space of operators is given by supremum, the following is also true:

$$\forall x \in X : \ \|A_n x - A_m x\| < \varepsilon \|x\|.$$

Thus, $\{A_n x\}_{n=1}^{\infty}$ is a Cauchy sequence in Y. Therefore, there exists a limit; the limit preserves linear operations, so one can define an operator

$$\exists \lim_{n \to \infty} A_n x =: Ax.$$

Existence of this limit means the pointwise convergence $A_n \to A$. Then, in the written above

$$\forall n, n \geqslant N \ \|A_n x - A_m x\| < \varepsilon \|x\|$$

take the limit as $m \to \infty$:

$$\|A_n x - Ax\| \leqslant \varepsilon \|x\|,$$

and, taking the supremum over the unit sphere in X, we get

$$\|A_n - A\| \leqslant \varepsilon,$$

so $A_n \to A$ converges uniformly.

It is easy to see that A is a bounded operator:

$$\|A\| = \|A - A_n + A_n\| \leqslant \|A - A_n\| + \|A_n\|;$$

the first summand is less than ε for $n \geqslant N$, and the second one is bounded $\forall n$, therefore, $\|A\| < \infty$. □

Linear Functionals and Dual Spaces

One of the benefits of the previous theorem is that the space of operators from X to the field is complete:

Definition 7.4. *Let X be a normed space over a field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. $B(X,\mathbb{K}) =: X^*$ is called a **dual space** to X. An element $f \in X^*$ is called a **functional**. The norm is X^* is given by*

$$\|f\| = \sup_{\|x\| \leqslant 1} |f(x)| = \sup_{\|x\| \neq 0} \frac{|f(x)|}{\|x\|}.$$

The corollary from the previous theorem:

Corollary 7.1. *X^* is Banach for any normed space X.*

Now we will describe the dual spaces to some specific normed spaces.

Theorem 7.3. *$c_0^* \cong \ell_1$ (here $\cong$ stands for the isometric isomorphism).*

Remark 7.2. *What does it mean? For any $f \in c_0^*$, we have a unique $y \in \ell_1$ corresponding to f, and the formula for the action of the function f on x is the following:*

$$f(x) = \sum_{k=1}^{\infty} x_k y_k,$$

moreover, $\|f\|_{c_0^*} = \|y\|_{\ell_1}$.

Proof. Let $y \in \ell_1$. We will construct a functional $f_y(x)$ such that

$$f_y(x) = \sum_{k=1}^{\infty} x_k y_k.$$

The functional is obviously linear as the sum is linear. Now let us find the bound for $|f_y(x)|$:

$$|f_y(x)| \leqslant \sum_{k=1}^{\infty} |x_k||y_k| \leqslant \sup_{k\geqslant 1}|x_k| \cdot \sum_{k=1}^{\infty} |y_k|,$$

where the first component is just the norm of x in c_0, i.e., $\|x\|_{c_0}$ and the second one is $\|y\|_{\ell_1}$, thus,

$$\|f_y\|_{c_0^*} \leqslant \|y\|_{\ell_1}.$$

Consider $x^n := (\operatorname{sgn} y_1, \operatorname{sgn} y_2, \ldots, \operatorname{sgn} y_n, 0, 0, \ldots) \in c_0$ with an obvious inequality for the norm: $\|x^n\| \leqslant 1$. For such a sequence,

$$|f_y(x^n)| = \sum_{k=1}^{n} |y_k| \to \sum_{k=1}^{\infty} |y_k| \quad \text{as} \quad n \to \infty,$$

so $\|f_y\|_{c_0^*} = \|y\|_{\ell_1}$.

Now we should start from the functional and provide an element of ℓ_1. Let $f \in c_0^*$. We know that $e_k = (0, 0, \ldots, 0, \overset{k}{1}, 0, \ldots)$ is a basis in c_0. Define

$$y_k := f(e_k).$$

If we take $x = (x_1, x_2, \ldots, x_n, \ldots) \in c_0$, we know that

$$\sum_{k=1}^{n} x_k e_k \to x.$$

f is continuous, therefore,

$$f\Big(\sum_{k=1}^{n} x_k e_k\Big) \to f(x);$$

the functional is linear, so, by the definition of y_k,

$$f(\sum_{k=1}^{n} x_k e_k) = \sum_{k=1}^{n} x_k y_k,$$

where $\sum_{k=1}^{n} x_k y_k \to \sum_{k=1}^{\infty} x_k y_k$, so

$$f(x) = \sum_{k=1}^{\infty} x_k y_k.$$

Why $y \in \ell_1$? We know that $\|f\| < \infty$, where

$$\|f\| = \sup_{\|x\|\leqslant 1} |f(x)| \geqslant |f((\operatorname{sgn} y_1, \operatorname{sgn} y_2, \ldots, \operatorname{sgn} y_n, 0, 0, \ldots))| = \sum_{k=1}^{n} |y_k| \ \forall n \in \mathbb{N}.$$

Taking the limit as $n \to \infty$,

$$\|f\| \geqslant \sum_{k=1}^{\infty} |y_k| \quad \Rightarrow \quad y \in \ell_1.$$

In the first step of the proof, we showed that $\|f_y\| = \|y\|$. □

Consider the following example:

Example 7.1. *Find the norm of the functional in c_0:*

$$f(x) = \sum_{k=1}^{\infty} \frac{x_k}{2^k}, \quad \|f\| - ?$$

Here $y_k = 1/2^k$, so

$$\|f\| = \sum_{k=1}^{\infty} \frac{1}{2^k} = 1.$$

Now we formulate the theorem on the structure of the dual space to ℓ_p.

Theorem 7.4. *Let $1 \leqslant p < \infty$. Then $\ell_p^* \cong \ell_q$, where $\frac{1}{p} + \frac{1}{q} = 1$.*

Remark 7.3. *This means that there is a one-to-one correspondence between $f \in \ell_p^*$ and $y = (y_1, \ldots, y_k, \ldots) \in \ell_q$ such that*

$$\forall x \in \ell p: \; f(x) = \sum_{k=1}^{\infty} x_k y_k \quad \textit{and} \quad \|f\|_{\ell_p^*} = \|y\|_{\ell_q}.$$

Proof. The scheme is the same as in the previous theorem. Take $y \in \ell_q$ and construct a functional

$$f_y(x) = \sum_{k=1}^{\infty} x_k y_k \quad \text{for any} \quad x \in \ell_p.$$

First, we estimate the absolute value

$$|f_y(x)| = \Big|\sum_{k=1}^{\infty} x_k y_k\Big| \leqslant \sum_{k=1}^{\infty} |x_k y_k|$$

For this sum, we use the Hölder inequality:

$$\sum_{k=1}^{\infty} |x_k y_k| \leqslant \Big(\sum_{k=1}^{\infty} |x_k|^p\Big)^{1/p} \Big(\sum_{k=1}^{\infty} |y_k|^q\Big)^{1/q}$$

for $1 < p < \infty$, and

$$\sum_{k=1}^{\infty} |x_k y_k| \leqslant \sup_{k \geqslant 1} |y_k| \sum_{k=1}^{\infty} |x_k|^p$$

for $p = 1$. In both cases, we obtain

$$|f_y(x)| \leqslant \|x\|_{\ell_p} \cdot \|y\|_{\ell_q}.$$

It is known that the Hölder inequality is sharp; if $1 < p < \infty$, take

$$x_k = |y_k|^{q-1} \operatorname{sgn} y_k.$$

Since

$$\frac{1}{p} + \frac{1}{q} = 1 \;\Rightarrow\; q = p(q-1),$$

$|x_k|^p = |y_k|^q$, thus, $x \in \ell_p$, and for the functional we obtain

$$f_y(x) = \sum_{k=1}^{\infty} |y_k|^q,$$

and, therefore,

$$\frac{|f_y(x)|}{\|x\|} = \frac{\sum_{k=1}^{\infty} |y_k|^q}{\left(\sum_{k=1}^{\infty} |y_k|^q\right)^{1/p}} = \left(\sum_{k=1}^{\infty} |y_k|^q\right)^{1/q},$$

which means that $\|f_y\|_{\ell_p^*} = \|y\|_{\ell_q}$. For $p = 1$, $q = \infty$, the norm in ℓ_q is given by

$$\|y\| = \sup_{k \geqslant 1} |y_k|,$$

so there are two possibilities:

a) $\exists k_0$: $|y_{k_0}| = \|y\|$. Then we take

$$x = (0, 0, \dots, 0, \overset{k_0}{\operatorname{sgn}} y_{k_0}, 0, \dots), \quad \|x\|_{\ell_p} = 1,$$

and $f(y) = |y_{k_0}| = \|y\|_{\ell_\infty}$.

b) $\exists k_n$ such that $|y_{k_n}| \to \|y\|$. Then take

$$x^n = (0, 0, \dots, \overset{k_n}{\operatorname{sgn}} y_n, 0, \dots), \quad \|x^n\|_{\ell_p} = 1,$$

and then $f(x^n) = |y_{k_n}| \to \|y\|_{\ell_\infty}$ as $n \to \infty$, therefore, $\|f_y\| = \|y\|_{\ell_\infty}$.

Now we take a functional $f \in \ell_p^*$ and construct an element $y \in \ell_q$. We know that

$$e_k = (0, 0, \dots, \overset{k}{1}, 0, \dots), \quad k \in \mathbb{N},$$

is a basis in ℓ_p, $1 \leqslant p < \infty$. Then,

$$\forall x = (x_1, x_2, \dots) \in \ell_p : \quad x = \sum_{k=1}^{\infty} x_k e_k,$$

and the partial sum converges to this element:

$$\sum_{k=1}^{\infty} x_k e_k \to x \quad \text{as} \quad n \to \infty.$$

Define

$$y_k := f(e_k), \quad \text{thus,} \quad f\Big(\sum_{k=1}^{n} x_k e_k\Big) \to f(x),$$

where the left-hand side, by linearity, is a partial sum of the form

$$f\Big(\sum_{k=1}^{n} x_k e_k\Big) = \sum_{k=1}^{n} x_k y_k \to f\Big(\sum_{k=1}^{\infty} x_k e_k\Big) \quad \text{as} \quad n \to \infty.$$

Why $y \in \ell_q$? Again, there are two possibilities:

1) $1 < p < \infty$. The functional is bounded, i.e.,

$$\|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|} < \infty;$$

we consider a nonzero functional $f \neq 0$, so, obviously $y \neq 0$ as well. Take

$$x^n = (x_1, x_2, \dots, x_n, 0, \dots), \quad \text{where} \quad x_k = |y_k|^{q-1} \operatorname{sgn} y_k, \quad k = 1, 2, \dots, n.$$

Continue the estimation:

$$\sup_{x \neq 0} \frac{|f(x)|}{\|x\|} \geqslant \frac{|f(x^n)|}{\|x^n\|} = \frac{\sum_{k=1}^{n} |y_k|^q}{\left(\sum_{k=1}^{n} |y_k|^q\right)^{1/p}} = \Big(\sum_{k=1}^{n} |y_k|^q\Big) \quad \forall n \in \mathbb{N}.$$

Taking the limit as $n \to \infty$, we obtain

$$\|f\| \geqslant \Big(\sum_{k=1}^{n} |y_k|^q\Big) \ \Rightarrow \ y \in \ell_q.$$

2) $p = 1$. In this case, we must show that the sequence y is bounded, i.e., belongs to ℓ_∞. Take

$$x^n = (0, 0, \dots, 0, \overset{n}{\operatorname{sgn}} y_n, 0, \dots), \quad \|x^n\|_{\ell_1} \leqslant 1,$$

and $f(x^n) = |y_n|$. Since $|f(x^n)| \leqslant \|f\|$,

$$\forall n: \ |y_n| \leqslant \|f\| \ \Rightarrow \ y \in \ell_\infty.$$

Thus, for $f = f_y$, from the previous step of the proof, we have $\|f_y\|_{\ell_p^*} = \|y\|_{\ell_q}$. □

Corollary 7.2. *All spaces ℓ_p, $1 \leqslant p \leqslant \infty$, are complete.*

The following theorem, a more general one, claims that the structure of the dual spaces of L_p is similar. We will provide it without a proof:

Theorem 7.5. *Let (Ω, M, μ) be a measurable space, where μ stands for a σ-additive measure σ-finite measure, and $1 \leqslant p < \infty$. Then*

$$\Big(L_p(\Omega,\mu)\Big)^* \cong L_q(\Omega,\mu), \quad \frac{1}{p}+\frac{1}{q}=1,$$

where $\cong$ denotes the isometric isomorphism:

$$\Big(L_p(\Omega,\mu)\Big)^* \ni G \leftrightarrow g \in L_q(\Omega,\mu)$$

such that

$$\forall f \in L_p: \quad G(f)=\int_\Omega f(x)g(x)\,d\mu, \quad \|G\|_{L_p^*}=\|g\|_{L_q}.$$

Lecture 8. Linear Operators and Functionals in Normed Spaces: Exercises

Discussion of Self-Study Exercises from the Previous Lecture

We begin with a discussion of the homework from Lecture 6.

4) a) $M_1 = \left\{ f \in C^1[a,b]: \ |f(a)| \leqslant c_1 \text{ and } \int_a^b |f'(x)|\,dx \leqslant c_2 \right\}$ is not precompact. An example can be provided by $f_n(x) = x^n$ in $C[0,1]$ or

$$f_n = \left(\frac{x-a}{b-a}\right)^n$$

in $C[a,b]$. Since $f_n(0) = 0$, $\exists c_1$: $|f_n(0)| \leqslant c_1$. These functions are monotonic, so

$$\int_0^1 f_n'(x)\,dx = f_n(1) - f_n(0) = 1 \quad \Rightarrow \quad \exists c_2 : \int_0^1 |f_n'(x)|\,dx \leqslant c_2.$$

Thus, both conditions hold, but $\{f_n\}_{n=1}^{\infty}$ is not an equicontinuous family.

b) $M_2 = \left\{ f \in C^1[a,b]: \ |f(a)| \leqslant c_1 \text{ and } \int_a^b |f'(x)|^2\,dx \leqslant c_2 \right\}$. To find out whether this set is precompact in $C[a,b]$ or not, we must study its boundedness and equicontinuity. We know that

$$f(x) = \int_a^x f'(t)\,dt + f(a),$$

therefore,

$$|f(x)| \leqslant \int_a^x |f'(t)|\,dt + |f(a)| \leqslant \int_a^b |f'(t)|\,dt + c_1,$$

for which one can apply the Hölder or Cauchy–Bunyakovsky–Schwarz inequality:

$$\int_a^b |f'(t)|\,dt + c_1 \leqslant \left(\int_a^b |f'(t)|^2\,dt\right)^{1/2} \sqrt{b-a} + c_1 \leqslant \sqrt{c_2}\sqrt{b-a} + c_1,$$

so M_2 is bounded. Now check its equicontinuity. Let $|x-y| < \delta$. We know that

$$\Big|f(x) - f(y)\Big| = \left|\int_x^y f'(t)\,dt\right| \leqslant \left|\int_y^x |f'(t)|\,dt\right|,$$

to which we apply the Cauchy–Bunyakovsky–Schwarz inequality:

$$\left|\int_y^x |f'(t)|\,dt\right| \leqslant \left|\int_y^x |f'(t)|^2\,dt\right|^{1/2} \sqrt{|x-y|} \leqslant \sqrt{c_2}\sqrt{|x-y|},$$

so the functions in M_2 form an equicontinuous family, therefore, M_2 is precompact.

c) $M_3 = \left\{ f \in C^1[a,b] : \int_a^b \left(|f(x)|^2 + |f'(x)|^2 \right) dx \leqslant c \right\}$. One can show that $M_3 \subset M_2$ for some c_1, c_2. Let us do so. By the Newton–Leibniz formula,

$$f(x) = \int_a^x f'(t)\,dt + f(a),$$

or, rearranging it,

$$f(a) = \int_a^x f'(t)\,dt - f(x) \quad \Rightarrow \quad |f(a)| \leqslant \int_a^x |f'(t)|\,dt + |f(x)| \leqslant \int_a^b |f'(t)|\,dt + |f(x)|.$$

Integrating this inequality over $[a,b]$, we obtain

$$(b-a)|f(a)| \leqslant (b-a) \int_a^b f'(t)\,dt + \int_a^b |f(x)|\,dx,$$

and then, using the Cauchy–Bunyakovsky–Schwarz inequality,

$$(b-a) \int_a^b f'(t)\,dt + \int_a^b |f(x)|\,dx \leqslant \sqrt{b-a}\sqrt{b-a} \Big(\int_a^b |f'(t)|^2\,dt \Big)^{1/2} + \\ + \sqrt{b-a} \Big(\int_a^b |f(x)|\,dx \Big)^{1/2},$$

so $f(a)$ is bounded:

$$|f(a)| \leqslant \sqrt{b-a}\sqrt{c} + \frac{\sqrt{c}}{\sqrt{b-a}} =: c_1.$$

Now we must show that the derivative is bounded in the L_2-sense. By definition of M_3, we have

$$\int_a^b \left(|f(x)|^2 + |f'(x)|^2 \right) dx \leqslant c,$$

therefore,

$$\int_a^b |f'(x)|^2\,dx \leqslant c,$$

so $M_3 \subset M_2$ for c_1 as defined above, and $c_2 = c$, thus, M_3 is precompact.

5) Prove that the unit ball $B[0,1] \subset L_2[0,1]$ is not precompact in $L_1[0,1]$.

First, we show that $L_2[0,1] \subset L_1[0,1]$. By the Cauchy–Bunyakovsky–Schwarz inequality,

$$\int_0^1 1 \cdot |f(t)|\,dt \leqslant \Big(\int_0^1 |f(t)|^2\,dt \Big)^{1/2} \cdot \Big(\int_0^1 1\,dt \Big)^{1/2} = \|f\|_{L_2},$$

therefore, $f \in L_2[0,1] \Rightarrow f \in L_1[0,1]$.

Now, for $n = 1$, consider

$$f_1 = \chi_{\left[0,\frac{1}{2}\right]} - \chi_{\left[\frac{1}{2},1\right]},$$

see Fig. 8.1.

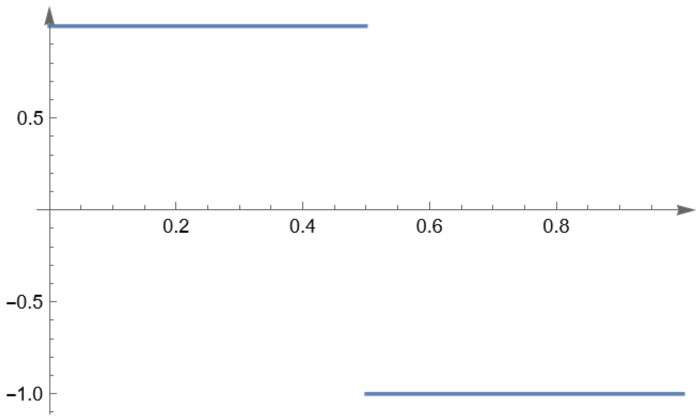

Рис. 8.1. Graph of f_1.

For an arbitrary n, we divide the interval $[0,1]$ into pieces of length $1/2^n$, where the values 1 and -1 alternate for $f_n(x)$, i.e.,

$$f_n = \chi_{\left[0,\frac{1}{2^n}\right]} - \chi_{\left[\frac{1}{2^n},\frac{2}{2^n}\right]} + \chi_{\left[\frac{2}{2^n},\frac{3}{2^n}\right]} - \chi_{\left[\frac{3}{2^n},\frac{4}{2^n}\right]} + \dots,$$

see an example in Fig. 8.2.

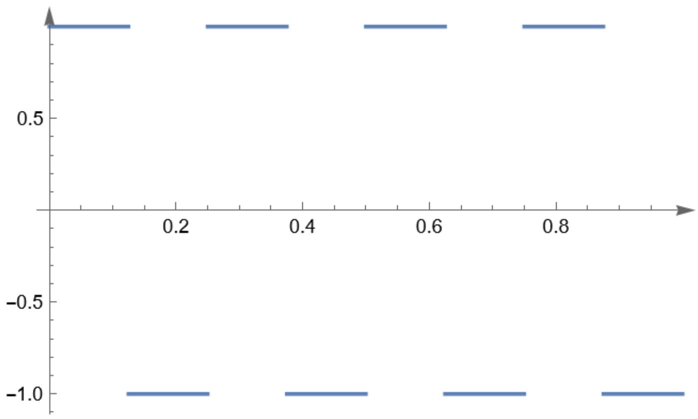

Рис. 8.2. Graph of f_n, $n = 3$.

What can we say about the norm of these functions in L_2 and in L_1?

$$\|f_n\|_{L_2[0,1]} = \|f_n\|_{L_1[0,1]} = 1,$$

since the absolute value of $f_n(x)$ equals 1 identically. Now, consider $\|f_n - f_m\|_{L_1[0,1]}$, see an example in Fig. 8.3.

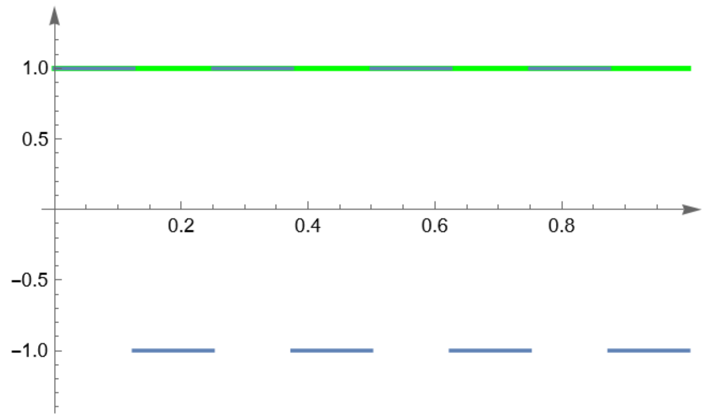

Рис. 8.3. Graphs of f_0 (green) and f_3 (blue).

One can see that $\|f_n - f_m\|_{L_1} = 1$, since half the length of the interval these functions coincide, so the difference is 0, while in the other half, they differ by 2. Thus, there is no Cauchy subsequence of f_n.

Exercises on Bounded Operators and Functionals

Now, we discuss some examples of bounded operators and functionals and consider some exercises.

1) Take $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n, \dots) \in \ell_\infty$, and define

$$A_\alpha x = (\alpha_1 x_1, \alpha_2 x_2, \dots, \alpha_n x_n, \dots) \quad \text{in} \quad \ell_2.$$

Find the norm $\|A_\alpha\|$.

Since we are in ℓ_2, it is convenient to write the squared norm. By definition,

$$\|A_\alpha x\|^2 = \sum_{k=1}^{\infty} |\alpha_k x_k|^2.$$

From this sum, one can take out the supremum of α_k:

$$\sum_{k=1}^{\infty} |\alpha_k x_k|^2 \leqslant \sup_{k\geqslant 1} |\alpha_k|^2 \sum_{k=1}^{\infty} |x_k|^2 = \|\alpha\|_{\ell_\infty}^2 \|x\|_{\ell^2}^2.$$

Thus, we obtained an upper bound for the norm of the operator:

$$\|A_\alpha\| \leqslant \|\alpha\|_{\ell_\infty}.$$

There are two possibilities:

a) $\exists k_0$: $|\alpha_{k_0} = \sup\limits_{k\geqslant 1}|\alpha_k|$. Then we take

$$x = (0, 0, \dots, \overset{k_0}{\operatorname{sgn}\alpha_{k_0}}, 0, \dots).$$

For this x,

$$\|Ax\| = |\alpha_{k_0}| \equiv \|\alpha\|_{\ell_\infty}.$$

b) $\nexists k_0$, but $\exists k_n \to \infty$:

$$|\alpha_{k_n}| \to \sup_{k\geqslant 1}|\alpha_k|.$$

In this case, we consider a sequence

$$x^n = (0, \dots, 0, \overset{k_n}{\operatorname{sgn} k_n}, 0, \dots) \in \ell_2,$$

so

$$\|A_\alpha x^n\| = |\alpha_{k_n}| \to \|\alpha\|_{\ell_\infty},$$

therefore, $\|A_\alpha\| = \|\alpha\|_{\ell_\infty}$.

2) In $C[-1,1]$, consider the functional F such that

$$\forall f \in C[-1,1]: \quad F(f) = \int_{-1}^{0} f(t)\,dt - \int_{0}^{1} f(t)\,dt.$$

Find the norm $\|F\|$.

We begin with the estimation

$$|F(f)| = \Big|\int_{-1}^{0} f(t)\,dt - \int_{0}^{1} f(t)\,dt\Big| \leqslant \int_{-1}^{0} |f(t)|\,dt + \int_{0}^{1} |f(t)|\,dt. \tag{8.1}$$

In $C[a,b]$, we have a very useful inequality:

$$\forall t \in [a,b]: \quad |f(t)| \leqslant \|f\|_{C[a,b]} = \max_{[a,b]}|f(x)|.$$

Using this, we conclude that each of the integrals on the right-hand side is bounded from above by $\|f\|$, so

$$|F(f)| \leqslant 2\|f\|_{C[a,b]}, \quad thus, \quad \|F\| \leqslant 2.$$

For what function can equality be achieved? If we take f_0 such that

$$f_0(x) = \begin{cases} 1, & x \in [-1, 0], \\ -1, & x \in (0, 1], \end{cases}$$

then we obtain the equality $F(f_0) = 2$. The problem here is that f_0 of the given form does not belong to $C[-1,1]$. We can approximate it by a sequence of continuous function taking a small neighborhood of zero for a linear function gluing the values together, for instance, consider the sequence of functions

$$f_n(x) = \begin{cases} 1, & x \in \left[-1, -\frac{1}{n}\right], \\ -nx, & x \in \left[-\frac{1}{n}, \frac{1}{n}\right], \\ -1, & x \in \left[\frac{1}{n}, 1\right], \end{cases}$$

see an example in Fig. 8.4.

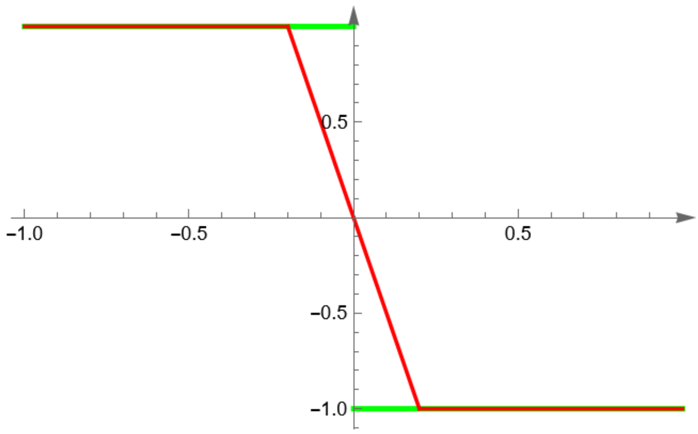

Рис. 8.4. Graphs of f_0 (green) and f_5 (red).

Obviously, $f_n \in C[-1,1]$, $\|f_n\| = 1$, and $f_n \to f$. The functional evaluated at this element gives $F(f_n) = 2 - 1/n \to 2$ as $n \to \infty$. Thus, its norm is indeed equal to 2.

3) Consider in ℓ_2 the operators of right and left shifts:

$$A_r x = (0, x_1, x_2, \ldots), \quad A_\ell = (x_2, x_3, x_4, \ldots).$$

What can be said about the norms of these operators?

These operators are closely related to the creation and annihilation operators that arise in Quantum Mechanics; usually, these operators are considered in two-sided ℓ_2.

It is clear that

$$\forall x: \quad \|A_r x\| = \|x\|, \quad \|A_\ell x\| \leqslant \|x\|.$$

For A_r, we immediately obtain $\|A_r\| = 1$. For A_ℓ, this only guarantees the bound $\|A_\ell\| \leqslant 1$. One can take the second basis vector $e_2 = (0,1,0,0,\dots)$, and, applying the operator, get that

$$A_\ell e_2 = e_1,$$

therefore, $\|A_\ell\| = 1$.

Consider these operators in $\ell_2(\mathbb{Z})$:

$$\ell_2(\mathbb{Z}) \ni x = (\dots, x_{-2}, x_{-1}, (x_0), x_1, x_2, \dots).$$

By taking an element to the brackets, we point out that it is the center of the sequence. $\ell_2(\mathbb{Z})$ is a Hilbert space with the norm and the inner product defined by

$$\|x\| = \Big(\sum_{k=-\infty}^{\infty} |x_k|^2\Big)^{1/2}, \quad (x,y) = \sum_{k=-\infty}^{\infty} x_k \overline{y_k}.$$

In this space, $\|A_r\| = \|A_\ell\| = 1$:

$$A_r x = (\dots, x_{-2}, (x_{-1}), x_0, \dots), \qquad A_\ell x = (\dots, x_0, (x_1), x_2, \dots),$$

so these two are examples of the unitary operators.

4) Let $g \in C[a,b]$ be some certain function. Consider the functional F_g in $C[a,b]$ defined by the formula

$$F_g(f) = \int_a^b f(x)g(x)\,dx \quad \forall f \in C[a,b].$$

Evedently, it is a linear functional. What is the norm of F_g?

First, we will provide a bound for $\forall f \in C[a,b]$ in terms of $\|f\|_{C[a,b]}$:

$$\Big|F_g(f)\Big| \leqslant \Big|\int_a^b f(x)g(x)\,dx\Big| \leqslant \int_a^b |f(x)g(x)|\,dx;$$

the following step is quite simple, we just take out the norm of f:

$$\int_a^b |f(x)g(x)|\,dx \leqslant \max_{[a,b]}|f(x)| \int_a^b |g(x)|\,dx.$$

The conjecture is that $\|F_g\| = \|g\|_{L_1}$. For $f(x) = \operatorname{sgn} g(x)$, $F_g(f) = \int_a^b |g(x)|\,dx$, but $f(x) \notin C[a,b]$. Even though, one can approximate it by a continuous family, for example, as in the following. Let $\varepsilon > 0$. Consider

$$f_\varepsilon(x) = \begin{cases} \varepsilon \operatorname{sgn} g(x), & \text{if} \quad |g(x)| \geqslant \varepsilon, \\ g(x), & \text{if} \quad |g(x)| < \varepsilon. \end{cases}$$

It is a continuous function, and $\|f_\varepsilon\|_{C[a,b]} = \varepsilon$ (if $g \not\equiv 0$). Consider

$$\widetilde{f}_\varepsilon(x) \equiv \frac{f_\varepsilon(x)}{\varepsilon} = \begin{cases} \operatorname{sgn} g(x), & \text{if} \quad |g(x)| \geqslant \varepsilon, \\ \dfrac{g(x)}{\varepsilon}, & \text{if} \quad |g(x)| < \varepsilon; \end{cases}$$

obviously, $\|\widetilde{f}_\varepsilon\| = 1$. Now evaluate the functional F_g at this function:

$$F_g(\widetilde{f}_\varepsilon) = \int_a^b \widetilde{f}_\varepsilon(x) g(x)\, dx = \int_{x\colon |g(x)| \geqslant \varepsilon} |g(x)|\, dx + \frac{1}{\varepsilon} \int_{x\colon |g(x)| < \varepsilon} g^2(x)\, dx.$$

Since the integrand of the second integral is positive, we can bound the sum from below by the first integral, that is,

$$F_g(\widetilde{f}_0) \geqslant \int_{x\colon |g(x)| \geqslant \varepsilon} |g(x)|\, dx.$$

Taking the limit as $\varepsilon \to 0$, we come to

$$F_g(\widetilde{f}_\varepsilon) \geqslant \int_a^b |g(x)|\, dx,$$

therefore,

$$\|F_g\|_{(C[a,b])^*} = \int_a^b |g(x)|\, dx \equiv \|g\|_{L_1[a,b]}.$$

Another way to find the norm of this functional is following. First, give a uniform approximation of g with polynomials p_n, using the Weierstrass approximation theorem. Second, approximate the sign of the polynomial p_n by a continuous function f_n, and evaluate the functional F_g at the function f_n.

5) Consider a functional f in c (recall that this is the space of converging sequences:

$$x = (x_1, x_2, \dots, x_n, \dots) \in c \quad \Leftrightarrow \quad \exists \lim_{n\to\infty} x_n = a,$$

where $a = a(x)$, and $\|x\| = \sup\limits_{k \geqslant 1} |x_k|$):

$$f(x) = \sum_{k=1}^{\infty} \frac{(-1)^k x_k}{2^k}.$$

Find the norm $\|f\|$.

Once again, first we estimate the functional in terms of $\|x\|$:

$$|f(x)| = \Big| \sum_{k=1}^{\infty} \frac{(-1)^k x_k}{2^k} \Big| \leqslant \sum_{k=1}^{\infty} \frac{|x_k|}{2^k} \leqslant \sup_{k \geqslant 1} |x_k| \cdot \sum_{k=1}^{\infty} \frac{1}{2^k} = \|x\|_c. \tag{8.2}$$

We have obtained that $\|f\| \leqslant 1$. The natural conjecture is that $\|f\| = 1$. If we analyze the first inequality in (8.2), that is,

$$\Big|\sum_{k=1}^{\infty} \frac{(-1)^k x_k}{2^k}\Big| \leqslant \sum_{k=1}^{\infty} \frac{|x_k|}{2^k},$$

we see that the equality is achieved for

$$x = (-1, 1, -1, \dots, (-1)^n, \dots),$$

which is not an element of c. One can take a sequence

$$x^n = (-1, 1, -1, 1, \dots, -1, 1, 0, 0, \dots) \in c_0 \subset c, \quad \|x^n\| = 1,$$

which has $2n$ nonzero coordinates. Evaluating the functional at this sequence, we get

$$f(x) = \sum_{k=1}^{2n} \frac{1}{2^k} \to 1 \quad \text{as} \quad n \to \infty,$$

therefore, $\|f\| = 1$.

6) Consider an operator

$$(Af)(x) = \int_a^n K(x,t) f(t)\, dt;$$

the function $K(x,t)$ is called an **integral kernel** of the operator A.

Let $K(x,t) \in C[a,b]^2$. Consider this operator on the space $C[a,b]$:

$$A : C[a,b] \to C[a,b].$$

Note that this is a continuous analog of the matrix operator. What does it mean?

Let $A = (a_{ij})_{i,j=1}^n$, $x = (x_1, x_2, \dots, x_n)$. Then

$$(Ax)_j = \sum_{i=1}^{n} a_{ij} x_i.$$

Replacing $j \to t$, $a_{ij} \to K(x,t)$, and $\sum \to \int$, we obtain $K(x,t) f(t)\, dt$.

Now, find the norm of A.

First, we would like to obtain a bound for Af in terms of $\|f\|$:

$$\|Af\| = \max_{[a,b]} \Big| \int_a^b K(x,t) f(t)\, dt \Big| \leqslant \max_{[a,b]} \int_a^b |K(x,t) f(t)|\, dt \leqslant \max_{[a,b]} |f(t)| \cdot \int_a^b |K(x,t)|\, dt.$$

Our conjecture is that $\|A\| = \int_a^b |K(x,t)|\, dt$.

We know that the function $K(x,t)$ is continuous; therefore, $\int_a^b |K(x,t)|\,dt$ is continuous. Therefore,

$$\exists x_0 \in [a,b]: \quad \int_a^b |K(x_0,t)|\,dt = \max_{x\in[a,b]} \int_a^b |K(x,t)|\,dt.$$

Consider problem 4 with $g(t) = K(x_0,t) \in C[a,b]$, where we have constructed $\widetilde{f}_\varepsilon$:

$$F(\widetilde{f}_\varepsilon) \to \int_a^b |g(t)|\,dt.$$

Now, take the family $\widetilde{f}_\varepsilon$ from problem 4 for the function $g(t) = K(x_0,t)$. Then

$$\|A\|_{C[a,b]\to C[a,b]} = \max_{x\in[a,b]} \int_a^b |K(x,t)|\,dt.$$

Self-Study Exercises

1) Show that $c^* \cong \ell_1 \oplus \mathbb{C}$ $(\cong \ell_1)$. The symbol $\cong$ stands for the isometric isomorphism

$$c^* \ni f \leftrightarrow (y,\alpha), \quad ,y \in \ell_1, \quad \alpha \in \mathbb{C},$$

and

$$f(x) = \alpha x_0 + \sum_{k=1}^{\infty} x_k y_k, \quad \|f\| = |\alpha| + \sum_{k=1}^{\infty} |y_k|.$$

2) Consider in ℓ_3 the functional

$$f(x) = \sum_{k=1}^{\infty} \frac{x_k}{k^{4/3}}.$$

Find the norm $\|f\|$.

3) In $C[-1,1]$, consider the functional

$$F(f) = \int_{-1}^{1} |x| f(x)\,dx + 2f\Big(-\frac{1}{2}\Big) - f\Big(\frac{1}{4}\Big).$$

Find the norm $\|F\|$.

4) Consider

$$(Af)(x) = \int_a^b K(x,t) f(t)\,dt,$$

a) $K(x,t) \in C[a,b]^2$, $A: L_1[a,b] \to L_1[a,b]$. Find the norm $\|A\|$.

b) $K(x,t) \in C[a,b]^2$, $A: L_1[a,b] \to C[a,b]$. Find the norm $\|A\|$.

c) $K(x,t) \in L_2[a,b]^2$, $A: L_2[a,b] \to L_2[a,b]$. Find the bound C for the norm: $\|A\| \leqslant C$.

5) Consider an operator

$$(Af)(x) = \int_0^x f(t)\,dt$$

a) in $C[0,1]$: find the norm $\|A\|$.

b) in $L_2[0,1]$: find a sharp bound C, $\|A\| \leqslant C$.

Lecture 9. The Hahn–Banach Theorem and the Corollaries

The Hahn–Banach Theorem

Theorem 9.1 (Hahn–Banach). *Let X be a linear space over a field $\mathbb{K}$ ($\mathbb{R}$ or $\mathbb{C}$), and $p : X \to [0,+\infty)$ be a seminorm. Let X_0 be a nontrivial subspace, and f_0 be a linear functional on X_0 such that*

$$\forall x \in X_0 : \quad |f_0(x)| \leqslant p(x).$$

Then there exists a linear functional f on X such that

$$f\Big|_{X_0} = f_0, \qquad \forall x \in X : \quad |f(x)| \leqslant p(x).$$

It is a general formulation of this theorem. For us, it will be convenient to use a particular case, formulating the theorem for a normed space.

Theorem 9.2 (The Hahn–Banach Theorem for normed space). *Let X be a normed space over a field $\mathbb{K}$ ($\mathbb{R}$ or $\mathbb{C}$), X_0 be a nontrivial subspace, and $f_0 \in X_0^*$. Then there exists a linear functional $f \in X^*$ such that*

$$f\Big|_{X_0} = f_0, \qquad \|f\| = \|f_0\|.$$

Remark 9.1. *This theorem is an obvious corollary of the previous one, as one plugs $p(x) = \|f_0\| \cdot \|x\|$.*

Why do we need X_0 to be nontrivial? A trivial subspace is either $\{0\}$ or the entire X. In the first case, $f_0 \equiv 0$, so its extension is zero functional. In the second one, we already have a functional on the entire space, so its extension is $f \equiv f_0$.

For simplicity, we will prove the theorem for a separable space, while it is valid otherwise as well.

Proof.

1) Suppose $\mathbb{K} = \mathbb{R}$. There exists $x_1 \notin X_0$. Consider a subspace

$$X_1 = \langle X_0, x_1 \rangle = \{x = x_0 + t x_1,\ x_0 \in X_0,\ t \in \mathbb{R}\}.$$

We are going to construct an extension of f_0 to this subspace: $f_1 \in X_1^*$ such that

$$f_1(x) = f_0(x_0) + t f_1(x_1) \equiv f_0(x_0) + t \cdot \alpha,$$

where $f_1(x_1) \equiv \alpha \in \mathbb{R}$. We need to verify that $\|f\| = \|f_0\|$; to obtain the same norm, one must choose α appropriately. Let $x', x'' \in X_0$. Consider

$$f_0(x') + f_0(x'') = f_0(x' + x'') \leqslant p(x' + x'') = p(x' - x_1 + x'' + x_1),$$

and use the triangle inequality for p:

$$p(x' - x_1 + x'' + x_1) \leqslant p(x' - x_1) + p(x'' + x_1),$$

so we have

$$f_0(x') + f_0(x'') \leqslant p(x' - x_1) + p(x'' + x_1).$$

Now we rearrange this inequality in such a way that x' is on the left-hand side and x'' is on the right-hand side:

$$f_0(x') - p(x' - x_1) \leqslant -f_0(x'') + p(x'' + x_1) \quad \forall x', x'' \in X_0. \tag{9.1}$$

Now, take the supremum on the left-hand side and the infimum om the right-hand side:

$$A := \sup_{x' \in X_0} \Big(f_0(x') - p(x' - x_1) \Big), \quad B := \inf_{x'' \in X_0} \Big(-f_0(x'') + p(x'' + x_1) \Big).$$

As (9.1) holds *for any* x', x'', for A and B we have $A \leqslant B$. Take $A \leqslant \alpha \leqslant B$. We have to verify that $|f(x)| \leqslant p(x)$. Consider $f_1(x) = f_0(x_0) + t\alpha$; let $t > 0$:

$$f_0(x_0) + t\alpha \leqslant f_0(x_0) + tB;$$

B is an infimum, so, taking the expression under the inf, we will increase the bound; take a specific element: $x'' = x_0/t$. Then

$$f_0(x_0) + tB \leqslant f_0(x_0) + t\Big(-f_0\Big(\frac{x_0}{t}\Big) + p\Big(\frac{x_0}{t} + x_1\Big)\Big);$$

f_0 is a linear functional on X_0, so one can take out $1/t$, which cancels out $f_0(x_0)$:

$$f_0(x_0) + t\Big(-t f_0(x_0) + p\Big(\frac{x_0}{t} + x_1\Big)\Big) = tp\Big(\frac{x_0}{t} + x_1\Big).$$

p is a seminorm, so one can take out a positive number; t is positive, therefore,

$$tp\Big(\frac{x_0}{t} + x_1\Big) = p(x),$$

which means

$$f_1(x) \leqslant p(x) \quad \forall x \in X_1;$$

the same bound can be obtained by taking a minus inside the functional:

$$-f_1(x) = f_1(-x) \leqslant p(-x) = |-1| \cdot p(x) \equiv p(x) \quad \Rightarrow \quad f_1(x) \geqslant -p(x),$$

and, finally,

$$|f_1(x)| \leqslant p(x) \quad \Leftrightarrow \|f_1\| = \|f_0\|.$$

For negative t, the proof is similar with α being replaced by A.

By definition of separability, we have a countable dense subset $\{x_k\}_{k=1}^{\infty}$, $\overline{\{x_k\}_{k=1}^{\infty}} = X$. Thus, one can construct a chain of subspaces

$$X_0 \subsetneq X_1 \subsetneq X_2 \subsetneq \dots \subsetneq X_n \subsetneq \dots$$

by extension with one element each. Without loss of generality, assume that

$$\{x_1, x_2, \dots, x_n\} \subset X_n.$$

Then, by definition of set operations,

$$X_\infty := \cup_{n=1}^{\infty} X_n.$$

This set may not coincide with X, but, since $\{x_k\}_{k=1}^{\infty}$ is dense in X, we have

$$\overline{X}_\infty = X.$$

By induction, one can construct functionals

$$f_2 \in X_2^*, \quad f_3 \in X_3^*, \quad \dots, \quad f_n \in X_n^*, \quad \dots$$

such that $\forall n$: $\|f_n\| = \|f_0\|$; on X_∞, define

$$f_\infty(x) = \lim_{n\to\infty} f_n(x);$$

this functional is well-defined since $\forall x \in X\ \exists n_0$: $x \in X_{n_0}$, so $f_\infty(x) = f_{n_0}(x)$.

For further developments, we need the following auxiliary statement:

Statement 9.1. *Let X, Y be normed spaces, and Y be a Banach space. Let $X_0 \subset X$, $\overline{X}_0 = X$, be a nontrivial subspace, and $A_0 \in B(X_0, Y)$. Then*

$$\exists! A \in B(X,Y): \quad \|A_0\| = \|A\|.$$

Proof. The difference between extensions of operators and functionals is that to define an extension of an operator, one must require that it is defined on a dense subset.

Now, take $x \in X \backslash X_0$; it is a limit point of X_0, therefore,

$$\exists x_n \in X_0 : \quad x_n \to x.$$

Estimate the norm

$$\left\|A_0 x_n - A_0 x_m\right\| \leqslant \|A_0\| \|x_n - x_m\| \to 0 \quad \text{as} \quad n, m \to \infty,$$

i.e., the sequence $A_0 x_n$ is Cauchy along with x_n; Y is a complete space, thus,

$$\exists \lim_{n\to\infty} A_0 x_n.$$

What can we say about this operator? A_0 is linear, $\lim$ preserves linear operations, so this expression depends linearly on x, and one can define

$$Ax := \lim_{n\to\infty} A_0 x_n.$$

It is clear that this construction is well-defined: the sequence $x_n \to x$ is not unique, but if we take $x_n' \to x$, then, by combining the elements of the sequences like

$$x_1,\ x_1',\ x_2,\ x_2',\ \ldots,\ x_n,\ x_n',\ \ldots,$$

we see that

$$A_0 x_1,\ A_0 x_1',\ A_0 x_2,\ A_0 x_2',\ \ldots,\ A_0 x_n,\ A_0 x_n',\ \ldots,$$

is a Cauchy sequence, so the limit is unique. The norm is preserved due to the fact that it is continuous, so

$$\|Ax\| = \underset{n\to\infty}{\|} A_0 x_n\| \leqslant \underset{n\to\infty}{\|} A_0\| \cdot \|x_n\| = \|A_0\| \cdot \|x\|,$$

and since $A\big|_{X_0} = A_0$, the norm is the same. □

Now, let us return to the proof of the Hahn–Banach theorem. We have f_∞ on X_∞, $\overline{X}_\infty = X$, and $\|f_\infty\| = \|f_0\|$. Using the auxiliary statement, we conclude that $\exists! f \in X^*$ with $\|f\| = \|f_\infty\| = \|f_0\|$.

Thus, we completed the proof for real separable Hilbert spaces. What if it is complex?

2) Suppose $\mathbb{K} = \mathbb{C}$. In this case, we proof is based on Linear Algebra. We have $f_0 \in X_0^*$, with complex X_0. Consider a *realification* of X_0: $X_0^{\mathbb{R}}$, i.e., the space where only multiplication by real numbers is allowed. As for the functional, we decompose it into

$$f_0(x) = \operatorname{Re} f_0(x) + i \operatorname{Im} f_0(x) \equiv \varphi_0(x) + i \operatorname{Im} f_0(x).$$

Thus, we have a real functional $\varphi_0(x)$ on a real subspace $X_0^{\mathbb{R}}$, so one can construct an extension $\varphi(x)$ by step 1 on the space $X^{\mathbb{R}}$:

$$\varphi\Big|_{X_0^{\mathbb{R}}} = \varphi_0 \quad \text{and} \quad |\varphi_0(x)| \leqslant |f_0(x)| \leqslant p(x).$$

The imaginary part can be in fact recovered from the real one. Why? We would like to construct a functional

$$f(x) = \varphi(x) + i \operatorname{Im} f(x). \tag{9.2}$$

Recall that in $X_0^{\mathbb{R}}$, there are all the elements of X_0, but we allow multiplication only by real numbers; this means that ix belongs to $X_0^{\mathbb{R}}$ along with x, so, for

$$f(ix) = \varphi(ix) + i \operatorname{Im} f(ix),$$

by linearity, one can take the right-hand side of (9.2) with a factor i, that is,

$$if(x) = i\varphi(x) - \operatorname{Im} f(x),$$

therefore, $\operatorname{Im} f(x) = \varphi(ix)$, and the entire functional takes the form

$$f(x) = \varphi(x) - i\varphi(x).$$

Now, for f, we must check the preserving of the bound. Let $f(x) = re^{i\theta}$. Then $e^{-i\theta} f(x) = f(e^{-i\theta}x)$ is real. Therefore, for non-real $f(x)$, the same bound as for $f(e^{-i\theta}x)$ is valid, where $f(e^{-i\theta}x) = r \in \mathbb{R}$; for this, we obtain

$$|f(x)| = |e^{-i\theta} f(x)| = |f(e^{-i\theta}x)| = |\varphi(x)| \leqslant p(x),$$

which means that $\|f\| = \|f_0\|$. □

Corollaries of the Hahn–Banach Theorem

Why is the Hahn–Banach theorem so important? In fact, for the space L_p with $0 < p < 1$, which is a *quasi-Banach space* (but not a Banach space due to the lack of subadditivity in its quasi-norm), the Hahn–Banach theorem does not apply in its usual form. Here, only the zero functional exists as a continuous linear functional, since L_p, $p < 1$, fails the norm structure required for the extension theorem, highlighting the theorem's necessity for Banach spaces.

The Hahn–Banach theorem is fundamental in Functional Analysis due to its important corollaries as well. We will consider some of them.

Corollary 9.1. *Let X, $X \neq \{0\}$, be a normed space. Then*

$$\forall x \neq 0 \ \exists f_x \in X^* : \quad \|f_x\| = 1, \quad f_x(x) = \|x\|.$$

Proof. Consider $X_0 = \langle x \rangle = \{y = \alpha x,\ \alpha \in \mathbb{C}\}$, and $X_0 \ni f_0 = \alpha\|x\|$. It is obvious that $f_0(x) = \|x\|$. The Hahn–Banach theorem allows one to construct an extension of a *bounded* functional, so we have to check the boundedness of f_0:

$$\frac{|f_0(y)|}{\|y\|} = \frac{|\alpha|\|x\|}{|\alpha|\|x\|} = 1.$$

By the Hahn–Banach theorem, construct an extension f_x of f_0. □

Corollary 9.2. *Let X, $X \neq \{0\}$, be a normed space. Then*

$$\forall x, y \in X,\ x \neq y,\ \Rightarrow \exists f \in X^*,\ \|f\| = 1,\ f(x) \neq f(y).$$

This means that weak topology on the normed space is Hausdorff.

Proof. Consider $z = x - y \neq 0$. By the previous corollary,

$$\exists f_z \in X^*,\ \|f_z\| = 1,\ f_z(z) = \|z\| \neq 0.$$

By the linearity, $0 \neq f_z(z) = f_z(x) - f_z(y)$. □

So we have enough functionals to distinguish the elements of X.

Before formulation of the next corollary, let us look what we have. We have X, a normed space, and

$$X \to X^* \to X^{**}.$$

In the case $\dim X < \infty$, we know that there is a canonical isomorphism

$$X \cong X^{**}.$$

If $\dim X = \infty$, we are only able to construct a canonical embedding $X \hookrightarrow X^{**}$. This means the following. Let $x \in X$, $f \in X^*$, and $F \in X^{**}$. We can take x and associate to it a functional $F_x \in X^{**}$ such that

$$F_x(f) = f(x).$$

In the finite-dimensional case, it is a bijection; in the infinite-dimensional one, this is not the case.

Corollary 9.3. *The canonical embedding $X \hookrightarrow X^{**}$ is an isometry.*

Proof. By the definition of the canonical embedding,

$$\|F_x\| = \sup_{\|f\|=1} |F_x(f)| = \sup_{\|f\|=1} |f(x)| \leqslant \sup_{\|f\|=1} \|f\| \cdot \|x\| = \|x\|.$$

We have an inequality only at a single step; at the other steps, there are equalities. Recall that, by the first corollary, there exists f_x, $\|f_x\| = 1$, such that $f_x(x) = \|x\|$. Therefore, due to this property, this inequality is sharp, and equality can be achieved for f_x. □

Reflexive Spaces

Definition 9.1. *A normed space X is called **reflexive** if the canonical embedding $X \hookrightarrow X^{**}$ is bijection.*

Note that it is sufficient to require that the embedding be a surjection. As we have already learned, it is obviously is injection since it preserves the norm.

Example 9.1. *Consider the following examples:*

1) *All finite-dimensional spaces are reflexive.*

2) $c_0^* \cong \ell_1$, $\ell_p^* \cong \ell_q$ *for* $1 \leqslant p < \infty$, $1/p + 1/q = 1$. *In particular,* $\ell_1^* \cong \ell_\infty$, *so* $c_0^{**} \cong \ell_\infty$, *therefore,* c_0 *is **not** reflexive.*

3) *If* $1 < p < \infty$, *then* $1 < q < \infty$, *and* $\ell_p \cong \ell_p^{**}$ *since* ℓ_p *is dual to* ℓ_q, *and vice versa.*

Corollary 9.4. *Let X be reflexive. Then*

$$\forall f \in X^* \ \exists x \in X : \quad \|x\| = 1 \quad \text{and} \quad f(x) = \|f\|.$$

Proof. The proof requires only corollary 9.1; it claims that

$$\forall f \in X^*, \ f \neq 0 \ \exists F \in X^{**} : \quad \|F\| = 1 \quad \text{and} \quad F(f) = \|f\|.$$

By the definition of reflexive space, to $F = F_x \in X^{**}$, there corresponds $x \in X$, which is in fact the same: $x = F_x$; therefore, $F(f) = f(x)$, which completes the proof. □

Now, we have description for $\ell_p \cong \ell_q$, $L_p^*(\Omega, \mu) = L_q(\Omega, \mu)$, and $c_0^* \cong \ell_1$. We have also the space of continuous functions; it would be nice to describe the dual space to $C[a,b]$ as well.

Dual Space to $C[a,b]$

First, we state the result, and then provide all the necessary constructions.

Theorem 9.3.

$$\Big(C[a,b]\Big)^* \cong BV_0[a,b].$$

Let f be defined on $[a,b]$. Let $T = \{t_k\}_{k=1}^n$ be a partition of $[a,b]$:

$$a = t_0 < t_1 < \cdots < t_n = b.$$

By definition, a **variation** of $f(x)$ on T is

$$V_T f := \sum_{k=1}^{n} |f(t_k) - f(t_{k-1})|.$$

A **total** variation is the supremum with respect to T:

$$V_a^b f := \sup_T V_T f.$$

We say that $f \in BV$ (f is a **function of bounded variation**) on $[a,b]$ if $V_a^b f < \infty$.

For instance, if function f is monotonic, than $V_a^b f = |f(b) - f(a)|$; if $f \in C^1[a,b]$, one can rewrite

$$V_T f = \sum_{k=1}^{n} |f(t_k) - f(t_{k-1})| = \sum_{k=1}^{n} \frac{|f(t_k) - f(t_{k-1})|}{t_k - t_{k-1}} (t_k - t_{k-1}) \to \int_a^b |f'(t)|\, dt,$$

so $f \in C^1[a,b] \Rightarrow f \in BV[a,b]$.

$BV[a,b]$ is a normed and complete space; the norm can be given by

$$\|f\| = V_a^b f + |f(a)|.$$

What is BV_0? It is an additional normalization of functions from BV, which we are to point out:

$$BV_0[a,b] := \{g \in BV[a,b],\ g(a) = 0 \text{ and } \forall x \in (a,b):\ g(x-0) = g(x)\}.$$

Now, we are ready to discuss Theorem 9.3. To any $G \in \Big(C[a,b]\Big)^*$, there corresponds $g \in BV_0[a,b]$ such that the action of G is a Riemann–Stieltjes integral of f of the form

$$G(f) = \int_a^b f(t)\, dg.$$

The Riemann–Stieltjes integral can be represented as

$$\int_a^b f(t)\, dg = \sum_k f(t_k)\big(g(t_k) - g(t_{k-1})\big);$$

a function of bounded variation can be represented as a difference of two increasing functions; any increasing function continuous from the left generates a σ-additive measure.

To construct a functional G, it is sufficient to consider a regular BV, not BV_0, but in that case there is no isomorphism of the spaces. For example, take $t_* \in (a,b)$ and $G(f) = f(t_*)$. It is a linear functional. Then the function $g \in BV_0[a,b]$ is

$$g(x) = \begin{cases} 0, & x \in [0, t_*], \\ 1, & x \in (t_*, 1], \end{cases}$$

see Fig. 9.1, while one could include t_* to the right interval with $g(t_*) = 1$ and $g(t_* - 0) = 0$, and both functions would be fine. To exclude these extra options and establish an isomorphism, one should require the functions from BV_0 to be continuous either from the left or from the right.

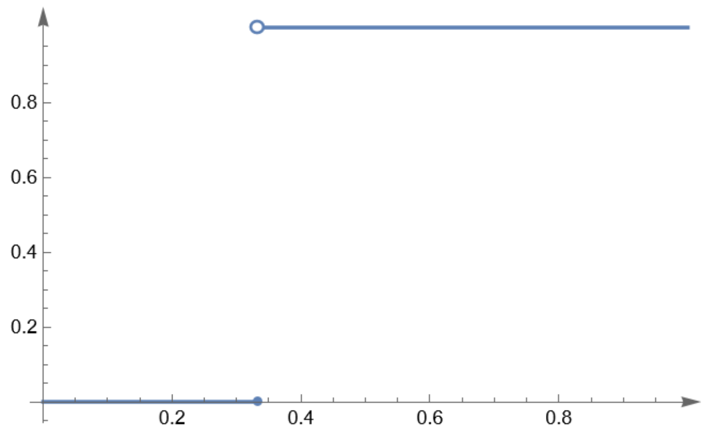

Рис. 9.1. Graphs of $g(x)$.

It is also clear that

$$\|G\|_{(C[a,b])^*} = V_a^b g \equiv \|g\|_{BV_0[a,b]}.$$

Lecture 10. $\big(C[a,b]\big)^*$. Norms of Functionals

Discussion of Self-Study Problems from the Previous Lecture

We begin with discussion of the homework from Lecture 8.

1) $c^* \cong \ell_1 \oplus \mathbb{C}$ such that

$$c^* \ni f \leftrightarrow (y,\alpha),\ y = (y_1, y_2, \dots) \in \ell_1,\ \alpha \in \mathbb{C}.$$

It is clear that $\ell_1 \oplus \mathbb{C} \cong \ell_1$, and one could redefine α to be y_0, so

$$f(x) = \sum_{k=1}^{\infty} x_k y_k + x_0\alpha \equiv \sum_{k=0}^{\infty} x_k y_k, \quad \|f\| = \|y\|_{\ell_1} + |\alpha| \equiv \|\{y_k\}_{k=0}^{\infty}\|_{\ell_1}; \tag{10.1}$$

in fact, c^* distinguishes from c_0^* by a one-dimensional space, so it is convenient to write it with α as well.

Take $x \in c$; then

$$\exists \lim_{n\to\infty} x_n = a, \quad \text{and for} \quad e_0 = (1,1,1,\dots): \quad x - ae_0 \in c_0.$$

For this element, $\exists y = (y_1, y_2, \dots) \in \ell_1 \equiv c_0^*$:

$$f(x - ae_0) = \sum_{k=1}^{\infty} (x_k - a) y_k, \quad \|f\| = \|y\|_{\ell_1}.$$

Expanding $f(x - ae_0)$ by linearity, one can rewrite it as

$$f(x) - af(e_0) = \sum_{k=1}^{\infty} x_k y_k - a\sum_{k=1}^{\infty} y_k,$$

so we obtain

$$f(x) = \sum_{k=1}^{\infty} x_k y_k + a\Big(f(e_0) - \sum_{k=1}^{\infty} y_k\Big);$$

comparing it with (10.1), we see that $x_0 := a$. Note that the sum of y_k here converges since $y \in \ell_1$. With α of the form

$$\alpha = f(e_0) - \sum_{k=1}^{\infty} y_k,$$

we obtain (10.1); one can easily see that $\|f\| = \|y\|_{\ell_1} + |\alpha|$. First, we will provide an upper bound:

$$|f(x)| = \Big|\sum_{k=1}^{\infty} x_k y_k + x_0\alpha\Big| \leqslant \sup_{k\geqslant 1}|x_k| \sum_{k=1}^{\infty} |y_k| + |x_0||\alpha_0| \leqslant \|x\|_c\Big(\|y\|_{\ell_1} + \alpha\Big).$$

To demonstrate that this upper bound is, in fact, sharp, we will evaluate the functional at the elements of the sequence

$$x^n = (\operatorname{sgn} y_1, \operatorname{sgn} y_2, \ldots, \operatorname{sgn} y_n, \operatorname{sgn}\alpha, \operatorname{sgn}\alpha, \operatorname{sgn}\alpha, \ldots), \quad \|x^n\| \leqslant 1.$$

For $f(x^n)$, we have

$$f(x^n) = \sum_{k=1}^{n} |y_k| + \operatorname{sgn}\alpha \cdot \sum_{k=1}^{\infty} y_k + |\alpha|,$$

where

$$\sum_{k=1}^{n} |y_k| \to \sum_{k=1}^{\infty} |y_k|, \quad \operatorname{sgn}\alpha \cdot \sum_{k=1}^{\infty} y_k \to 0 \quad \text{as} \quad n \to \infty,$$

where the second one holds since the left-hand side is a tail of a converging series. Thus,

$$f(x^n) \to \|y\|_{\ell_1} + \alpha.$$

2) Find the norm of the functional

$$f(x) = \sum_{k=1}^{\infty} \frac{x_k}{k^{4/3}} \ \in \ \ell_3^*.$$

By the theorem on isometric isomorphism, $\ell_p^* \cong \ell_q$, $1/p + 1/q = 1$, and since $p = 3$, $q = 3/2$. The norm of f is

$$\|f\|_{\ell_3^*} \equiv \|f\|_{\ell_{3/2}} = \left(\sum_{k=1}^{\infty} \left(\frac{1}{k^{4/3}}\right)^{3/2}\right)^{2/3} = \left(\sum_{k=1}^{\infty} \frac{1}{k^2}\right)^{2/3} = \left(\frac{\pi^2}{6}\right)^{2/3}.$$

3) Find the norm of the functional

$$F(f) = \int_{-1}^{1} |x| f(x)\, dx + 2f\left(-\frac{1}{2}\right) - f\left(\frac{1}{4}\right) \ \in \ \big(C[-1,1]\big)^*.$$

It is more interesting to consider the functional with x instead of $|x|$:

$$F(f) = \int_{-1}^{1} x f(x)\, dx + 2f\left(-\frac{1}{2}\right) - f\left(\frac{1}{4}\right) \ \in \ \big(C[-1,1]\big)^*.$$

The answer would be the same since at the first step, one takes the integrand under the absolute value. Now, obtain an upper bound:

$$|F(f)| \leqslant \int_{-1}^{1} |x||f(x)|\, dx + 2\left|f\left(-\frac{1}{2}\right)\right| + \left|f\left(\frac{1}{4}\right)\right| \leqslant \|f\| \left(\int_{-1}^{1} |x|\, dx + 3\right) = 4\|f\|,$$

since in $C[-1,1]$, the norm is the maximum, and, therefore, the value at any specific point is bounded by the maximum from above.

The next step is to analyze the formula of the functional in order to determine for which element the equality in the upper bound can hold. It is convenient to take f_0 such that $\|f_0\| = 1$. One can take

$$f_0 = \begin{cases} \operatorname{sgn} x, & x \in [-1,1],\ x \notin \left\{-\frac{1}{2}, \frac{1}{4}\right\}, \\ 1, & x = -\frac{1}{2}, \\ -1, & x = \frac{1}{4}, \end{cases}$$

see Fig. 10.1.

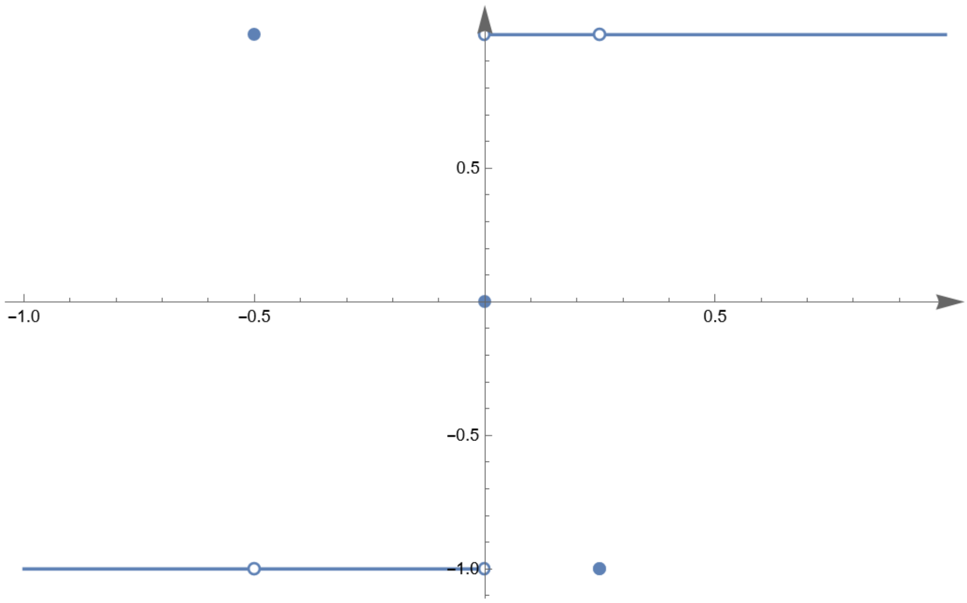

Рис. 10.1. Graphs of $f_0(x)$.

This function is not continuous on $[-1,1]$. It is not a problem, since we can construct a sequence of continuous functions f_n that approximate the given discontinuous function by connecting the discontinuities in small neighborhoods of the points of discontinuity, for instance, like this:

$$f_n = \begin{cases} \operatorname{sgn} x, & x \in [-1,1] \setminus \left(\left(-\frac{1}{2} - \frac{1}{n}, -\frac{1}{2} + \frac{1}{n}\right) \cup \left(-\frac{1}{n}, \frac{1}{n}\right) \cup \left(\frac{1}{4} - \frac{1}{n}, \frac{1}{4} - \frac{1}{n}\right)\right) \\ -2n\left|x + \frac{1}{2}\right| + 1, & x \in \left(-\frac{1}{2} - \frac{1}{n}, -\frac{1}{2} + \frac{1}{n}\right), \\ nx, & x \in \left(-\frac{1}{n}, \frac{1}{n}\right), \\ 2n\left|x - \frac{1}{4}\right| - 1, & x \in \left(\frac{1}{4} - \frac{1}{n}, \frac{1}{4} + \frac{1}{n}\right), \end{cases}$$

see Fig. 10.2. It is clear that $F(f_n) \to 4$, since $f_n \to f$ as $n \to \infty$.

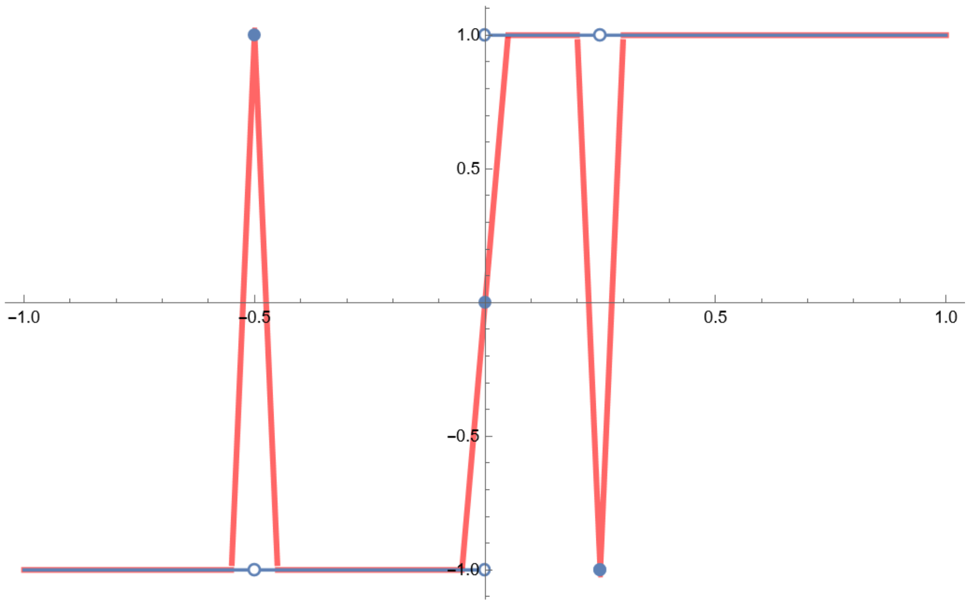

Рис. 10.2. Graphs of $f_0(x)$.

4) Consider

$$(Af)(x) = \int_a^b K(x,t)f(t)\,dt,$$

a) $K(x,t) \in C[a,b]^2$, $A : L_1[a,b] \to L_1[a,b]$. Find the norm $\|A\|$.

b) $K(x,t) \in C[a,b]^2$, $A : L_1[a,b] \to C[a,b]$. Find the norm $\|A\|$.

c) $K(x,t) \in L_2[a,b]^2$, $A : L_2[a,b] \to L_2[a,b]$. Find the bound C for the norm: $\|A\| \leqslant C$.

Now, begin with the item a).

a) First, as usual, we obtain an upper bound:

$$\|Af\| = \int_a^b \Big| \int_a^b K(x,t)f(t)\,dt \Big|\,dx \leqslant \int_a^b \int_a^b |K(x,t)| \cdot |f(t)|\,dt\,dx.$$

The functions $f(t)$, $K(x,t)$ are integrable. To continue the estimation, we use Fubini's theorem

$$\int_a^b \int_a^b |K(x,t)| \cdot |f(t)|\,dt\,dx \leqslant \int_a^b |f(t)| \Big| \int_a^b K(x,t)\,dx \Big|\,dt \leqslant$$

$$\leqslant \max_{t\in[a,b]} \Big(\int_a^b |K(x,t)|\,dx \Big) \cdot \int_a^b |f(t)|\,dt = \max_{t\in[a,b]} \Big(\int_a^b |K(x,t)|\,dx \Big) \|f\|_{L_1},$$

where the first factor is a candidate for being the norm of A:

$$\|A\|_{L_1 \to L_1} \leqslant \max_{t\in[a,b]} \Big(\int_a^b |K(x,t)|\,dx \Big).$$

Is this bound sharp? Is there a function for which the equalilty can be achieved? $K(x,t)$ is a continuous function, so, after the integration with respect to x, we obtain a continuous function in variable t; therefore,

$$\exists t_0 \in [a,b]: \quad \max_{t\in[a,b]}\left(\int_a^b |K(x,t)|\,dx\right) = \int_a^b |K(x,t_0)|\,dx.$$

Suppose that t_0 is an interior point of $[a,b]$ to consider two-sided neighborhoods of it (otherwise, neighborhoods are one-sided). We will integrate it with $f_n(t)$ of the form $f_n(t) = n\chi_{[t_0-1/(2n),t_0+1/(2n)]}$, see Fig. 10.3.

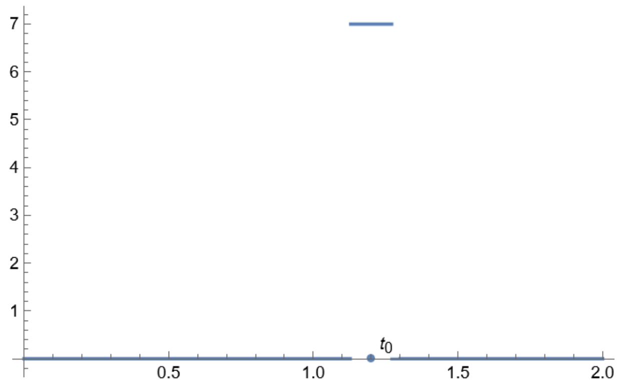

Рис. 10.3. Graphs of $f_7(x)$.

One can see that $\|f_n\| = 1$ in $L_1[a,b]$ (it is so-called *delta-sequence* since it tends to the delta-function). Substitute it to $\|Af_n\|$:

$$\|Af_n\| = \int_a^b \left|\int_{t_0-\frac{1}{2n}}^{t_0+\frac{1}{2n}} nK(x,t)\,dt\right| dx$$

It looks like one could use the mean value theorem for integrals:

$$\|Af_n\| = \int_a^b |K(x,t_n)|\,dx, \quad t_n \in \left[t_0 - \frac{1}{2n}, t_0 + \frac{1}{2n}\right].$$

As $n\to\infty$, $t_n \to t_0$, and

$$\int_a^b |K(x,t_n)|\,dx \to \int_a^b |K(x,t_0)|\,dx,$$

so

$$\|A\| = \int_a^b |K(x,t_0)|\,dx \equiv \max_{t\in[a,b]}\left(\int_a^b |K(x,t)|\,dx\right).$$

b) In this item, we use the similar approach:

$$\|Af\| = \max_{x\in[a,b]}\left|\int_a^b K(x,t)f(t)\,dt\right| \leqslant \max_{x\in[a,b]}\int_a^b |K(x,t)|\cdot|f(t)|\,dt \leqslant$$

$$\leqslant \max_{x,t\in[a,b]}|K(x,t)|\int_a^b |f(t)|\,dt \equiv \max_{x,t\in[a,b]}|K(x,t)|\|f\|_{L_1},$$

where the the first factor is a candidate for being the norm of $\|A\|$. Since $K(x,t)\in C[a,b]^2$,

$$\exists(x_0,t_0): \quad |K(x_0,t_0)| = \max_{x,t\in[a,b]}|K(x,t)|,$$

where we take the maximum with respect to x due to the fact that the image space, $C[a,b]$, has such a norm, while the maximum in t can be achieved using a delta-sequence, so example for which the upper bound gives the equality, is the same as in the previous item. Therefore,

$$\|A\| = \max_{x,t\in[a,b]}|K(x,t)|.$$

c) In this item, the problem was stated as follows: find an upper bound for $\|A\|$, $A : L_2[a,b] \to L_2[a,b]$ with $K(x,t)\in L_2[a,b]^2$, instead of the exact value. To eliminate the square roots, we will work with the squared norm:

$$\|Af\|^2 = \int_a^b \left|\int_a^b K(x,t)f(t)\,dt\right|^2 dx \leqslant \int_a^b \left(\int_a^b |K(x,t)|\cdot|f(t)|\,dt\right)^2 dx.$$

We have to transform this integral to take out the squared norm of f. Let us use the Cauchy–Bunyakovsky–Schwarz inequality:

$$\int_a^b \left(\int_a^b |K(x,t)|\cdot|f(t)|\,dt\right)^2 dx \leqslant \int_a^b \left(\int_a^b |K(x,t)|^2\,dt \cdot \int_a^b\int_a^b |f(t)|^2\,dt\right) dx =$$

$$= \|K(x,t)\|^2_{L_2[a,b]^2}\|f\|^2.$$

Thus,

$$\|A\| \leqslant \|K\|_{L_2[a,b]^2}.$$

5) Consider

$$(Af)(x) = \int_0^x f(t)\,dt$$

a) in $C[0,1]$:

$$\|Af\| = \max_{x\in[0,1]}\left|\int_0^x f(t)\,dt\right| \leqslant \max_{x\in[0,1]}\int_0^x |f(t)|\,dt \leqslant \|f\|.$$

For the function $f\equiv 1$, $Af = x$ and $\max|Af| = 1$, therefore, $\|A\| = 1$.

b) in $L_2[0,1]$. First, it is convenient to transform the operator to the integration with fixed limits:

$$\int_0^x f(t)\,dt = \int_0^1 K(x,t)f(t)\,dt.$$

What can we say about the function $K(x,t)$? In fact,

$$K(x,t) = \chi_{x \geqslant t} \equiv \begin{cases} 1, & 0 \leqslant t \leqslant x \leqslant 1, \\ 0, & 0 \leqslant x < t \leqslant 1, \end{cases}$$

and this is an example of so-called *triangle kernels*, see Fig. 10.4.

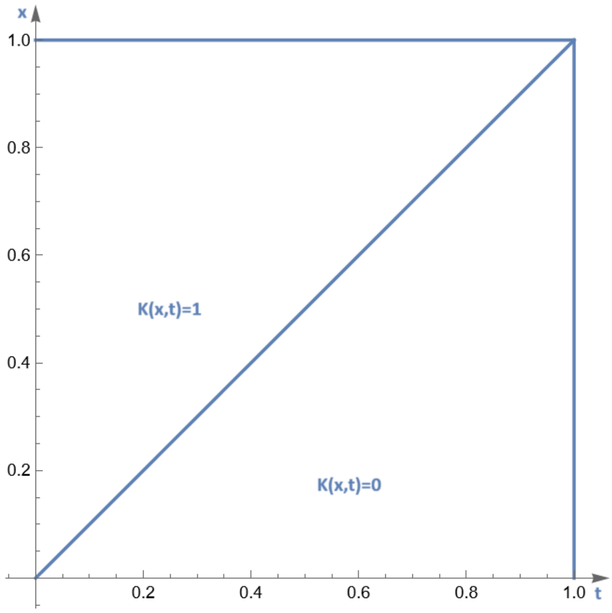

Рис. 10.4. Regions where $K(x,t)$ takes the values 0 and 1.

Using the results of 4c), we see that

$$\|A\| \leqslant \|K\|_{L_2[a,b]^2} = \frac{1}{\sqrt{2}}.$$

(Spoiler: In fact, the norm is less than this.)

Dual Space to C[a,b]

Theorem 10.1.

$$\Big(C[a,b]\Big)^* \cong BV_0[a,b] = \{g \in BV[a,b],\ g(a) = 0,\ \forall x \in (a,b) : g(x-0) = g(x)\},$$

such that

$$\Big(C[a,b]\Big)^* \ni G \leftrightarrow g \in BV_0[a,b]: \quad G(f) = \int_a^b f(t)\,dg(t), \quad \textit{and} \quad \|G\| = \|g\|_{BV_0}.$$

Let us comment on a minor issue. Why does g need to be continuous from the left only on the interval (a,b) , and why is it not required to be continuous at the endpoints? At the point a, there is no left-sided neighborhood in $[a,b]$. For the point b, the answer will appear later.

Note also that for the spaces ℓ_p, L_p, c_0, and c, the theorems on the isometric isomorphism of the dual space to some nice space make it more simple to find the norm of the functional in practice. Unfortunately, this is not true for $(C[a,b])^*$; in this space, it is easier to find the norm of the functional by definition.

Proof.

1) $\Leftarrow$. Let $g \in BV_0[a,b]$. Construct

$$G(f) = \int_a^b f(t)\,dg(t)$$

and try to estimate it:

$$|G(f)| \leqslant \int_a^b |f(t)|\,|dg(t)| \leqslant \|f\| \int_a^b |dg(t)| = \|f\| V_a^b g,$$

since by definition of the Riemann–Stieltjes integral, it is the limit of the sum with respect to all partitions of $[a,b]$:

$$\sum_k |g(t_k) - g(t_{k-1})|,$$

therefore, $\|G\| \leqslant \|g\|$. At this step, we will not try to obtain the equalilty of the norms, since one can do it at the second step, where we are to construct a function from BV_0 starting from a functional. As for now, it is sufficient to understand that to each function g from BV_0, there corresponds a functional $G \in (C[a,b])^*$.

2) Suppose that $G \in (C[a,b])^*$. Let us use the Hahn–Banach theorem. Recall that $C[a,b] \subset L_\infty[a,b]$, with the same norm: in L_∞, we have a supremum-norm, which coincides with the maximum for continuous functions. In L_∞,

$$\|f\|_{L_\infty} = \inf_{\mu(E)=0} \sup_{[a,b]\setminus E} |f(x)|.$$

Then, by the Hahn–Banach theorem, G can be extended to $\widetilde{G}$ in the entire L_∞. We can apply $\widetilde{G}$ to discontinuous functions, for instance,

$$\widetilde{G}\Big(\chi_{[a,t)}\Big), \quad \text{where} \quad \chi_{[a,t)} = \begin{cases} 1, & x < t, \\ 0, & x \geqslant t. \end{cases}$$

This is a function of t. We claim that this is the function we need:

$$\widetilde{G}\Big(\chi_{[a,t)}\Big) =: g(t).$$

Further,

$$\widetilde{G}\Big(\chi_{[t_1,t_2)}\Big) = \widetilde{G}\Big(\chi_{[a,t_2)} - \chi_{[a,t_1)}\Big) = g(t_2) - g(t_1).$$

Let $T = \{t_k\}_{k=0}^n$, $a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$ be some partition of $[a,b]$. Construct the function

$$f_T(x) = \begin{cases} \operatorname{sgn}\big(g(t_k) - g(t_{k-1})\big)\chi_{[t_{k-1},t_k)}, & k < n, \\ \operatorname{sgn}\big(g(b) - g(t_{n-1})\big), & x \in [t_{n-1}, b]. \end{cases}$$

One can write this function in the following way:

$$f_T(x) = \sum_{k=1}^n \operatorname{sgn}\big(g(t_k) - g(t_{k-1})\big)\chi_{[t_{k-1},t_k)}, \quad \|f_T\|_{L_\infty} \leqslant 1,$$

where for $k = n$, the last interval (in the subscript of χ) is closed: $[t_{n-1}, b]$. The functional $\widetilde{G}$ is linear, so

$$\widetilde{G}(f_T) = \sum_{k=1}^n |g(t_k) - g(t_{k-1})|,$$

and $\|\widetilde{G}\| \geqslant |\widetilde{G}(f_T)| = V_T g$ $(\forall T)$. Taking the supremum over all partitions, we obtain

$$\|\widetilde{G}\| \geqslant V_a^b g,$$

where, by the Hahn–Banach theorem, $\|G\| = \|\widetilde{G}\|$, so we obtain the inverse inequality for the norms.

It is clear that $g(a) = 0$.

We must also show that the action of G is integration with dg. Consider the integral

$$\int_a^b f_T\,dg = \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \operatorname{sgn}\big(g(t_k) - g(t_{k-1})\big)\,dg = \sum_{k=1}^n \operatorname{sgn}\big(g(t_k) - g(t_{k-1})\big) \int_{t_{k-1}}^{t_k} dg,$$

where

$$\int_{t_{k-1}}^{t_k} dg = g(t_k) - g(t_{k-1}),$$

so

$$\int_a^b f_T\,dg = V_T g.$$

Thus, one can see that

$$\widetilde{G}(f_T) = \int_a^b f_T\,dg,$$

since we obtain the same result on the left- and right-hand sides. Since both sides are linear in their arguments f_T, one can evaluate the functional at the linear combination of the functions of this kind

$$\widetilde{G}\Big(\sum_k c_k f_{T_k}\Big) = \sum_k c_k \widetilde{G}(f_{T_k}) = \sum c_k \int f_{T_k}\, dg = \int \Big(\sum_k c_k f_{T_k}\Big)\, dg$$

for some number of partitions T_k. One can see that any continuous function can be approximated in terms of step functions with any given accuracy, for instance,

$$x \approx \frac{[xn]}{n},$$

see Fig. 10.5.

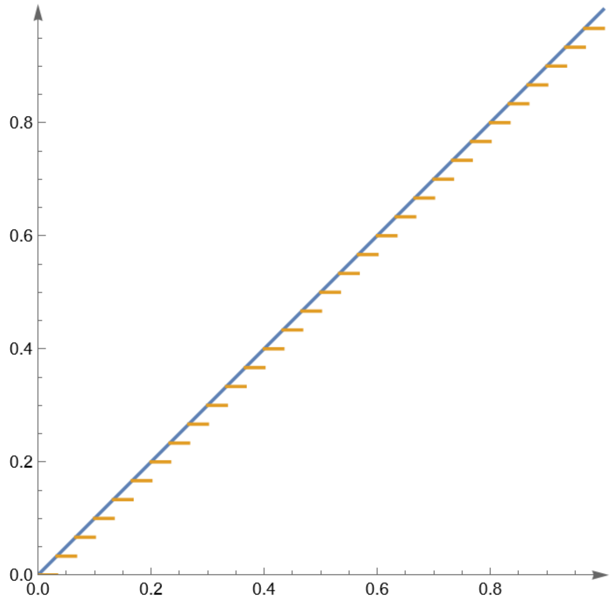

Рис. 10.5. Approximation of $f(x) = x$ with $\frac{[xn]}{n}$, $n = 30$.

For any $f \in C[a,b]$, define $f_n := f([xn]/n)$. It is obvious that $f_n \to f$ (pointwise). Then,

$$\widetilde{G}(f_n) = \int f_n\, dg,$$

where the left-hand side converges to $G(f)$ and the right-hand side converges to $\int f\, dg$. □

Consider an example from the homework:

$$F(f) = \int_{-1}^{1} |x| f(x)\, dx + 2f\Big(\frac{-1}{2}\Big) - f\Big(\frac{1}{4}\Big). \tag{10.2}$$

If we are to find the norm of the functional F, then we can rewrite it as

$$F(f) = \int_{-1}^{1} f(x)\, dg$$

and then find the total variation of g. Recall that

$$\alpha f(t_0) = \int f\, dg$$

with g as depicted in Fig. 10.6.

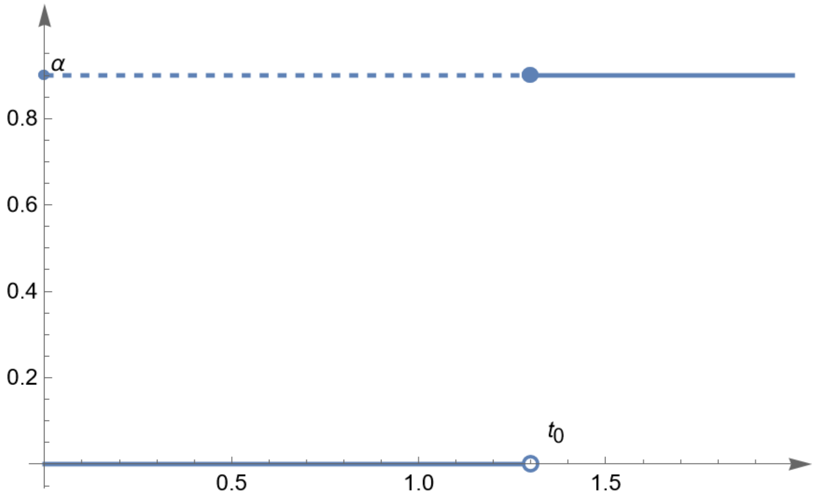

Рис. 10.6. Graphs of $g(x)$.

Further, rewrrite (10.2) as

$$F(f) = \int_{-1}^{0} -xf(x)\, dx + \int_{0}^{1} xf(x)\, dx + 2f\left(\frac{-1}{2}\right) - f\left(\frac{1}{4}\right).$$

The function g that corresponds to F is as in Fig. 10.6.

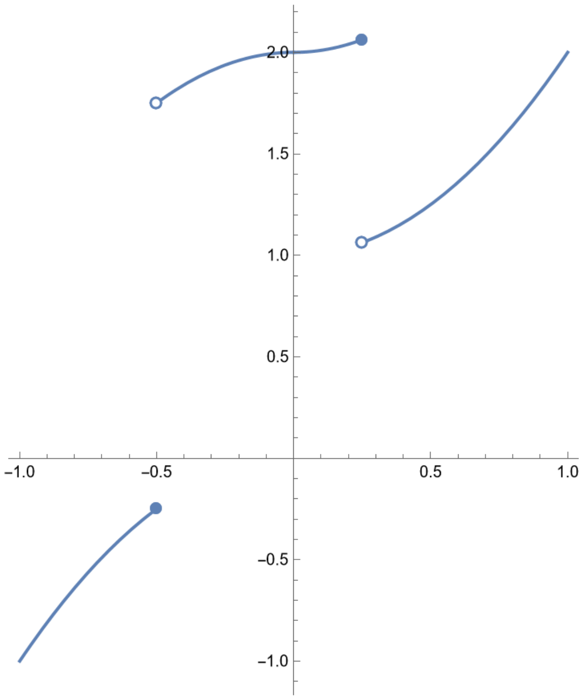

Рис. 10.7. Graphs of $g(x)$.

Computing the total variation of this function is not very convenient, so it is easier to find the norm of F by definition.

Self-Study Problems

1) Let X be a normed space and $X_0 \subset X$ be a nontrivial closed subspace. Let $x \notin X_0$, and

$$\operatorname{dist}(x, X_0) := \inf_{x_0 \in X_0} \|x - x_0\| = d > 0.$$

Show that

$$\exists f \in X^*, \ \|f\| = 1: \quad f(x) = d, \quad f\Big|_{X_0} = 0.$$

2) Let X be a Banach space. Prove that if X^* is separable, then X is separable as well.

3) $f(x) = x^\alpha \sin\frac{1}{x}$. For which α does f belong to $BV[0,1]$?

4) Consider

$$M = \left\{ f \in C[0,1] : \int_0^1 f(x)\,dx = 0 \right\}.$$

Find $\operatorname{dist}(1, M)$.

5) Let X be a normed space, $X_0 = \operatorname{Ker} f$, $f \in X^*$. Prove that $\operatorname{dist}(x, \operatorname{Ker} f) = \frac{|f(x)|}{\|f\|}$.

Lecture 11. The Dual of a Hilbert Space. Modes of Convergence in Normed Spaces

The Dual of a Hilbert Space. Riesz Representation Theorem

In the previous lectures, we proved the following. Let $1 \leqslant p < \infty$, and let (Ω, M, μ) be a measurable space. Then

$$\Big(L_p(\Omega,\mu)\Big)^* \cong L_q(\Omega,\mu),$$

where

$$\frac{1}{p}+\frac{1}{q}=1.$$

Remark 11.1. *One can see that for $p=2$, we have $q=2$; thus, $L_2(\Omega,\mu)$ is dual to itself. Furthermore, $L_2(\Omega,\mu)$ is a Hilbert space.*

The situation when the dual space coincides with the space itself is quite typical for Hilbert spaces. Next, we will prove some auxiliary lemmas, and then come to a general statement on the dual space for a Hilbert space H.

Lemma 11.1. *Let X and Y be normed spaces and $A \in B(X,Y)$ be a bounded operator. Then*

$$\operatorname{Ker} A := \{x \in X : \ Ax = 0\}$$

is a closed linear subspace in X.

Proof. We will prove that $\operatorname{Ker} A$ is linear and closed separately.

1) The linearity of $\operatorname{Ker} A$ is obvious, since A is a linear operator and the spaces X, Y are linear. Consider $x_1, x_2 \in \operatorname{Ker} A$. Consider, for $\alpha, \beta \in \mathbb{C}$ (or $\mathbb{R}$),

$$A(\alpha x_1 + \beta x_2) = \alpha A x_1 + \beta A x_2,$$

where $Ax_j = 0$ by the definition of $\operatorname{Ker} A$, so $A(\alpha x_1 + \beta x_2) = 0$, which means that

$$\alpha x_1 + \beta x_2 \in \operatorname{Ker} A.$$

2) Next, we prove that $\operatorname{Ker} A$ is closed. Let x_0 be a limit point of $\operatorname{Ker} A$. By definition,

$$\exists \{x_n\}_{n=1}^{\infty} : \quad x_n \in \operatorname{Ker} A, \quad x_n \to x$$

as $n \to \infty$. Since A is bounded, it is continuous; therefore, $Ax_n \to Ax_0$. Since

$$x_n \in \operatorname{Ker} A, \quad Ax_n = 0,$$

we have $Ax_0 = 0$, which means that $x_0 \in \operatorname{Ker} A$, so $\operatorname{Ker} A$ is closed. □

Since linear functional is a particular linear operator, the following statement simply follows Lemma 11.1.

Corollary 11.1. *Let X be a normed space and $f \in X^*$ be a linear bounded functional. Then* $\operatorname{Ker} f$ *is a closed linear subspace in X.*

Of course, the infinite-dimensional case is of utmost interest, since in the finite-dimensional one, one can describe the dual space in terms of linear algebra.

We will consider only Hilbert spaces, although one can consider Banach spaces slightly modifying all the constructions.

Lemma 11.2. *Let H be a Hilbert space. Define*

$$H_0 := \operatorname{Ker} f, \quad f \in H^*.$$

Then $\dim H_0^{\perp} \leqslant 1$.

Remark 11.2. *If X is a Banach space and $f \in X^*$, for*

$$X_0 := \operatorname{Ker} f,$$

we can consider the quotient space X/X_0; since X_0 is closed, it is a well-defined space with quotient norm. It turns out that

$$\dim X/X_0 \leqslant 1.$$

Proof of Lemma 11.2.

1) Consider $f \equiv 0$. Then

$$H_0 \equiv \operatorname{Ker} f = H,$$

and $H_0^{\perp} = \{0\}$, that is, $\dim H_0^{\perp} = 0$.

2) Consider $f \not\equiv 0$. Then

$$\exists x \in H, x \neq 0: \quad f(x) \neq 0.$$

Next, we proceed the proof by contradiction. Suppose that there exist x_1, x_2 in $H_0^{\perp}$ that are linearly independent; this means that $\dim H_0^{\perp} > 1$. Recall that the orthogonal complement of a subspace is a closed linear subspace; thus, $H_0^{\perp}$ is linear and closed. Consider the element

$$x := f(x_1)x_2 - f(x_2)x_1 \in H_0^{\perp},$$

where $f(x_j) \neq 0$ since $x_j \in H_0^{\perp}$, $H_0 = \operatorname{Ker} f$. Evaluating f at this element, we obtain $f(x) = 0$. Therefore, $x \in H_0$. Therefore, $x = 0$, while the coefficients $f(x_1)$ and $f(x_2)$ in x_2 and x_1 respectively are nonzero, which contradicts the linear independence of x_1 and x_2. □

Now we are all set to formulate the Riesz representation theorem.

Theorem 11.1 (The Riesz Representation Theorem). *Let H be a Hilbert space. Then $H^* \cong H$:*

$$\forall f \in H^* \ \exists! y \in H: \quad \forall x \in H \ f(x) = (x, y).$$

Remark 11.3. *Here $\cong$ denotes a linear isometric isomorphism, which is conjugate for H over $\mathbb{C}$. The term isometric means that*

$$\|f\|_{H^*} = \|y\|_H;$$

next, for f, $g \in H^*$, for which $f \leftrightarrow y$, $g \leftrightarrow z$, we have

$$\alpha f + \beta g \leftrightarrow \overline{\alpha} y + \overline{\beta} z.$$

Note also that $y \in H$ corresponding to $f \in H^*$ is called the Riesz representation of f.

Next, for ℓ_2, we proved

$$\forall f \in \ell_2^* \ \exists! y \in \ell_2: \quad f(x) = \sum_{i=1}^{\infty} x_i y_i;$$

as opposed to this, now the isomorphism is established not through the sum, but via the inner product:

$$(x, y) = \sum_{i=1}^{\infty} x_i \overline{y}_i.$$

Proof. Define

$$H_0 := \operatorname{Ker} f.$$

Then, by Lemma 11.2, $\dim H_0^\perp$ is either 0 or 1. Consider the following cases.

1) $f \equiv 0$; then we have $H_0 = H$, so $y = 0$.

2) $f \not\equiv 0$; then $\dim H_0^\perp = 1$. Define

$$H_1 := H_0^\perp.$$

Then, due to the properties of the orthogonal complement, one can decompose the space as

$$H = H_0 \oplus_\perp H_1.$$

This means that $\forall x \in H$ being decomposed as $x = x_0 + x_1$, $x_j \in H_j$, we have $f(x_0) = 0$, so it is sufficient to describe the action of f only at x_1. H_1 is a one-dimensional space; let e_1 be a basis in H_1: $H_1 = \langle e_1 \rangle$. Then

$$x = x_0 + \alpha e_1, \quad f(x) = \alpha f(e_1).$$

The idea is to look for $y \in H_1$, and then we will the uniqueness. Let us set $y = \beta e_1$ and calculate the inner product with x:

$$(x,y) = (x_0 + \alpha e_1, \beta e_1);$$

since $x_0 \perp e_1$, we have

$$(x_0 + \alpha e_1, \beta e_1) = \alpha\overline{\beta}\|e_1\|^2.$$

We want the right-hand side to coincide with $f(x)$ for any $\alpha \in \mathbb{C}$. Then, one solve the condition $f(x) = \alpha\overline{\beta}\|e_1\|^2$ for β:

$$\beta = \frac{\overline{f}(e_1)}{\|e_1\|^2},$$

so

$$y = \frac{\overline{f}(e_1)}{\|e_1\|^2} e_1.$$

Let us prove the uniqueness of such an element. Suppose that there exists $z \neq y$:

$$f(x) = (x,z) \quad \forall x \in H.$$

Then, by the linearity of the inner product,

$$0 = (x, y-z),$$

so $y - z \perp H$, which is possible only for $z = y$.

Next, we must prove that the map $f \mapsto y$ is an isometry. By the Cauchy–Bunyakovsky,

$$|f(x)| = |(x,y)| \leqslant \|x\|\,\|y\|;$$

this implies

$$\|f\| \leqslant \|y\|.$$

Taking $x = y/\|y\|$, $\|x\| = 1$, we get $f(x) = \|y\|$, so

$$\|f\|_{H^*} = \|y\|_H,$$

which completes the proof. □

Let us consider the following examples that demonstrate the power of the Riesz theorem:

Example 11.1. *In $L_2[0,1]$, consider*

$$F(f) = \int_0^1 \sqrt{x} f(x)\,dx.$$

One can see that $F(f) = (f,g)$ for $g = \sqrt{x}$. Therefore,

$$\|F\| = \|g\| = \sqrt{\int_0^1 x\,dx} = \frac{1}{\sqrt{2}}.$$

We will consider more complicated examples later. Now, let us introduce a new concept.

Reproducing Kernel Hilbert Spaces

The concept of the reproducing kernel Hilbert space (RKHS) is highly useful in the geometry of Hilbert spaces and Mathematical Physics. In fact, an RKHS can be thought of as a generalization of Green's functions for differential equations.

Definition 11.1. *A Hilbert space H is called a* ***reproducing kernel Hilbert space(RKHS)*** *if*

1) H is a space of ***functions*** *on some set Ω.*

2) $\forall a \in \Omega\ \forall f \in H$ the point evaluation $F_a(f): f \mapsto f(a)$ is a continuous linear functional.

By the Riesz representation theorem, for $F_a \in H^*$, there exists a function $g_a \in H$ such that

$$F_a(f) = (f, g_a).$$

Let us take another point $b \in \Omega$ and $F_b(f) = f(b)$. Then,again by the Riesz representation theorem, there exists $g_b \in H$ such that

$$F_b(f) = (f, g_b).$$

Here, $a, b \in \Omega$ are just some fixed points. Next, one can consider the following function:

$$K(a,b) = (g_b, g_a): \quad \Omega \times \Omega \to \mathbb{C}.$$

Definition 11.2. *$K(a,b)$ is called a* ***reproducing kernel****.*

One can see that, by the reproducing properties of g_a and g_b,

$$K(a,b) = g_b(a) = \overline{g_a(b)}.$$

In fact, all the convenience of reproducing kernel Hilbert spaces is concentrated at this property.

Let us consider the following examples.

Example 11.2. *1) $L_2[a,b]$ is **not** a reproducing kernel Hilbert space. Note that L_2 fails to meet even the first condition in the definition of an RKHS, as it does not consist of functions in the classical sense but rather equivalence classes of functions differing on sets of measure zero; also, the elements of L_2 cannot be evaluated at a point.*

2) $\mathbb{R}^n$ is a reproducing kernel Hilbert space. For $x,y \in \mathbb{R}^n$, $x = (x_1,x_2,\dots,x_n)$, $y = (y_1,y_2,\dots,y_n)$, we have

$$(x,y) = \sum_{i=1}^{n} x_i y_i.$$

This is a space of functions on the set $\Omega = (1,2,\dots,n)$, and the value at the point $k \in \Omega$ is given by $x_k = (x,e_k)$, where $e_k = (0,\dots,0,\overset{k}{1},0,\dots,0)$. One can see that

$$K(i,j) = (e_i,e_j) = \delta_{ij},$$

so the reproducing kernel is the identity matrix.

3) Consider $W_2^1[-1,1]$ over $\mathbb{R}$ with inner product

$$(f,g) = \int_{-1}^{1} fg\,dx + \int_{-1}^{1} f'g'\,dx.$$

Recall that

$$W_2^1 = [-1,1] = \{f \in AC[-1,1],\ f' \in L_2[-1,1]\}.$$

In this space, the point evaluation is well-defined; take $a \in [-1,1]$ and consider $F_a(f) = f(a)$. By the Riesz representation theorem,

$$F_a(f) = (f,g_a).$$

What is g_a here and what is the reproducing kernel $K(a,b)$? What is the norm of F_a? By the Riesz representation theorem, a point evaluation functional can be represented in the form of the inner product:

$$f(a) = \int_{-1}^{1} fg_a\,dx + \int_{-1}^{1} f'g_a'\,dx.$$

The idea is to integrate by parts. Decomposing the second integral,

$$\int_{-1}^{1} f'g_a'\,dx = \int_{-1}^{a} f'g_a'\,dx + \int_{a}^{1} f'g_a'\,dx,$$

we can take the derivative to g_a assuming that it has an additional derivative (recall that $g_a \in W_2^1$):

$$fg_a'\Big|_{-1}^{a} - \int_{-1}^{a} fg_a''\,dx + fg_a'\Big|_{a}^{1} - \int_{a}^{1} fg_a''\,dx.$$

Next, as there are no integrals on the left-hand side, let us try to cancel the integrals on the right-hand side. Decomposing the first integral and combining them with the expression for the second one above, we get

$$f(a) = \int_{-1}^{a} (fg_a - fg_a'')\,dx + \int_{a}^{1} (fg_a - fg_a'')\,dx + fg_a'\Big|_{-1}^{a} + fg_a'\Big|_{a}^{1}. \tag{11.1}$$

Suppose that g_a *satisfies*

$$g_a = g_a'' \quad on \quad [-1,a], \qquad\qquad g_a = g_a'' \quad on \quad [a,1] \tag{11.2}$$

separately (not on the entire interval as a whole); then the integrals in (11.1) *disappear, and we arrive at*

$$f(a) = f(a)g_a'(a-0) - f(-1)g_a'(-1) + f(1)g_a'(1) - f(a)g_a(a+0).$$

Further, we must cancel the second and the third terms (for any f*), which gives*

$$g_a'(1) = 0, \qquad g_a'(-1) = 0. \tag{11.3}$$

Combining the first and the last terms, we get $f(a)\big(g_a(a-0) - g_a(a+0)\big)$*, so for this expression to be equal to* $f(a)$*, it is necessary that*

$$g_a'(a-0) - g_a'(a+0) = 1. \tag{11.4}$$

Additionally, as $g_a \in W_2^1[-1,1]$*, and, consequently,* $g_a \in AC[-1,1]$*,*

$$g_a(a-0) = g_a(a+0). \tag{11.5}$$

Instead of writing a solution of (11.2) *in terms of exponential functions, let us use the hyperbolic functions:*

$$g_a = A\cdot\cosh(x+1) \quad on \quad [-1,a)$$

and

$$g_a = B\cdot\cosh(x-1) \quad on \quad [a,1];$$

these are solutions to the differential equations (**??**) *along with conditions* (11.3) *for the derivative. Next, the continuity condition* (11.5) *becomes*

$$A\cdot\cosh(a+1) = B\cdot\cosh(a-1),$$

so $B = A\cdot\cosh(a+1)/\cosh(a-1)$*. Next, for the jump of derivative* (11.4)*, we have*

$$A\sinh(a+1) - \frac{A\cdot\cosh(a+1)}{\cosh(a-1)}\sinh(a-1) = 1.$$

Multiplying by $\cosh(a-1)$, *we get*

$$A\Big(\sinh(a+1)\cosh(a-1) - \cdot\cosh(a+1)\sinh(a-1)\Big) = \cosh(a-1),$$

or simply

$$A = \frac{\cosh(a-1)}{\sinh 2}, \qquad B = \frac{\cosh(a+1)}{\sinh 2}.$$

Thus,

$$g_a(x) = \begin{cases} \dfrac{\cosh(a-1)}{\sinh 2}\cosh(x+1), & x \in [-1, a), \\ \dfrac{\cosh(a+1)}{\sinh 2}\cosh(x-1), & x \in [a, 1]. \end{cases}$$

By the Riesz representation theorem, the norm of the functional F_a *coincides with the norm of* g_a*:*

$$\|F_a\| = \|g_a\|.$$

One can find the norm of g_a *in a straightforward way, through the integration (by calculating the inner product with itself). However, it is easier to use the reproducing property:*

$$\|g_a\| = \sqrt{(g_a, g_a)} = \sqrt{g_a(a)} = \sqrt{\frac{\cosh(a-1)\cosh(a+1)}{\sinh 2}}.$$

In case $a = 0$, *we have*

$$\|F_a\| = \sqrt{\frac{\cosh^2 1}{\sinh 2}} = \sqrt{\frac{\coth 1}{2}}.$$

Next, for the reproducing kernel $K(a,b)$ *we have*

$$K(a,b) = g_a(b) = g_b(a) = \begin{cases} \dfrac{\cosh(a-1)\cosh(b+1)}{\sinh 2}, & -1 \leqslant b < a, \\ \dfrac{\cosh(a+1)\cosh(b-1)}{\sinh 2}, & a \leqslant b \leqslant 1. \end{cases}$$

Modes of Convergence in Normed Spaces

I. Let X be a normed space.

1. We have the concept of *norm* convergence (convergence *with respect to the norm*):

$$x_n \xrightarrow{\|\cdot\|} x \quad \text{if} \quad \|x_n - x\| \to 0.$$

2. *Weak* convergence: $x_n \rightharpoonup x$ (or $x_n \xrightarrow{w} x$).

Definition 11.3. $x_n \rightharpoonup x$ *if* $\forall f \in X^*$*:* $f(x_n) \to f(x)$.

One can see that this kind of convergence is indeed weaker than the norm convergence, as the norm convergence implies the weak one:

Statement 11.1. *If* $x_n \xrightarrow{\|\cdot\|} x$ *then* $x_n \rightharpoonup x$.

Proof. Consider

$$\Big|f(x_n) - f(x)\Big| = \Big|f(x_n - x)\Big| \leqslant \|f\| \cdot \|x_n - x\|;$$

since $\|f\| < \infty$ as $f \in X^*$, for $\|x_n - x\| \to 0$ we have

$$|f(x_n) - f(x)| \to 0. \qquad \square$$

The converse is not true; consider the following example. Let H be a Hilbert space, $\dim H = \infty$, and $\{e_n\}_{n=1}^{\infty}$ be an orthonormal system. Let us prove that $e_n \rightharpoonup 0$.

For $f \in H^*$, by the Riesz representation theorem, we have

$$f(e_n) = (e_n, y), \quad y \in H.$$

One can see that $(e_n, y) = \overline{y}_n$, where y_n is the corresponding Fourier coefficient. It is known that $y_n \to 0$ as $n \to \infty$, since $\{y_n\}_{n=1}^{\infty} \in \ell_2$ (see the Bessel inequality). As we can see from Statement 11.1, the weak limit of a norm-converging sequence coincides with the limit with respect to the norm. However,

$$e_n \overset{\|\cdot\|}{\not\to} 0,$$

since $\|e_n\| = 1$.

II. Let X be a normed space with the dual space X^*. Consider the convergences of $f_n \in X^*$.

1. The norm convergence: $f_n \xrightarrow{\|\cdot\|} f \Leftrightarrow \|f_n - f\| \to 0$.
2. The weak convergence: $f_n \rightharpoonup f \Leftrightarrow \forall F \in (X^*)^* = X^{**}$: $F(f_n) \to F(f)$.
3. In X^*, we have an additional convergence, called the *weak** (*weak-star*) convergence. It is defined by

$$f_n \overset{*}{\rightharpoonup} f \quad \Leftrightarrow \quad \forall x \in X: \quad f_n(x) \to f(x), \tag{11.6}$$

i.e., it is the pointwise convergence.

One can see that 1. $\Rightarrow$ 2. $\Rightarrow$ 3.; the first implication is proved above. To prove the second one, let us recall one of the corollaries of Hahn–Banach theorem: the embedding $X \hookrightarrow X^{**}$ is isometry; that is, the isometry is determined by

$$x \mapsto F_x, \qquad F_x(f) = f(x).$$

One can see that in (11.6), we have

$$F_x(f_n) \to F_x(f),$$

and the weak convergence is valid for all F, while the weak-star convergence holds for F_x, $x \in X$. If X is reflexive, i.e., $X \cong X^{**}$, the weak convergence coincides with the weak-star one.

Consider the example that the weak-star convergence is not equivalent to the weak one. We know that $c_o^* \cong \ell_1$ and $\ell_1^* \cong \ell_\infty$. Let us consider the weak and weak-star convergence in ℓ_1 as a dual to c_0.

Take $f_n \in c_0^*$. To f_n, there corresponds

$$e_n = (0, 0, \dots, 0, \overset{n}{1}, 0, \dots) \in \ell_1.$$

Next,

$$\forall x \in c_0 : \quad f_n(x) = x_n \to 0 \quad \text{as} \quad n \to \infty,$$

so $f_n \xrightarrow{*} 0$. However, taking $F \in \ell_1^* \cong \ell_\infty$,

$$F = (1, 1, 1, \dots, 1, \dots) \in \ell_\infty,$$

we can see that $F(f_n) = 1 \not\to 0$, so $f)n \not\rightharpoonup 0$.

Modes of Convergence in $B(X,Y)$

In the space of bounded operators, we have the norm convergence. However, as an operator is defined by its action on $x \in X$, it is convenient to define another kind of convergence.

Definition 11.4. *Let X and Y be Banach spaces, and $B(X,Y)$ be the space of bounded linear operators $X \to Y$, $A_n, A \in B(X,Y)$.*

*1) **Norm convergence** is defined by*

$$A_n \xrightarrow{\|\cdot\|} A : \quad \|A_n - A\| \to 0.$$

2) ***Strong operator convergence*** *is defined by*

$$A_n \xrightarrow{s} A \quad \Leftrightarrow \quad \forall x \in X: \quad A_n x \xrightarrow{\|\cdot\|_Y} Ax.$$

3) ***Weak operator convergence*** *is defined by*

$$A_n \rightharpoonup A \quad \Leftrightarrow \quad \forall x \in X,\ \forall f \in Y^*: \quad f(A_n x) \to f(Ax).$$

Note that in a Hilbert space H with $A_n, A \in B(H)$, we have

$$A_n \rightharpoonup A \quad \Leftrightarrow \quad (A_n x, y) \to (Ax, y).$$

Let us prove that 1) $\Rightarrow$ 2) $\Rightarrow$ and the converse implications do not hold.

Statement 11.2. *If $A_n \xrightarrow{\|\cdot\|} A$ then $A_n \xrightarrow{s} A$.*

Proof. Consider

$$\|A_n x - Ax\| = \|(A_n - A)x\| \leqslant \|A_n - A\| \cdot \|x\|,$$

which proves the statement. □

Statement 11.3. *If $A_n \xrightarrow{s} A$ then $A_n \rightharpoonup A$.*

Proof. Consider

$$\big|f(A_n x) - f(Ax)\big| \leqslant \big|f(A_n x - Ax)\big| \leqslant \|f\| \cdot \|A_n x - Ax\| \to 0,$$

which proves the statement. □

Further, let us prove that the strong operator convergence does not imply the convergence with respect to the norm. In ℓ_2, consider

$$A_n x = (x_1, x_2, \ldots, x_n, 0, \ldots).$$

One can see that $A_n \xrightarrow{s} I$ since

$$\forall x: \quad \|A_n x - Ix\| = \Big(\sum_{k=n+1}^{\infty} |x_k|\Big)^{1/2} \to 0.$$

It is easy to check that this sequence is not a Cauchy sequence in the norm sense:

$$\|A_n - A_m\| \not\to 0;$$

consider, for $n > m$,

$$\|A_m x - A_n x\| = \Big(\sum_{k=m+1}^{n} |x_k|^2\Big)^{1/2} \leqslant \|x\|,$$

that is, $\|A_n - A_m\| \leqslant 1$. Taking

$$\|(A_n - A_m)e_{m+1}\| = 1,$$

we can see that $\|A_m - A_n\| = 1 \not\to 0$.

Finally, let us show that the weak operator convergence does not imply the strong one. In ℓ_2, consider

$$A_n x = (0, \dots, \overset{n}{0}, x_1, x_2, \dots);$$

one can see that $A_n = A_r^n$. Let us show that $A_n \rightharpoonup 0$. Indeed, $\forall x, y$:

$$\big|(A_n x, y)\big| = \Big|\sum_{i=1}^{\infty} x_i y_{n+i}\Big| \leqslant \Big(\sum_{i=1}^{\infty} |x_i|^2\Big)^{1/2} \Big(\sum_{i=n+1}^{\infty} |y_i|^2\Big)^{1/2} \to 0,$$

since the second factor is a tail of the converging sequence. However, $A_n \overset{s}{\not\to} 0$, since

$$\|A_n x\| = \|x\| \not\to 0.$$

Lecture 12. Reproducing Kernels and Weak Convergence: Exercises

Discussion of Self-Study Problems from the Previous Lecture

We begin with discussion of the homework from Lecture 10.

1) Let X be a normed space and $X_0 \subset X$ be a nontrivial closed subspace. Let $x \notin X_0$, and

$$\operatorname{dist}(x, X_0) := \inf_{x_0 \in X_0} \|x - x_0\| = d > 0.$$

Show that

$$\exists f \in X^*,\ \|f\| = 1: \quad f(x) = d, \quad f\Big|_{X_0} = 0.$$

$f(x) = d$ is a hint for constructing a functional. We will construct an extension of this functional to the space $X_1 = \langle x, X_0 \rangle = \{y = x_0 + \alpha x,\ x_0 \in X_0,\ \alpha \in \mathbb{C}\}$ such that

$$f_1(y) = f_1(x_0 + \alpha x) = f_1(x_0) + \alpha f_1(x) = \alpha d,$$

as $f_1(x_0) \equiv f(x_0) = 0$. For the norm of that functional, if $\alpha \neq 0$, we have

$$\|f_1\|_{X_1^*} = \sup_{y \neq 0} \frac{|f(y)|}{\|y\|} = \sup_{x_0 \neq 0} \frac{|\alpha| d}{\|x_0 + \alpha x\|} = \sup_{x_0 \neq 0} \frac{d}{\|\frac{x_0}{\alpha} + x\|} = \frac{d}{\inf\limits_{x_0 \neq 0} \|\frac{x_0}{\alpha} + x\|} = \frac{d}{d} = 1.$$

Then, f is an extension of f_1 obtained by the Hahn–Banach theorem.

2) Let X be a Banach space. Prove that if X^* is separable, then X is separable as well. By the definition of a separable space,

$$\exists \{f_k\}_{k=1}^{\infty}: \quad \overline{\{f_k\}_{k=1}^{\infty}} = X^*.$$

Then one can claim that

$$\forall k \in \mathbb{N}: \quad \exists x_k \in X,\ \|x_k\| = 1, \quad |f_k(x_k)| \geqslant \frac{\|f_k\|}{2}.$$

(Since the norm is the supremum over the unit sphere, there exists elements that gives at least half the norm.)

Consider

$$X_0 = \Big\{x = \sum_{k=1}^{N} c_k x_k,\ n \in \mathbb{N},\ c_k \in Q \text{ for } \mathbb{R},\ \text{or } \alpha_k + i\beta_k,\ \alpha_k, \beta_k \in \mathbb{Q} \text{ for } \mathbb{C}\Big\}.$$

It is a countable set. Let us show that $\overline{X}_0 = X$ by contradiction.

Let $\overline{X}_0 \neq X$; then it is a closed nontrivial subspace. By the previous problem,

$$\exists f \in X^*,\ \|f\| = 1: \quad f\Big|_{X_0} = 0.$$

Since $\{f_k\}_{k=1}^{\infty} = X^*$, there exists a subsequence $\{f_{k_n}\}_{n=1}^{\infty}$ such that

$$f_{k_n} \to f.$$

Further,

$$\|f - f_{k_n}\| \geqslant |(f - f_{k_n})(x_{k_n})\| = |f_{k_n}(x_{k_n})| \geqslant \frac{\|f_{k_n}\|}{2} \to \frac{1}{2},$$

and, since the norm is a continuous function,

$$f_{k_n} \to f \quad \Rightarrow \quad \|f_{k_n}\| \to \|f\|.$$

We showed that the distance between f and f_{k_n} tends to $1/2$ and $f_{k_n} \to f$, which is incompatible. Therefore, $\overline{X}_0 = X$.

3) $f(x) = x^{\alpha} \sin \frac{1}{x}$. For which α does f belong to $BV[0,1]$?

The idea is simple if one depicts these functions. For $\alpha > 1$, there is a pair of parabolas that bound the function from above and below, see Fig. 12.1.

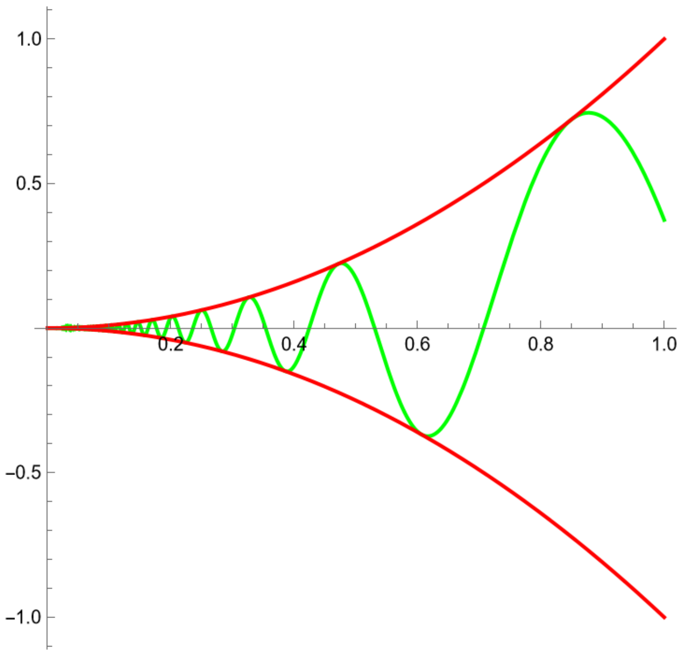

Рис. 12.1. $f(x)$ (green) is bounded by a pair of parabolas (red).

If $\alpha < 0$, the function is not bounded, and, therefore, have an infinite total variation. If $\alpha \in (0,1]$, then there are two parabolas that bound the function and have reverse convexity, see Fig. 12.2.

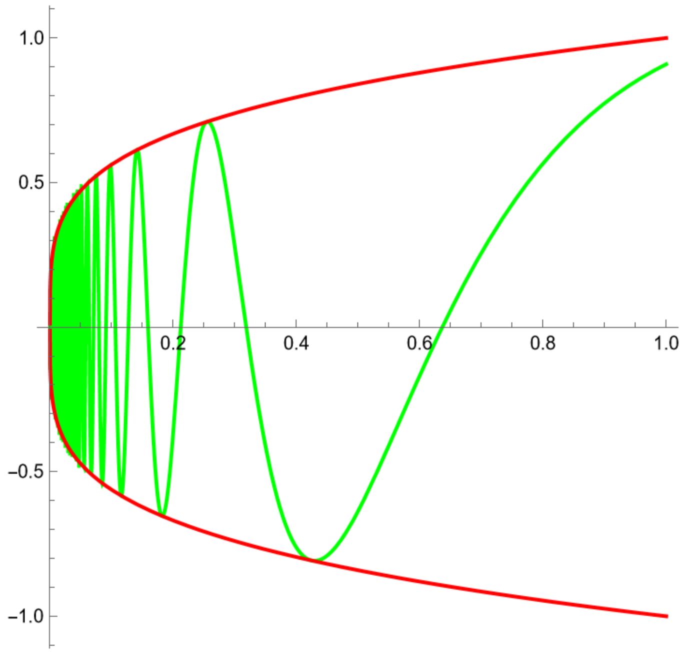

Рис. 12.2. $f(x)$ (green) is bounded by a pair of parabolas (red).

In this case, the oscillation is larger, and the total variation is infinite. We will show it now.

Take

$$x_n' = \frac{1}{\frac{\pi}{2}+2\pi n}, \quad x_n'' = \frac{1}{-\frac{\pi}{2}+2\pi n}, \quad n = 1,2,\dots,$$

and calculate the variation for these points (it is less than the total variation). Denote the partition in these points by T; then

$$V_T f = \sum_{n=1}^{\infty} \frac{1}{\left(\frac{\pi}{2}+2\pi n\right)^{\alpha}} + \frac{1}{(-\frac{\pi}{2}+2\pi n)^{\alpha}} = \infty.$$

For the case $\alpha > 1$, unfortunately, these points do not represent the maximums and minimums of f. Let us find them. Solve $f'(x) = 0$:

$$\alpha x^{\alpha-1} \sin\left(\frac{1}{x}\right) - x^{\alpha-2} \cos\left(\frac{1}{x}\right) = 0 \quad \Leftrightarrow \quad \tan\left(\frac{1}{x}\right) = \frac{1}{\alpha x}.$$

Substituting $t = 1/x$, we arrive at $\tan t = t/\alpha$. In Fig. 12.3, one can see that $t_n \asymp \frac{\pi}{2} + \pi n$ for large n, and, therefore, for x_n, we have a similar series (although in this case, it is a series of asymptotic values).

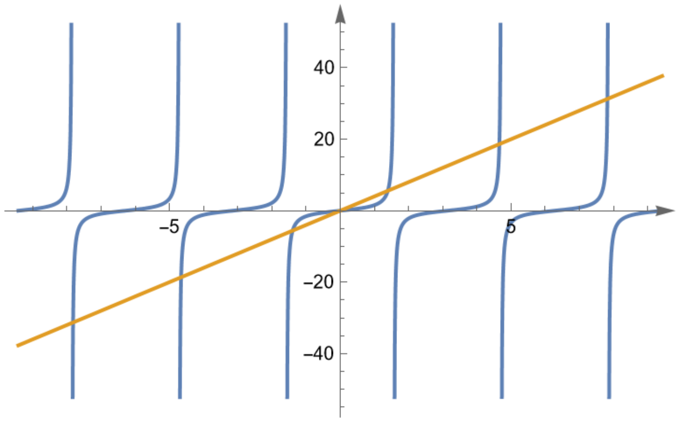

Рис. 12.3. Graphs of $\tan t$ (blue) and t/α (orange).

For these values, the series converges, so the function has finite variation.

4) In $C[0,1]$, consider

$$X_0 = \left\{ f \in C[0,1] : \int_0^1 f(t)\,dt = 0 \right\}.$$

Find $\mathrm{dist}(f_0, X_0)$, $f_0(x) \equiv 1$.

The next problem provides a way to an solve this one. Take a functional

$$F(f) = \int_0^1 f(t)\,dt, \quad F \in (C[0,1])^*.$$

Then $X_0 = \mathrm{Ker}\, f$, so

$$\mathrm{dist}(f_0, X_0) = \frac{|F(f_0)|}{\|F\|} = 1.$$

Let us show it:

$$\left| \int_0^1 f(t)\,dt \right| \leqslant \int_0^1 |f(t)|\,dt \leqslant \|f\|.$$

For $f_0(x) \equiv 1$, $F(f) = 1$. Therefore, our previous calculation is confirmed, and the answer is 1. In the derivations, we used the results of the next problem, so now we must solve it as well.

5) Let X be a normed space, $X_0 = \mathrm{Ker}\, f$, $f \in X^*$. Prove that $\mathrm{dist}(x, \mathrm{Ker}\, f) = \frac{|f(x)|}{\|f\|}$.

Consider $x^* \notin X_0$; then

$$\mathrm{dist}(x^*, X_0) = \frac{f(x^*)|}{\|f\|}.$$

Now, write out two inequalities. First, $|f(x^*)| = |f(x^* - x_0)|\ \forall x_0 \in X_0$. Then

$$|f(x^* - x_0)| \leqslant \|f\| \cdot \|x^* - x_0\| \quad \Leftrightarrow \quad \|x^* - x_0\| \geqslant \frac{|f(x^*)}{\|f\|}.$$

Now we must obtain the inverse inequality. Take $\varepsilon > 0$; $\exists z$:

$$|f(z)| \geqslant \frac{\|f\|}{1+\varepsilon}.$$

From this, we construct another element: $\exists w$ such that $f(w) = 1$:

$$w = \frac{z}{f(z)}, \quad \|w\| \leqslant \frac{1+\varepsilon}{\|f\|}.$$

Consider $y = x - wf(x)$, and evaluate the functional f at this element:

$$f(y) = f(x - wf(x)) = f(x) - f(w)f(x) = f(x) - f(x) = 0 \quad \Rightarrow \quad y \in X_0.$$

Now find the distance between y and x:

$$\|y - x\| = \|w\| \cdot |f(x)| \leqslant 1 + \varepsilon \frac{|f(x)|}{\|f\|}.$$

Taking the infimum with respect to y, we obtain

$$\operatorname{dist}(x, X_0) \leqslant \|y - x\| = \|w\| \cdot |f(x)| \leqslant (1+\varepsilon)\frac{|f(x)|}{\|f\|}.$$

In the limit as $\varepsilon \to 0$, we obtain the inverse inequality, so

$$\operatorname{dist}(x, X_0) = \frac{|f(x)|}{\|f\|}.$$

Exercises on Reproducing Kernels and Weak Convergence

1) In $W_2^1[-1,1] = \{f \in AC[-1,1],\ f' \in L_2[-1,1]\}$, consider the functional

$$F(f) = f(a), \quad a \in [-1,1].$$

Find

a) g_a: $F(f) = (f, g_a)_{W_2^1}$,

b) Reproducing kernel $K(a,b) = (g_a, g_b)$,

c) $\|F\|$.

For simplicity, consider the problem for the Sobolev space over $\mathbb{K} = \mathbb{R}$, since here one can omit the annoying conjugation that does not affect the core idea of the solution, although complicates the calculations.

Consider

$$F(f) = (f, g_a) \equiv (f, g) = \int_{-1}^{1} f(x)g(x)\,dx + \int_{-1}^{1} f'(x)g'(x)\,dx$$

(we will omit the index a keeping it in mind). Let us assume that g has the second derivative. The whole idea of evaluation the function at some point through the integration is based on th integration by parts. So, we decompose the second integral into two

$$\int_{-1}^{1} f'(x)g'(x)\,dx = \int_{-1}^{a} f'(x)g'(x)\,dx + \int_{a}^{1} f'(x)g'(x)\,dx$$

and integrate by parts, assigning the derivatives to g instead of f:

$$-fg'\Big|_{-1}^{a} - \int_{-1}^{a} f(x)g''(x)\,dx + fg'\Big|_{a}^{1} - \int_{a}^{1} f(x)g''(x)\,dx =$$
$$= f(a)g'(a)(a-0) - f(-1)g'(-1) + f(1)g'(1) - f(a)g'(a+0) -$$
$$- \int_{-1}^{a} f(x)g''(x)\,dx - \int_{a}^{1} f(x)g''(x)\,dx.$$

Here, we write the left and right limits for $g'(a \pm 0)$, since no one guaranties that this function is continuous: $g' \in L_2$, so there may be points of discontinuity. On the interval $(-1, a)$, one must take the left limit, while on $(a, 1)$ we take the right one.

From all this calculation, we should obtain just $f(a)$. What is the condition for the function g? The integral part must disappear; at the points ± 1, it must have vanishing derivative, for the nonintegral terms to disappear as well. Thus,

$$g(x) - g''(x) = 0 \quad \text{for} \quad x \in [-1, a] \text{ and } [a, 1],$$

and also

$$g'(1) = 0, \quad g'(-1) = 0, \quad g'(a+0) - g'(a-0) = 1.$$

For this differential equation, exponential functions are often taken as a basis. It is more convenient to take the hyperbolic sine and cosine in this case (for the boundary conditions that we have here). The hyperbolic sine vanishes at 0; so one could take the hyperbolic cosine with the shifted argument:

$$-A\cosh(x+1), \quad B\cosh(x-1).$$

For these functions, the boundary condition at ± 1 are automatically met. Additional condition, for the function g to belong to $W_2^1[-1,1]$ (and, therefore, to $AC[-1,1]$), is $g(a-0) = g(a+0)$. Thus,

$$-A\cosh(a+1) = \quad B\cosh(a-1) \equiv B\cosh(1-a), \quad -1 \leqslant a \leqslant 1.$$

For instance, we can take

$$B = \frac{A\cosh(a+1)}{\cosh(1-a)}.$$

Now, plug it into the condition for the jump of the derivative:

$$A\sinh(a+1) - B\sinh(a-1) = 1 \quad\Leftrightarrow\quad A\sinh(a+1) - \frac{A\cosh(a+1)}{\cosh(1-a)}\sinh(a-1) = 1,$$

or, equivalently,

$$\frac{A(\sinh(a+1)\cosh(1-a) + \cosh(a+1)\sinh(1-a))}{\cosh(1-a)} = 1,$$

therefore, after applying the formulas of sum for hyperbolic functions, we obtain

$$A = \frac{\cosh(1-a)}{\sinh 2}, \quad \text{and} \quad B = \frac{\cosh(a+1)}{\sinh 2}.$$

Finally, we have the complete data:

$$g(x) = \begin{cases} \dfrac{\cosh(1-a)\cosh(x+1)}{\sinh 2}, & x \in [-1,a), \\ \dfrac{\cosh(a+1)\cosh(x-1)}{\sinh 2}, & x \in [a,1]. \end{cases}$$

To write down the reproducing kernel, take g_a and g_b (recall that we have omitted the index a in $g = g_a$ that denotes the point at which we take the value of f), and then

$$K(a,b) = (g_b, g_a).$$

We know that, by Riesz representation theorem,

$$\|F\| = \|g_a\| = \sqrt{(g_a, g_a)} = \sqrt{g_a(a)} = \sqrt{\frac{\cosh(a+1)\cosh(a-1)}{\sinh 2}}.$$

Now, put here $a = 0$. Then

$$\|g_0\| = \sqrt{\frac{\cosh^2 1}{\sinh 2}} = \sqrt{\frac{\cosh^2 1}{2\sinh 1\cosh 1}} = \sqrt{\coth 1}2.$$

2) Consider in $C[0,1]$ the set of functions $f_n(t) = t^n$. What can we say about the convergence?

Consider the functional

$$F_{t_0}(f) = f(t_0) \equiv \int_0^1 f(t)\,dg$$

for the step function with a step of height 1 at $t = t_0$. Let us evaluate it at the sequence f_n:

$$F_{t_0}(f_n) = t_0^n \to 0 = \begin{cases} 0, & t_0 \in [0,1), \\ 1, & t = 1. \end{cases}$$

For weak convergence, we should have $F_t(f)$ as the right-hand side, if $f_n \rightharpoonup f$. But the function on the right-hand side is discontinuous, so $f_n \not\rightharpoonup f$, and, therefore, $f_n \nrightarrow f$.

3) Consider the same set of functions, but now in $(L_p[0,1])^* \cong L_q[0,1]$ for $1 \leqslant p < \infty$.

   a) First, suppose that $1 < p < \infty$. Then, working with f_n as with functions from $L_q[0,1]$, we get

$$\|f_n\|_{L_q} = \left(\int_0^1 t^{nq}\,dt\right)^{1/q} = \left(\frac{1}{nq+1}\right)^{1/q} \to 0 \quad \text{as} \quad n \to \infty.$$

   Thus, we have weak and $*$-weak convergence.

   b) Now suppose that $p = 1$. In this case, $\|f_n\|_{L_\infty} = 1$ since the supremum is 1. It is not obvious if f_n converges to any function from $L_\infty[0,1]$. Let us begin with the weakest convergence – $*$-weak convergence. It means that

$$f_n \in (L_1)^*.$$

   How to evaluate this at some function? Like that:

$$f_n(g) = \int_0^1 t^n g(t)\,dt. \tag{12.1}$$

   The integrand $t^n g(t)$ converges to zero: $t^n g(t) \to 0$ (almost everywhere). Further, provide an upper bound

$$|t^n g(t)| \leqslant |g(t)| \in L_1.$$

   By Lebesgue's dominated convergence theorem, for $f_n \to f$ a.e. with a bound $\exists g \in L_1$: $|f_n| \leqslant g$, $f \in L_1$ and

$$\lim_{n\to\infty} \int f_n\,d\mu = \int \lim_{n\to\infty} f_n\,d\mu \equiv \int f\,d\mu.$$

   Therefore, we have weak convergence to zero: $f_n \overset{*}{\rightharpoonup} 0$. Note that $f_n \underset{\Omega}{\to} f$ means that $\exists A$, $\mu(A) = 0$ such that $\forall x \in \Omega \backslash A$: $f_n(x) \to f(x)$. Taking the limit inside the integral in (12.1), we obtain

$$f_n(g) = \int_0^1 \lim_{n\to\infty} f_n g(t)\,dt \to 0.$$

   Thus, we have $*$-weak convergence to the zero functional: $f_n \overset{*}{\to} 0$.

   To study the weak convergence, one must take the functionals from the second dual space $F \in L_1^{**} = L_\infty^*$, and this problem is nontrivial since the structure of this space is quite complicated. Although, to prove that the weak convergence is violated, one can take a single functional.

   All $f_n(t) = t^n$ are continuous. Note that we are considering f_n as an element of L_∞. So, it is convenient to take the functional of evaluating at a point, that is, $F_{t_0}(f_n) = f_n(t_0)$ in $C[0,1]$, and construct its extension to the entire L_∞ using the Hahn–Banach theorem:

$$F_{t_0}(f_n) \longrightarrow \widetilde{F}_{t_0}(f_n) = t_0^n \to \begin{cases} 0, & t_0 \in [0,1), \\ 1, & t_0 = 1, \end{cases}$$

and the limit is not equal to $\widetilde{F}(0)$.

Self-Study Exercises

1) Consider the space $\overset{\circ}{W}{}_2^1[0,1] = \{f \in W_2^1[0,1] : \ f(0) = f(1) = 0\}$ (the Sobolev space with Dirichlet boundary conditions). Due to the boundary conditions, it is possible to prove that
$$(f,g) =_{\overset{\circ}{W}{}_2^1} \int_0^1 f'(x)\overline{g}'(x)\,dx.$$
For $a \in (0,1)$, consider $F_a(f) = f(a)$.

   a) Find $g = g_a$ such that $f(a) = (f, g_a)$.

   b) Find the norm $\|F_a\|$.

2) In the Bergman space
$$AL_2(\mathbb{D}) = \left\{ f \in \mathcal{A}(|z| < 1) : \iint\limits_{x^2+y^2<1} |f(z)|^2\,dx\,dy < \infty, \ z = x + iy \right\},$$
the inner product is given by
$$(f,g) = \iint\limits_{|z|<1} f(z)\overline{g(z)}\,dx\,dy.$$

   a) Check that $\{z^k\}_{k=0}^\infty$ is an ONS. Note that the power series for AL_2-functions uniformly converge on any compact set:
   $$f(z) = \sum_{k=0}^\infty a_k z^k.$$
   From the uniform convergence, one can derive the convergence in the integral sense, so it is the way to see that this space is complete.

   b) Consider the functional $F_{z_0}(f) = f(z_0)$. Find the norm $\|F_{z_0}\|$. Note that near the boundary, the behavior of an analytic function may be quite bad, and one can see it through this functional.

   c) Try to find the reproducing kernel.

3) Consider $f_n(t) = \sin(\pi n t)$ in $C[0,1]$. Study the convergence with respect to norm and weak convergence.

4) Consider $f_n(t) = \sin(\pi n t)$ in $(L_p[0,1])^*$. Study the convergence with respect to norm and weak convergence.

5) Consider A_r, A_ℓ, and A_α in ℓ_2. Find the adjoint operators.

6) Consider
$$(Af)(x) = \int_a^b K(x,t) f(t)\, dt$$
in $L_2[a,b]$. Find the adjoint operator.

7) Consider
$$(Af)(x) = \int_0^x f(t)\, dt$$
in $L_2[0,1]$. Find the adjoint operator.

Lecture 13. Adjoint, Self-Adjoint, and Normal Operators. Hellinger–Toeplitz Theorem

Banach Adjoint Operators

Linear Algebra usually deals with *Hilbert* adjoint operators, which we are to discuss a little later. Now we begin with the definition of the Banach adjoint operator.

Definition 13.1. *Let X, Y be Banach spaces, and $A \in B(X,Y)$. An **adjoint operator** $A' : Y^* \to X^*$ is an operator such that*

$$\forall f \in Y^* \quad \forall x \in X \quad (A'f)(x) := f(x).$$

Remark 13.1. *Banach adjoint satisfies the following properties:*

1) $A' \in \mathcal{L}(Y^*, X^*)$.

2) $A' \in B(Y^*, X^*)$. Moreover, norm of the operator coincides with the norm of its adjoint. We will prove that.

Statement 13.1. $\|A'\| = \|A\|$.

Proof. By definition:

$$\|A'\| = \sup_{\|f\|=1} \|A'f\|;$$

$(A'f)$ is a functional, so we use the norm of the dual space:

$$\sup_{\|f\|=1} \|A'f\| = \sup_{\|f\|=1} \sup_{\|x\|=1} |(A'f)(x)|,$$

which can be rewritten as

$$\sup_{\|f\|=1} \sup_{\|x\|=1} |(A'f)(x)| = \sup_{\|f\|=1} \sup_{\|x\|=1} |f(Ax)|$$

by definition of A'. Then, one can write the upper bound:

$$\sup_{\|f\|=1} \sup_{\|x\|=1} |f(Ax)| \leqslant \sup_{\|f\|=1} \sup_{\|x\|=1} \|f\| \cdot \|Ax\| = \|A\|. \tag{13.1}$$

There is only one place where we have an inequality. To prove the equality, we will use the first corollary of the Hahn–Banach theorem:

$$\forall x \neq 0 \quad \exists f \in X^* : \quad \|f\| = 1, \quad f(x) = \|x\|.$$

For $A \neq 0$ (note that $A = 0$ is trivial to consider), there exists x such that $Ax \neq 0$. For this x, there exists a functional $f \in Y^*$ with unit norm such that $f(Ax) = \|Ax\|$. Then, for this functional, we obtain an equality in (13.1), so $\|A'\| = \|A\|$. □

Consider an example of finding the adjoint operator. Typical examples of Banach adjoint operators arise in such spaces as ℓ_p, $p \neq 2$, and $C[a,b]$.

Example 13.1. *Consider*

$$A : C[0,2] \to C[0,2], \qquad (Af)(x) = \begin{cases} f(x), & x \in [0,1], \\ f(1), & x \in (1,2], \end{cases}$$

and see Fig. 13.1.

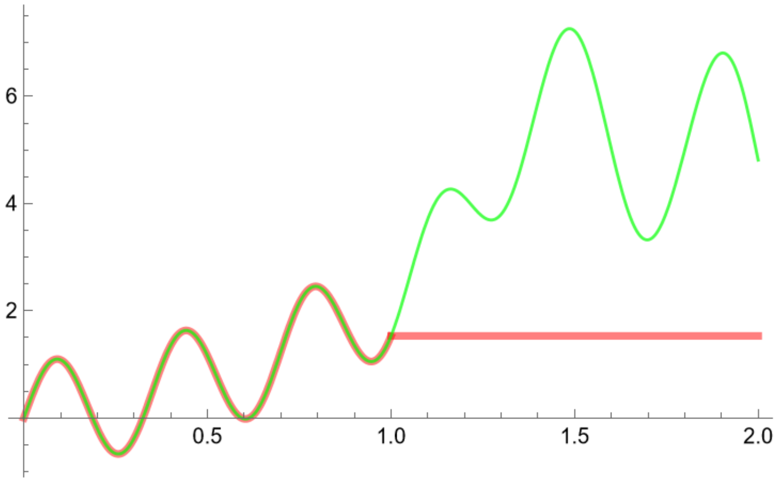

Рис. 13.1. $f(x)$ (green) and $(Af)(x)$ (red).

What is the adjoint operator? To answer this question, it is important to choose an appropriate language for description of action of the adjoint operator. It acts in the dual space, so we must construct an operator

$$A' : (C[0,2])^* \to (C[0,2])^*, \qquad (C[0,2])^* \ni G \mapsto W \in (C[0,2])^*.$$

By Riesz's theorem, these spaces are isometrically isomorphic $BV_0[0,2]$*:*

$$G(f) = \int_0^2 f(t)\,dg, \quad (C[0,2])^* \ni G \leftrightarrow g \in BV_0[0,2],$$
$$W(h) = \int_0^2 h(t)\,dw, \quad (C[0,2])^* \ni W \leftrightarrow w \in BV_0[0,2].$$

Thus, we can describe the action of A' *on functions from* $BV_0[0,2]$*. We start with* $A'G = W$*:*

$$(A'G)(f) = W(f) = \int_0^2 f(t)\,dw(t),$$

where, by the definition of the adjoint operator,

$$(A'G)(f) = G(Af) = \int_0^2 (Af)(t)\,dg(t) = \int_0^1 f(t)\,dg(t) + f(1)\int_0^1 dg(t),$$

which gives

$$(A'G)(f) = \int_0^1 f(t)\,dg(t) + f(1)\big(g(2) - g(1+0)\big).$$

Now we must obtain the image of g *under* A'. *It is clear that* $w(t) = g(t)$ *for* $t \in [0,1]$, *since we have the integration from* 0 *to* 1. *Then, we have an evaluation at the point* 1*:* $f(1)$*; so, the second term can be represented in terms of the step function with step* $g(2) - g(1+0)$. *Further, between* $t = 1$ *and* $t = 2$, *the function must be constant since there is no integration term over this interval. Thus, we obtain*

$$w(t) = \begin{cases} g(t), & t \in [0,1], \\ g(2) - g(1+0) + g(1), & t \in (1,2], \end{cases}$$

see Fig. 13.2.

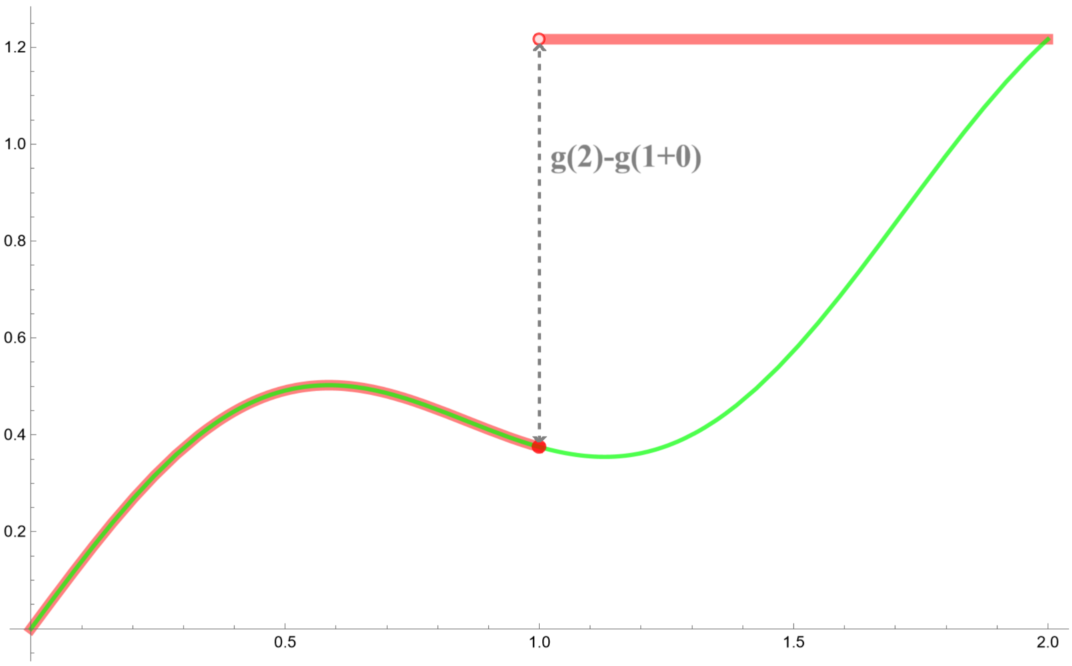

Рис. 13.2. $g(t)$ (green) and $w(t)$ (red).

This is the complete description of A'.

Hilbert Adjoint Operators

Definition 13.2. *Let* H *be a Hilbert space,* $A \in B(H)$. *Define* $A^* : H \to H$ *by*

$$(Ax, y) = (x, A^*y).$$

A^* *is called an* ***adjoint operator*** *of the operator* A.

Why does this equality define an operator? It is quite simple to explain in Linear Algebra, where one can introduce a basis, write the operator A in the matrix form, then write down this equality and see that it defines an operator A^* with a matrix, which is obtained from A by conjugate transpose. Unfortunately, this cannot be generalized to infinite-dimensional spaces. For separable spaces, one can try to describe this construction using infinite-dimensional matrices, though it cannot be applied to nonseparable spaces.

To prove that the adjoint operator is well-defined, one should use Riesz's theorem. For given A and fixed y, consider the left-hand side as a functional: $f(x) = (Ax, y)$. It is linear, and

$$|f(x)| \leqslant \|Ax\| \cdot \|y\| \leqslant \|A\| \cdot \|x\| \cdot \|y\|,$$

therefore, f is bounded. Thus, due to Riesz's theorem, there exists $z \in H$ such that $f(x) = (x, z)$, so we see that

$$(Ax, y) = (x, z).$$

The inner product is sesquilinear with respect to the second argument, but y and z are both second arguments, so z depends on y linearly; let us substitute a linear combination of y_j to the second argument:

$$(Ax, \alpha y_1 + \beta y_2) = \overline{\alpha}(Ax, y_1) + \overline{\beta}(Ax, y_2) = \overline{\alpha}(x, z_1) + \overline{\beta}(x, z_2) = (x, \alpha z_1 + \beta z_2).$$

Therefore, this construction defines a linear operator, and we put, by definition, $z := A^* y$.

Lemma 13.1. *Let* $A \in B(H)$, *where* H *is a Hilbert space. Then*

$$\|A\| = \sup_{\|x\|=\|y\|=1} |(Ax, y)|.$$

Proof.

1) In one direction, we simply write the upper bound

$$|(Ax, y)| \leqslant \|Ax\| \cdot \|y\| \leqslant \|A\| \cdot \|x\| \cdot \|y\|,$$

from which, taking the supremum over two unit spheres, we obtain

$$\sup_{\|x\|=\|y\|=1} |(Ax, y)| \leqslant \|A\|.$$

2) In the other direction, we can consider the supremum over a part of the unit sphere $\|y\| = 1$:

$$\sup_{\|x\|=\|y\|=1} |(Ax, y)| \geqslant \sup_{\|x\|=1, Ax\neq 0, y=Ax/\|Ax\|} \left(Ax, \frac{Ax}{\|Ax\|}\right) = \sup_{\|x\|=1, Ax\neq 0} \|Ax\| = \|A\|,$$

where the last equality is indeed an equality since the vectors x such that $Ax = 0$ do not contribute to the supremum. □

One can see that the order of arguments in the inner product have no influence on the value of $|(Ax, y)|$, therefore,

Theorem 13.1 (Corollary)**.** $\|A\| = \|A^*\|$.

Self-Adjoint Operators

Definition 13.3. *An operator A is called **self-adjoint** if $A = A^*$.*

This notion is quite important, especially in Quantum Mechanics, where observables are some self-adjoint operators, and the values of the observable are points of the spectrum of the corresponding self-adjoint operator.

One can see that self-adjoint operators can be defined only in Hilbert spaces, since the Banach adjoint acts in the dual space. There is also a minor difference between the Banach and Hilbert adjoint operators. Let us multiply the original operator by a constant. Then

$$(\alpha A)' = \alpha A', \qquad (\alpha A)^* = \overline{\alpha} A^*.$$

It is similar to substitution of variables for the tensor field, where the vector and functional components change with respect to different laws.

Example 13.2. *1) In ℓ_2, consider the operators A_ℓ, A_r. It is clear that $A_r^* = A_\ell$, $A_\ell^* = A_r$. Moreover, for a bounded operator A,*

$$A^{**} = A,$$

which is not exactly true in the case of unbounded A.

2) In ℓ_2, consider $A_\alpha x = (\alpha_1 x_1, \dots, \alpha_n x_n, \dots)$, $\alpha \in \ell_\infty$. The adjoint operator is $A_\alpha^* = A_{\overline{\alpha}}$ since

$$(A_\alpha x, y) = \sum_{k=1}^{\infty} \alpha_k x_k \overline{y}_k = \sum_{k=1}^{\infty} x_k \overline{\overline{\alpha}_k y_k} = (x, A^* y).$$

One can see that A_α is self-adjoint iff the sequence α is real-valued.

Definition 13.4. *Let $U:\ H \to H$. U is called a **unitary operator** if*

1) U is bijection,

2) $\forall x, y \in H$: $(Ux, Uy) = (x, y)$.

For example, A_r is not unitary since it is not a bijection. Although, in two-sided ℓ_2, that is $\ell_2(\mathbb{Z})$, both A_r and A_ℓ are unitary.

Jumping ahead, a bijective bounded operator has a bounded inverse. One can see that

$$(Ux,Uy) = (x,U^*Uy) = (x,y),$$

where the equality holds for **any** x and y. so $U^*U = I$; therefore, $U^{-1} = U^*$.

The inverse operator for A_α, $\alpha_k \neq 0\ \forall k$, is $A_{1/\alpha}$. Therefore, for A_α to be unitary, it is necessary that $\overline{\alpha}_k = 1/\alpha_k$, which means $|\alpha_k| = 1$: $\alpha_k = e^{i\theta_k}$.

Now, let us move on to projections. Let X be a Banach space and $X_0 \subset X$ be a closed subspace. Suppose that there exists X_1 (note that it is not unique) such that $X = X_0 \oplus X_1$ (note that X_1 is closed as well). Then, any $x \in X$ can be decomposed into $x = x_0 + x_1$, $x_j \in X_j$.

Definition 13.5. *P is a **projection operator** onto X_0 along X_1 if $Px = x_0$.*

This is a geometric definition of the projection. One can also give an algebraic one in the following way:

$$P^2 = P,$$

so the projection is an idempotent operator.

Let us provide an example where X_1 is not unique. Consider $X = \mathbb{R}^2$, $X_0 = \langle(1,0)\rangle$. Then one can see that $X_1 = \langle(0,1)\rangle$ and $X_1' = \langle(1,1)\rangle$ are both closed, and, for both of them, the sums are direct: $X = X_0 \oplus X_1 = X_0 \oplus X_1'$.

We can consider this construction in a Hilbert space as well. A Hilbert space has an additional geometric structure, represented by orthogonality. So, it is possible to consider orthogonal projections.

Theorem 13.2. *Let H be a Hilbert space, and $H = H_0 \oplus H_1$, where H_i are closed. Let P be a projection onto X_0 along X_1. Then*

$$H_0 \perp H_1 \iff P = P^*.$$

Proof. Note that $P^2 = P$ since P is a projection. Note also that $I - P$ is a projection onto H_1 along H_0:

$$x = x_0 + (x - x_0) = Px + (I-P)x.$$

We will first prove the theorem in $\Leftarrow$ direction. Let $Px = x_0 \in H_0$, $(I-P)y = y_1 \in H_1$. We have to prove that $(x_0,y_1) = 0$. (x_0,y_1) can be rewritten as $(Px,(I-P)y)$, and then we use that P is self-adjoint:

$$\big(Px,(I-P)y\big) = \big(x,P(I-P)y\big) = \big(x,(P-P^2)y\big) = 0,$$

since $P - P^2 = 0$. In $\Rightarrow$ direction, the proof is also simple. We must verify that P can be taken from the first argument to the second one in (Px, y). By the definition of P,

$$(Px, y) = (x_0, y) = (x_0, y_0 + y_1)$$

and, since $(x_0, y_1) = 0$, $(x_1, y_0) = 0$,

$$(x_0, y_0 + y_1) = (x_0, y_0) = (x_0 + x_1, y_0) = (x, Py). \quad \square$$

Normal Operators

Definition 13.6. *Let H be a Hilbert space, $A \in B(H)$. A is* ***normal*** *if $A^*A = AA^*$.*

Example 13.3. *1) $A = A^* \Rightarrow A$ is normal.*

2) $U^* = U^{-1} \Rightarrow U$ is normal.

3) A_α in ℓ_2 is normal:

$$A_\alpha^* A_\alpha = A_\alpha A_\alpha^* = A_{|\alpha|^2}.$$

4) A_r, A_ℓ are not normal:

$$A_r^* A_r = A_\ell A_r = I, \qquad A_r A_r^* = A_r A_\ell = P_{e_1^\perp},$$

where $P_{e_1^\perp} x = (0, x_2, x_3, \dots)$.

Properties of normal operators are quite close to ones for self-adjoint operators. There is an analogy of some sort: a self-adjoint operator is similar to multiplication by a real-valued function, while a normal operator is similar to multiplication by a complex-valued one.

Theorem 13.3 (Properties of Normal Operators)**.** *1) If A is normal, then $\forall \lambda \in \mathbb{C}$: $A - \lambda I$ is normal.*

2) If A is normal, then $\forall x \in H$: $\|Ax\| = \|A^*x\|$.

Proof. Property 1 is obvious. Let us prove property 2:

$$\|Ax\|^2 = (Ax, Ax) = (A^*Ax, x) = (AA^*x, x) = (A^*x, A^*x) = \|A^*x\|^2. \quad \square$$

Quadratic Form Associated to an Operator

Definition 13.7. *Let $A \in B(H)$, where H is a Hilbert space. The form (Ax,x) is called a* ***quadratic form associated to an operator*** *A.*

For an arbitrary operator, this form is quite useless, though it has a lot of applications in case the operator is self-adjoint. It is clear that for $A = A^*$ the form (Ax,x) is real-valued $\forall x \in H$ since $(Ax,x) = (x,Ax) = \overline{(Ax,x)}$.

Recall that the norm of an operator can be represented in the form of supremum of $|(Ax,y)|$ over two unit spheres. For self-adjoint operators, one can find the norm via taking supremum of the quadratic over a single sphere:

Theorem 13.4. *Let $A = A^*$ in a Hilbert space H. Then*

$$\|A\| = \sup_{\|x\|=1} |(Ax,x)|.$$

Proof. Denote the right-hand side by C:

$$C := \sup_{\|x\|=1} |(Ax,x)|.$$

1) For any bounded operator A,

$$|(Ax,x)| \leqslant \|Ax\| \cdot \|x\| \leqslant \|A\| \cdot \|x\|^2,$$

therefore, $C \leqslant \|A\|$.

2) For $A = A^*$, consider two quadratic forms, with $x \pm y$:

$$\Big(A(x+y), x+y\Big) - \Big(A(x-y), x-y\Big) = (Ax,x) + (Ax,y) + (Ay,x) + (Ay,y) -$$
$$-(Ax,x) + (Ax,y) + (Ay,x) - (Ay,y) =$$
$$= 2(Ax,y) + 2(Ay,y) = 4\,\mathrm{Re}(Ax,y),$$

thus,

$$\mathrm{Re}(Ax,y) = \frac{1}{4}\Big\{\Big(A(x+y), x+y\Big) - \Big(A(x-y), x-y\Big)\Big\}.$$

Let us try to estimate an absolute value of this expression:

$$|\,\mathrm{Re}(Ax,y)| \leqslant \frac{1}{4}\Big\{\Big|\Big(A(x+y), x+y\Big)\Big| + \Big|\Big(A(x-y), x-y\Big)\Big|\Big\}.$$

It is clear that $|(Ax,x)| \leqslant C\|x\|^2$, where C is the supremum of the left-hand side over the unit sphere. Therefore,

$$|\,\mathrm{Re}(Ax,y)| \leqslant \frac{C}{4}\big(\|x+y\|^2 + \|x-y\|^2\big),$$

and, due to the parallelogram law for the inner product, this implies

$$|\operatorname{Re}(Ax,y)| \leqslant \frac{C}{2}\left(\|x\|^2 + \|y\|^2\right). \tag{13.2}$$

Now we must choose y in an appropriate way. First, its norm must be equal to the norm of x, for the right-hand side to be equal to $C\|x\|^2$. Second, we want (Ax,y) to be real. Taking

$$y = \frac{Ax}{\|Ax\|}\|x\|,$$

we see that $\|y\| = \|x\|$, and inequality (13.2) becomes

$$\|Ax\|\|x\| \leqslant C\|x\|^2,$$

therefore, $\|Ax\| \leqslant C\|x\|$. Taking the supremum over the unit sphere, we obtain

$$\|A\| \equiv \sup_{\|x\|=1} \|Ax\| \leqslant C. \quad \square$$

Boundedness and Weak Boundedness of Sets in Normed Spaces

Consider the so-called *uniform boundedness principle*, which will be necessary in further developments.

Theorem 13.5 (Banach–Steinhaus)**.** *Let X be a Banach space and Y be a normed space. Let $\{A_\alpha\}_{\alpha\in\Lambda}$ be a family of bounded operators, $A_\alpha \in B(X,Y)$, and $\forall x \in X$: $\|A_\alpha x\| \leqslant c(x)$ with $c(x)$ independent of α. Then*

$$\sup_{\alpha\in\Lambda}\|A_\alpha\| < \infty.$$

Although the proof of this theorem is not very difficult, we will omit it.

A **bounded set** $M \subset X$, where X is a normed space, can be defined as follows:

$$\exists C > 0: \quad \forall x \in M \quad \|x\| \leqslant C.$$

A set M is called a **weakly bounded set**, if

$$\forall f \in X^*: \quad \forall x \in M: \quad |f(x)| \leqslant C(f).$$

Note that the bound on the right-hand side depends only on f and is independent of $x \in M$.

A surprising fact is that there is no difference between these two concepts:

1) It is obvious that a bounded set is weakly bounded:

$$|f(x)| \leqslant \|f\| \cdot \|x\| \leqslant C \cdot \|f\| \equiv \widetilde{C}(f).$$

2) In the opposite direction,

Statement 13.2. *A weakly bounded set is bounded.*

Proof. Consider a family of functionals $F_x : X^* \to \mathbb{C}$, $x \in M$, $F_x \in X^{**}$. The action of these functionals is defined by the canonical embedding:

$$\forall f \in X^* : \quad F_x(f) = f(x).$$

By Corollary 3 of the Hahn–Banach theorem,

$$\|F_x\| = \|x\|,$$

therefore, $\forall x \in M$ F_x is bounded. By weak boundedness of M, one can write

$$|F_x(f)| \equiv |f(x)| \leqslant C(f),$$

where the right-hand side is independent of x. By the Banach–Steinhaus theorem, we conclude

$$\sup_{x \in M} \|F_x\| < \infty,$$

where the left-hand side is equal to $\|x\|$, so M is bounded. □

Hellinger–Toeplitz Theorem

Consider typical operators from Quantum Mechanics, more precisely, the position and momentum operators. An interesting fact is that these operators are unbounded in L_2. In further lectures, we will see that symmetric unbounded operators must have some domain (a subset of the entire Hilbert space where it is well-defined). For now, consider the position operator

$$A : L_2(\mathbb{R}) \to L_2(\mathbb{R}), \quad (Af)(x) = xf(x).$$

For $f \in L_2(\mathbb{R})$, the function $xf(x)$ may not belong to $L_2(\mathbb{R})$, so one must define domain of the operator A:

$$\mathcal{D}(A) = \{f \in L_2(\mathbb{R}) : \quad xf \in L_2(\mathbb{R})\}.$$

Now, it is time to formulate the following theorem:

Theorem 13.6 (Hellinger–Toeplitz)**.** *Let H be a Hilbert space, $A \in \mathcal{L}(H)$, and $\forall x, y \in H$:*

$$(Ax, y) = (x, Ay).$$

Then A is bounded.

Note that in example above, A is symmetric since x is a real-valued function:

$$(Af,g) = \int_{\mathbb{R}} xf(x)\overline{g(x)}\,dx = \int_{\mathbb{R}} f(x)\overline{xg(x)}\,dx = (f,Ag).$$

This operator is also unbounded. Therefore, due to the Hellinger–Toeplitz theorem, it cannot be defined on the entire space $L_2(\mathbb{R})$, so it has some domain.

Proof (of the Hellinger–Toeplitz Theorem) by contradiction. Let A be unbounded. Then

$$\exists x_n: \quad \|x_n\| = 1, \quad \|Ax_n\| \geqslant n.$$

Consider functionals

$$f_n(x) = (Ax, x_n).$$

One can see that $|f_n(x)| = |(x, Ax_n)| \leqslant \|x\| \cdot \|Ax_n\|$, therefore, f_n is bounded: $\|f_n\| \leqslant \|Ax_n\|$ (while the bound depends on n). Using the symmetry of A, we can obtain another bound:

$$|f_n(x)| = |(Ax, x_n)| \leqslant \|Ax\| \cdot \|x_n\| = \|Ax\|$$

with a bound independent of n. Then, by the Banach–Steinhaus theorem, this family is uniformly bounded:

$$\sup_n \|f_n\| < \infty.$$

At the same time,

$$f_n\Big(\frac{Ax_n}{\|Ax_n\|}\Big) = \Big(\frac{Ax_n}{\|Ax_n\|}, Ax_n\Big) = \|Ax_n\| \geqslant n.$$

Therefore, the family is not bounded, which gives us a contradiction. □

In further, when we will proceed to studying unbounded operators with more depth, we will consider the operator

$$Af = -if',$$

$f \in L_2[0,1]$. This operator is unbounded since, being applied to $\sin \pi nx$, $\|\sin nx\| = \sqrt{2}$, it gives $\|A \sin \pi nx\| = \pi n\sqrt{2}$. What is the domain of this operator? The most natural one is the Sobolev space:

$$\mathcal{D}(A) = \big\{ f \in W_2^1[0,1],\ f(0) = f(1) = 0 \big\}.$$

Consider the following inner product:

$$(Af,g) = \int_0^1 -if'(x)\overline{g(x)}\,dx = -if(x)\overline{g(x)}\Big|_0^1 + \int_0^1 if(x)\overline{g'(x)}\,dx = (f,Ag).$$

In fact, this operator is not self-adjoint since the condition $f(0) = f(1) = 0$ is very restrictive, and the domain of the self-adjoint operator must be broader.

Lecture 14. Adjoint Operators: Exercises

Discussion of Self-Study Problems from the Previous Lecture

We will begin with discussion of the self-study problems from Lecture 12.

1) Consider the space $\overset{\circ}{W}{}_2^1[0,1] = \{f \in W_2^1[0,1]: \ f(0) = f(1) = 0\}$ (the Sobolev space with Dirichlet boundary conditions). Consider a functional $F_a(f) = f(a)$, $a \in (0,1)$. By Riesz's theorem, $f(a) = (f, g_a)$. The aim is to find the function g_a, to find the norm of F_a, and to find the reproducing kernel.

Note that

$$(f, g_a) = \int_0^1 f'(x)\overline{g_a'(x)}\,dx.$$

The idea is to use the integration by parts. The catch is that, in this case, the existence of higher derivatives of the function g_a is required. Nevertheless, it is the only simple way to find g_a, so we will try it anyway. First, decompose the integral into the sum of two, and then integrate by parts, taking the boundary conditions into account:

$$\int_0^1 f'(x)\overline{g_a'(x)}\,dx = \int_0^a f'(x)g_a'(x)\,dx + \int_a^1 f'(x)g_a'(x)\,dx =$$
$$= f(a)g_a'(a-0) - \int_0^a f(x)\overline{g_a''(x)}\,dx - f(a)g'(a+0) - \int_a^1 f(x)\overline{g_a''(x)}\,dx.$$

Thus, we must impose the following conditions for g_a:

$$g_a''(x) = \begin{cases} 0, & x \in [0,a), \\ 0, & x \in (a,1], \end{cases} \qquad g_a'(a-0) - g_a'(a+0) = 1, \qquad g_a(a-0) = g_a(a+0).$$

where the conditions for the second derivative are considered independently on given intervals (g_a is expected to be piecewise linear). Let us substitute g_a of the form

$$g_a(x) = \begin{cases} Ax, & x \in [0,a), \\ B(1-x), & x \in (a,1]. \end{cases}$$

This function automatically satisfies the boundary conditions. The conditions for g_a and g_a' allows one to find A and B. Substituting the continuity condition, we get

$$Aa = B(1-a) \quad \Rightarrow \quad B = \frac{Aa}{1-a}.$$

The condition for the first derivative, with $g_a'(a-0) = A$ and $g_a'(a+0) = -B$, gives

$$A + B = 1, \quad \Leftrightarrow \quad A + \frac{Aa}{1-a} = 1,$$

so $A = 1-a$, $B = a$. Whence, we finally obtain

$$g_a(x) = \begin{cases} x(1-a), & x \in [0,a), \\ a(1-x), & x \in [a,1], \end{cases}$$

where the value $x = a$ is included into the second interval (note that g_a is continuous, so, in fact, it does not matter where to include it).

By Riesz's theorem, $\|F_a\| = \|g_a\|$. So,

$$\|F_a\| = \|g_a\| \equiv \sqrt{(g_a, g_a)} = \sqrt{g_a(a)} = \sqrt{a(1-a)}.$$

The greatest possible value of this norm is 1/2.

By definition,

$$K(a,b) = (g_b, g_a) = g_b(a).$$

For convenience, we write out the formula for g_b:

$$g_b(x) = \begin{cases} x(1-b), & x \in [0,b), \\ b(1-x), & x \in [b,1], \end{cases}$$

and, using this, write $K(a,b)$:

$$K(a,b) = \begin{cases} a(1-b), & a < b, \\ b(1-a), a > b. \end{cases}$$

This is exactly the reproducing kernel of this space. With this kernel, one can consider another Hilbert space, where inner product has the kernel function as weight.

2) In the Bergman space

$$AL_2(\mathbb{D}) = \left\{ f \in \mathcal{A}(|z| < 1) : \iint\limits_{x^2+y^2<1} |f(z)|^2 \, dx\, dy < \infty,\ z = x+iy \right\},$$

inner product is given by

$$(f,g) = \iint\limits_{|z|<1} f(z)\overline{g(z)}\, dx\, dy.$$

It is a Hilbert space.

a) Consider $\{z^k\}_{k=0}^{\infty}$.

$$(z^k, z^n) = \iint\limits_{x^2+y^2<1} z^k \overline{z^n}\, dx\, dy \overset{z=re^{i\varphi}}{=} \int_0^1 \int_0^{2\pi} r^{k+n+1} e^{i(k-n)\varphi}\, dr\, d\varphi.$$

The integral of $e^{i(k-n)\varphi}$ with respect to φ over the period is equal to 0 for $k \neq n$. Thus,

$$(z^k, z^n) = \begin{cases} 0, & k \neq n, \\ \dfrac{\pi}{n+1}, & k = n, \end{cases}$$

where, for $k = n$, we have the squared norm of z^n, so

$$e_n = \sqrt{\frac{n+1}{\pi}} z^n$$

is an orthonormal basis: it is a closed system with $(e_i, e_j) = \delta_{ij}$, and Taylor series for any analytic function converges uniformly to this function on any compact subset of the given domain.

b) Consider the expansion

$$f(z) = \sum_{k=0}^{\infty} a_k z^k.$$

Multiplying and dividing each term by the norm of z^k, we obtain the Fourier series with respect to the system $\{e_k\}_{k=1}^{\infty}$:

$$f(z) = \sum_{k=0}^{\infty} a_k \sqrt{\frac{\pi}{k+1}} e_k,$$

and then, write out the Fourier series for g:

$$g(z) = \sum_{k=0}^{\infty} b_k e_k.$$

Consider the point evaluation functional:

$$F(f) = f(z_0).$$

By Riesz's theorem,

$$F(f) = f(z_0) = (f, g) = \sum_{k=0}^{\infty} a_k \sqrt{\frac{\pi}{k+1}} \overline{b}_k;$$

one can see that, due to the convergence of series for f,

$$\sum_{k=0}^{\infty} a_k \sqrt{\frac{\pi}{k+1}} \overline{b}_k = \sum_{k=0}^{\infty} a_k z_0^k.$$

From this, we can obtain that

$$b_k = \sqrt{\frac{k+1}{\pi}}\overline{z}_0^k.$$

By Parseval's identity,

$$\|F\|^2 = \|g\|^2,$$

so

$$\|F\| = \sqrt{\frac{1}{\pi}\sum_{k=0}^{\infty}(k+1)|z_0|^{2k}}.$$

We will calculate the sum using the substitution $|z_0|^2 = t < 1$ (note that the series converges uniformly in the unit ball):

$$\sum_{k=0}^{\infty}(k+1)|z_0|^{2k} = \Big(\sum_{k=0}^{\infty}t^{k+1}\Big)' = \Big(\frac{t}{1-t}\Big)' = \frac{1}{(1-t)^2}.$$

Therefore,

$$\|F\| = \frac{1}{\sqrt{\pi}(1-|z_0|^2)}.$$

One can see that as we are approaching the boundary, the norm of F tends to infinity. This is exactly why, in complex analysis, it is often necessary for a function to be not only analytic within a circle but also continuous all the way to the boundary.

c) Let us try to find the reproducing kernel $K(z,w) = (g_w, g_z)$. Let

$$g_w = \sum_{k=0}^{\infty}c_k e_k, \qquad g_z = \sum_{k=0}^{\infty}b_k e_k.$$

Then

$$K(z,w) = (g_w, g_z) = \sum_{k=0}^{\infty}c_k\overline{b}_k = \sum_{k=0}^{\infty}\frac{k+1}{\pi}(z\overline{w})^k,$$

since

$$b_k = \sqrt{\frac{k+1}{\pi}}\overline{z}_k, \qquad c_k = \sqrt{\frac{k+1}{\pi}}\overline{w}_k.$$

Further,

$$\sum_{k=0}^{\infty}\frac{k+1}{\pi}(z\overline{w})^k = \frac{1}{\pi}\sum_{k=0}^{\infty}(k+1)(z\overline{w})^k.$$

Then, using the same trick with $z\overline{w} = t$, $t < 1$, we obtain

$$K(z,w) = \frac{1}{\pi}\Big(\sum_{k=0}^{\infty}t^{k+1}\Big)' = \frac{1}{\pi(1-z\overline{w})}.$$

Exercises on Adjoint Operators

Let us find the adjoint operator for a multiplication operator:

Exercise 14.1. *A_φ: $L_2[a,b] \to L_2[a,b]$, where $\varphi \in L_\infty[a,b]$ is a certain function, and $(A_\varphi f)(x) = \varphi(x)f(x)$.*

1) Find A_φ^*.

2) When is it a self-adjoint operator?

3) When is it unitary?

1) To find A_φ^*, we will use the definition:

$$(A_\varphi f, g) = \int_a^b \varphi(x)f(x)\overline{g(x)}\,dx = \int_a^b f(x)\overline{\overline{\varphi(x)}g(x)}\,dx,$$

thus,

$$A_\varphi^* g = \overline{\varphi(x)}g(x) = A_{\overline{\varphi}}.$$

2) It is clear that this operator is self-adjoint iff the function φ is real-valued almost everywhere: $(A_\varphi = A_\varphi^* \equiv A_{\overline{\varphi}}) \Leftrightarrow (\varphi = \overline{\varphi}$ a.e.).

3) Similarly, A_φ is unitary iff $|\varphi(x)| = 1$ a.e.

Note also that the multiplication operator is normal in L_2 for any $\varphi \in L_\infty$.

Consider a slightly more difficult problem:

Exercise 14.2. *A_φ: $C[0,1] \to C[0,1]$, $(A_\varphi f)(x) = f(0)\cdot x + \int_0^x f(t)\,dt$.*

Find the Banach adjoint operator A'.

Recall that $A' : (C[0,1])^* \to (C[0,1])^*$. As before, we will describe the action of this operator on the space $BV_0[0,1]$ instead of $(C[0,1])^*$, since there is a one-to-one correspondence between the functions from these spaces.

For convenience, decompose A into $A_1 + A_2$, where

$$(A_1 f)(x) = f(0)\cdot x, \qquad (A_2 f)(x) = \int_0^x f(t)\,dt.$$

It is clear that $(A+B)' = A' + B'$, and similarly, for Hilbert adjoint operators, $(A+B)^* = A^* + B^*$. For the composition of operators, we have

$$(A_1A_2\ldots A_n)^* = A_n^*A_{n-1}^*\ldots A_1^*. (A_1A_2\ldots A_n)^* = A_n^*A_{n-1}^*\ldots A_1^*. (A_1A_2\ldots A_n)^* = A_n^*A_{n-1}^*\ldots A_1^*.$$

First, we will find A_1':

$$A_1' : G_1 \mapsto W_1, \qquad (A_1', G_1)(f) = W_1(f) = \int_0^1 f\, dw_1.$$

By definition,

$$G_1(A_1 f) = \int_0^1 f(0) t\, dg_1$$

Comparing the right-hand sides of these equalities, we can guess the action of the operator. Let us first equate the right-hand sides:

$$f(0) \int_0^1 t\, dg_1 = \int_0^1 f\, dw_1.$$

One can see that w_1 is a step function, see Fig. 14.1.

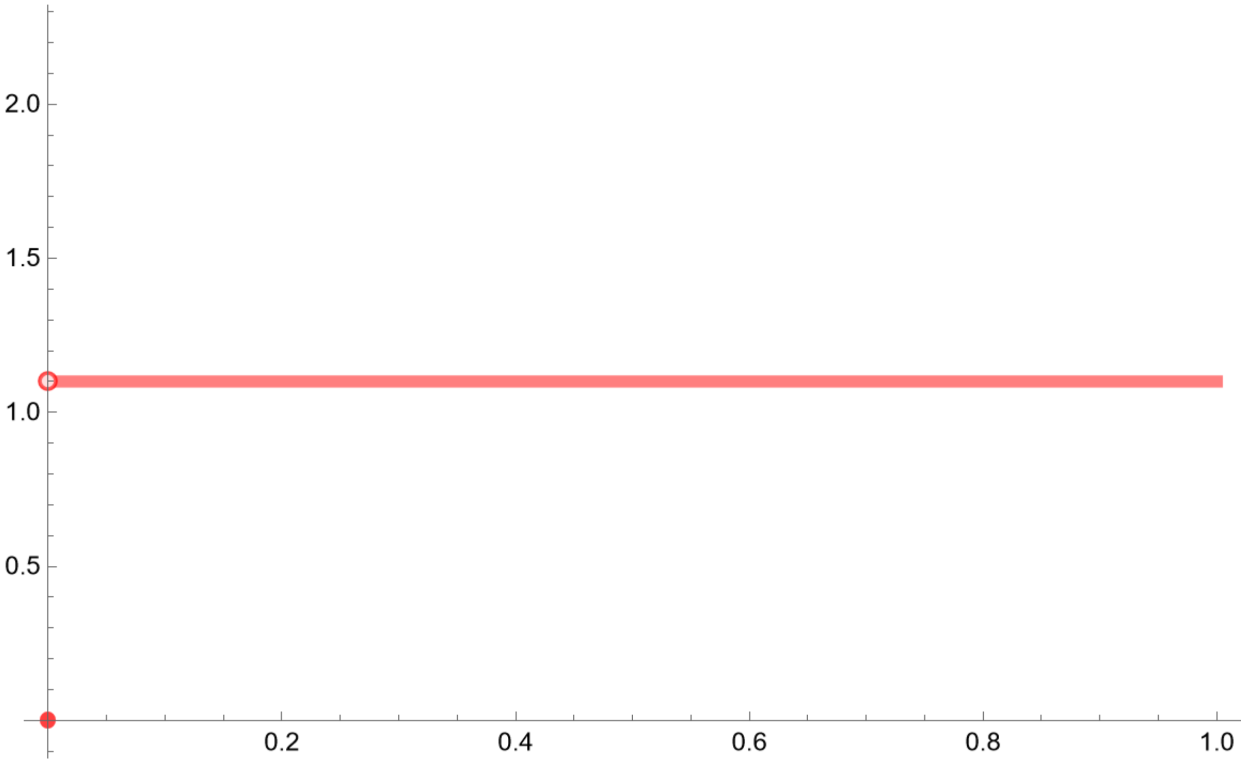

Рис. 14.1. Graph of $w_1(t)$.

The value of this jump is equal to $f(0)$.

Now let us find out how A_2 acts. One can see that

$$(A_2' G_2)(f) = W_2(f) = \int_0^1 f(t)\, dw_2,$$

and, on the other hand,

$$(A_2' G_2)(f) = G_2(A_2 f) = \int_0^1 \left(\int_0^x f(t)\, dt \right) dg_2.$$

Comparing these two formulas, we can see what is the action of A_2. Integrating the last equality by parts, we get

$$\int_0^1 \Big(\int_0^x f(t)\,dt\Big)\,dg_2 = \int_0^1 f(t)\,dt \cdot g_2(x)\Big|_0^1 - \int_0^1 f(x)g_2(x)\,dx = \\ = \int_0^1 f(t)\,dt \cdot g_2(1) - \int_0^1 f(x)g_2(x)\,dx,$$

where t can be replaced with x inside the integral:

$$\int_0^1 f(t)\,dt \equiv \int_0^1 f(x)\,dx \quad \Rightarrow \quad \int_0^1 f(t)\,dt \cdot g_2(1) - \int_0^1 f(x)g_2(x)\,dx \equiv \\ \equiv \int_0^1 f(x)\,dx \cdot g_2(1) - \int_0^1 f(x)g_2(x)\,dx,$$

so one can rewrite it as a single integral

$$\int_0^1 f(x)\,dx \cdot g_2(1) - \int_0^1 f(x)g_2(x)\,dx = \int_0^1 f(x)\,d\Big(g_2(1)x - \int_0^1 g(t)\,dt\Big).$$

This implies that

$$w_2(x) = g_2(1)x - \int_0^x g_2(t)\,dt.$$

Thus, finally,

$$A'g = w_1 + w_2 = g(1)x - \int_0^x g(t)\,dt + \int_0^1 t\,dg \cdot \chi_{(0,1]}.$$

Self-Study Exercises

1) In $L_2[0,1]$, consider
$$(Af)(x) = \int_0^x K(x,t)f(t)\,dt.$$
Find A^*. The answer must be written as $(A^*g)(x)$.

2) Apply the results of the previous problem to the following operator in $L_2[0,1]$
$$(Af)(x) = \int_0^x f(t)\,dt$$
to find A^*.

3) In $C[0,1]$, consider
$$(Af)(x) = x^2 f(0) + x\int_0^1 f(t)\,dt + f(1).$$
Find A'.

4) Let $AB - BA = I$ in a Banach space X. (Consider, e.g., $A = d/dx$, $Bf = xf$, then $AB - BA = I$.) Prove that at least one of operators A, B is unbounded.

   The results of this exercise demonstrate that Quantum Mechanics is a complicated field of study, as it inevitably deals with unbounded operators. For example, a relation of this kind, up to a constant factor, holds for the position and momentum operators. This relation is known as the Heisenberg uncertainty principle in Quantum Mechanics.

Lecture 15. Compact Operators. Inverse Operator

Compact operators. Set of Compact Operators $C(X,Y)$. Properties of Compact Operators

Definition 15.1. *Let X, Y be Banach spaces, and $A \in B(X,Y)$. A is called* **compact** *if, for any bounded set $M \subset X$, the image $AM = \{Ax,\ x \in M\}$ is precompact in Y.*

Recall that in infinite-dimensional spaces, there exist bounded sets that are not precompact; the unit ball is the standard example of such a set. Compact operators, in contrast, have the remarkable property of "compressing" bounded sets, transforming them in a way that resembles the behavior of sets in finite-dimensional spaces, even though the setting remains infinite-dimensional.

Note that in finite-dimensional spaces, all operators are compact. This is one of the examples below:

Example 15.1. *1)* $\dim X$, $\dim Y < \infty$*; given some norm, all operators become compact.*

2) If $\dim Y < \infty$ *and $A \in B(X,Y)$, then A is compact.*

Before considering the next example, recall the definitions of range and rank of an operator:

$$\operatorname{Rn} A := \{y \in Y :\ \exists x \in X\, s.t.\, y = Ax\}, \quad \operatorname{rk} A := \dim \operatorname{Rn} A.$$

3) Let $A \in B(X,Y)$ and $\operatorname{rk} A < \infty$. *Then A is compact.*

The condition that A is bounded is necessary; there are examples of unbounded operators of rank 1.

Sometimes, the definition of a compact operator given above is not convenient, since, to establish that A is compact, one must show that it makes **any** bounded set compact. To resolve this issue, the following theorem can be employed.

Theorem 15.1. *Let X, Y be Banach spaces, $A \in B(X,Y)$. Then A is compact iff the set $AB_X[0,1]$, $B_x[0,1] = \{x \in X : \|x\| \leqslant 1\}$, is precompact in Y.*

Proof. In direction $\Rightarrow$, the proof is obvious: claims that A is compact, we see that $B_X[0,1]$ is a particular compact set.

Thus, our aim is to prove the inverse. Let M be a bounded set in X. This means that

$$\exists R > 0 : \quad \forall x \in M \quad \|x\| \leqslant R \quad (M \subset B_X[0,R]).$$

As $AB_X[0,1]$ is precompact, due to the Hausdorff criterion, $\forall\varepsilon > 0$ there exists a finite ε-net $y_1, y_2, \dots, y_m$ for $AB_X[0,1]$. The idea of the proof is to construct an ε-net for the image of an arbitrary bounded set M. This set lies inside the ball of radius R; then $Ry_1, Ry_2, \dots, Ry_m$ is an εR-net for AM:

$$\forall x \in M \quad \exists i: \quad \|Ax - Ry_i\| = R\left\|A\frac{x}{R} - y_i\right\| < R\varepsilon,$$

since $\|x/R\| < 1$. □

Definition 15.2. *$C(X,Y)$ is the* ***space of all compact operators*** *from X to Y.*

Now let us discuss the properties of compact operators.

Theorem 15.2. *Let X, Y be Banach spaces, and A, $B \in C(X,Y)$. Then*

$$\alpha A + \beta B \in C(X,Y).$$

This means that the space of compact operators is a **linear** supspace of the space of bounded operators.

Proof. Let $y_1, y_2, \dots, y_m$ be an ε-net for $AB_X[0,1]$ and $z_1, z_2, \dots, z_n$ be an ε-net for $BB_X[0,1]$. The idea is to prove that $\{\alpha y_i + \beta z_j\}_{i,j=1}^{m,n}$ is a net for $(\alpha A + \beta B)B_X[0,1]$. $\forall x \in B_X[0,1]$,

$$\|(\alpha A + \beta B)x - (\alpha y_i + \beta z_j)\| \leqslant |\alpha|\|Ax - y_i\| + |\beta|\|Bx - z_j\|,$$

and $\exists i$: $\|Ax - y_i\| < \varepsilon$, $\exists j$: $\|Bx - z_j\| < \varepsilon$; therefore,

$$|\alpha|\|Ax - y_i\| + |\beta|\|Bx - z_j\| < (|\alpha| + |\beta|)\varepsilon,$$

so $\{\alpha y_i + \beta z_j\}_{i,j=1}^{m,n}$ is an $(|\alpha| + |\beta|)\varepsilon$-net for $(\alpha A + \beta B)B_X[0,1]$. □

Theorem 15.3. *Let X, Y, Z, and W be Banach spaces, and $A \in C(X,Y)$, $B \in B(Y,Z)$, $C \in B(W,X)$. Then*

$$BA \in C(X,Z), \qquad AC \in C(W,Y).$$

In other words, this means that the composition of a bounded and a compact operator (in any order) is compact.

From the Algebra course, it is known that the space of bounded operators forms an algebra. Naturally, the space of compact operators is a subalgebra of it, as established in the previous theorem. Moreover, this theorem implies that the space of compact operators forms a two-sided ideal within the algebra of bounded operators, provided that the operators act in the same space.

Proof. Consider the set $AB_X[0,1]$; by the property of compact operators, it is a precompact set in Y. Thus, $\forall\varepsilon > 0$ there exists a finite ε-net $y_1, y_2, \dots, y_m$ for $AB_X[0,1]$. Then, one can claim that $By_1, By_2, \dots, By_m$ is $\|B\|\varepsilon$-net for $(BA)(B_X[0,1])$. Why so? Let $\|x\| \leqslant 1$, $x \in X$. Consider

$$\|BAx - By_i\|_Z \leqslant \|B\| \cdot \|Ax - y_i\|,$$

and there exists $i \in \{1, 2, \dots, m\}$ such that $\|Ax - y_i\| < \varepsilon$; therefore,

$$\|BAx - By_i\|_Z < \|B\| \cdot \varepsilon.$$

The proof for AC is simpler. $CB_W[0,1]$ is a bounded set, since C is bounded. Then $A(CB_W[0,1])$ is a precompact set in Y. □

Theorem 15.4. *Let X, Y be Banach spaces, $\{A_n\}_{n=1}^{\infty}$, $A_n \in C(X,Y)$ $\forall n$, and $A_n \to A$ with respect to norm. Then $A \in C(X,Y)$.*

Proof. Take $\varepsilon > 0$. We know that $\exists N = N(\varepsilon)$: $\forall n \geqslant N$ $\|A_n - A\| < \varepsilon$. Now, take some $n \geqslant N$. Then $A_n B_X[0,1]$ is precompact in Y, so there exists a finite ε-net $y_1, y_2, \dots, y_m$. Let us find out where A maps the elements $y_1, \dots, y_m$. Take $x \in X$, $\|x\| \leqslant 1$; then

$$\|Ax - y_i\| = \|Ax - A_n x + A_n x - y_i\| \leqslant \|Ax - A_n x\| + \|A_n x - y_i\| \leqslant \|A - A_n\| \cdot \|x\| + \|A_n x - y_i\|,$$

where $\|x\| \leqslant 1$, so $\|A - A_n\| \cdot \|x\| \leqslant \varepsilon$, and $\exists i$: $\|A_n x - y_i\| < \varepsilon$, so

$$\|Ax - y_i\| \leqslant 2\varepsilon. \qquad \square$$

The following is a concise formulation of these theorems, provided the operators act in a single space.

Theorem 15.5. *$C(X)$ is a closed two-sided ideal in $B(X)$.*

Let us give an example of an ideal in the space of $n \times n$-matrices. Let

$$M_n \ni A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

and $B \in M_n$ such that

$$B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \vdots & \vdots \\ b_{(j-1),1} & \vdots & b_{(j-1),n} \\ 0 & \dots & 0 \\ b_{(j+1),1} & \vdots & b_{(j+1),n} \\ \vdots & \vdots & \vdots \\ b_{n1} & \dots & b_{nn}, \end{pmatrix}$$

i.e., $b_{jk} = 0$ $(\forall k)$. One can see that the space of the matrices of that form is a left ideal: in BA, the j-th row vanishes as well.

Now, recall what a bounded operator does to a weakly convergent sequence.

Remark 15.1. *Let X, Y be Banach spaces, $A \in B(X,Y)$, and $x_n \rightharpoonup x$ in X. Then $Ax_n \rightharpoonup Ax$ in Y.*

To demonstrate this, one can take an arbitrary $f \in Y^*$ and prove that $f(Ax_n) \to f(Ax)$. The left-hand side is

$$f(Ax_n) = (A'f)(x_n), \quad \text{where} \quad A'f = g \in X^*, \quad A' : Y^* \to X^*.$$

Thus, since $g(x_n) \to g(x)$, where $g = A'f$, we get $(A'f)(x) = f(Ax)$ on the right-hand side.

That is, a bounded operator preserves the weak convergence. In fact, a compact operator makes the convergence stronger:

Theorem 15.6. *Let X, Y be Banach spaces, $A \in C(X,Y)$, and $x_n \rightharpoonup x$ in X. Then*

$$Ax_n \xrightarrow{\|\cdot\|} Ax.$$

Proof by contradiction. Let $Ax_n \not\to Ax$. Then

$$\exists c > 0 \quad \exists n_k \to \infty : \quad \|Ax_n - Ax\| \geqslant c.$$

We know that $x_{n_k} \rightharpoonup x$ (and also $Ax_{n_k} \rightharpoonup Ax$, since A is bounded); therefore, $\{x_{n_k}\}$ is weakly bounded. Due to the Banach–Steinhaus theorem, the set $\{x_{n_k}\}$ is bounded, thus, $\{Ax_{n_k}\}$ is precompact, that is,

$$\exists n_{k_j} \to \infty : \quad Ax_{n_{k_j}} \to y \in Y,$$

and, simultaneously,

$$Ax_{n_{k_j}} \rightharpoonup Ax.$$

Additionally, we have

$$Ax_{n_{k_j}} \rightharpoonup y,$$

since the convergence with respect to norm implies the weak convergence. If $y \neq Ax$, then, due to the corollary of the Hahn–Banach theorem,

$$\exists f \in Y^* : \quad f(y) \neq f(Ax).$$

This gives us a contradiction, since

$$Ax_{n_{k_j}} \rightharpoonup y \quad \text{and} \quad Ax_{n_{k_j}} \rightharpoonup Ax.$$

Therefore, $y = Ax$. This is the final contradiction between the condition $\|Ax_{n_k} - Ax\| \geqslant c$ and $Ax_{n_{k_j}} \to Ax$, since n_{k_j} is a subsequence of n_k. □

There are examples of bounded operators that turn weak convergence into norm convergence, but are not compact; so this theorem is not a criterion for the compactness of an operator. However, if the space if reflexive, this becomes a criterion.

Example: Integral Operators in $C[a,b]$ and $L_2[a,b]$

Why compact operators are important? They arise in many applications, including Mathematical Physics, where they appear as inverse to differential operators.

Now, we consider the following integral operator

$$(Af)(x) = \int_a^b K(x,y) f(y)\, dy.$$

Theorem 15.7. *If $K(x,y) \in C[a,b]^2$, then $A \in C(C[a,b])$.*

Note that this is a sufficient condition, but not a criterion. However, it is quite close to necessary condition: $K(x,y)$ must be continuous on $[a,b]^2$ except for a finite number of continuous curves that are graphs of continuous functions.

Proof. We have to prove that the image of the unit ball is a precompact set. Consider $AB_{C[a,b]}[a,b] \equiv A\{f \in C[a,b] : \|f\| \leqslant 1\}$. Due to the Arzelà–Ascoli theorem, this set must be bounded and uniformly equicontinuous.

First, we show that A is bounded:

$$\max_{[a,b]} |(Af)(x)| = \max_{[a,b]} \left| \int_a^b K(x,t) f(t)\, dt \right| \leqslant \max_{[a,b]} \int_a^b \Big|K(x,t)\Big| |f(t)|\, dt,$$

where $|f(t)| \leqslant \|f\| \leqslant 1$, so

$$\max_{[a,b]} \int_a^b \Big|K(x,t)\Big| |f(t)|\, dt \leqslant \max_{[a,b]} \int_a^b \Big|K(x,t)\Big|\, dt,$$

therefore, the image of the ball is bounded as well.

Now, prove the equicontinuity. Take $\varepsilon > 0$. Note that any continuous function on a compact set is uniformly continuous, so is $K(x,t)$ on $[a,b]^2$:

$$\exists \delta > 0 \quad \forall (x_1,t_1), (x_2,t_2) \in [a,b]^2,\ |x_1 - x_2| + |t_1 - t_2| < \delta \quad \Rightarrow \quad |K(x_1,t_1) - K(x_2,t_2)| < \varepsilon.$$

Let us consider $|(Af)(x_1) - (Af)(x_2)|$ and try to estimate it, given $|x_1 - x_2| < \delta$:

$$|(Af)(x_1) - (Af)(x_2)| = \left| \int_a^b K(x_1,t) f(t)\, dt - \int_a^b K(x_2,t) f(t)\, dt \right| \leqslant$$
$$\leqslant \int_a^b \Big|K(x_1,t) - K(x_2,t)\Big| |f(t)|\, dt,$$

where $|K(x_1,t)-K(x_2,t)|<\varepsilon$ and $|f(t)|\leqslant\|f\|$, so

$$|(Af)(x_1)-(Af)(x_2)|<\varepsilon(b-a),$$

therefore, $AB_{C[a,b]}[a,b]$ forms an equicontinuous family. Thus, due to the Arzelà–Ascoli theorem, $AB_{C[a,b]}[a,b]$ is precompact, so A is a compact operator. □

Theorem 15.8. *If $K(x,y)\in L_2[a,b]^2$, then $A\in C(L_2[a,b])$.*

This time, the sufficient condition is far from being the necessary one.

Proof. The idea is to construct operators A_n, $A_n\xrightarrow{\|\cdot\|}A$, such that $A_n\in C(L_2[a,b])$.

The construction is simple: let $\{\varphi_i\}_{i=1}^{\infty}$ be an orthonormal basis in $L_2[a,b]$, then

$$\{\psi_{ij}(x,t):=\varphi_i(x)\varphi_j(t)\}_{i,j=1}^{\infty}$$

is an orthonormal basis in $L_2[a,b]^2$. The function $K(x,t)$ can be expanded into the Fourier series

$$K(x,t)=\sum_{i,j=1}^{\infty}c_{ij}\psi_{ij}(x,t).$$

Consider a partial sum

$$K_n(x,t)=\sum_{i,j=1}^{n}c_{ij}\psi_{ij}(x),$$

and the corresponding operator A_n:

$$(A_nf)(x)=\int_a^b K_n(x,t)f(t)\,dt.$$

One can see that each of A_n is of finite rank:

$$(A_nf)(x)=\int_a^b\sum_{i,j=1}^{n}c_{ij}\varphi_i(x)\varphi_j(t)f(t)\,dt=\sum_{i,j=1}^{n}\varphi_i(t)\cdot\Big(\sum_{i,j=1}^{n}\int_a^b\varphi_j(t)f(t)\,dt\Big),$$

so the image consists of linear combinations of φ_i, therefore, $\mathrm{rk}A_n\leqslant n$. Further,

$$\|A_n\|\leqslant\|K_n\|_{L_2[a,b]^2}\quad\Rightarrow\quad A_n\in C(L_2[a,b]).$$

Now, let us try to estimate $\|A_n-A\|$:

$$(A_n-A)f=\int_a^b\big(K_n(x,t)-K(x,t)\big)f(t)\,dt,$$

so

$$\|A_n-A\|\leqslant\|K_n-K\|_{L_2[a,b]^2}\to 0\quad\text{as}\quad n\to\infty,$$

therefore, $A\in C(L_2[a,b])$. □

Inverse Operator

Let X, Y be linear spaces, $A \in \mathcal{L}(X,Y)$.

Definition 15.3. *An operator$A_\ell^{-1}: Y \to X$ such that $A_\ell^{-1}A = I_X$ is called a **left inverse** of an operator A. $A_r^{-1}: Y \to X$ such that $AA_r^{-1} = I_Y$ is called a **right inverse** of an operator A.*

Note that, e.g., a left inverse is not unique, and, moreover, it may be nonlinear; we will provide some examples of nonlinear inverse operators a bit later.

One can see that if there exists a left inverse, then the operator A is injective ($\text{Ker}A = \{0\}$); if there exists a right inverse, then A is surjective ($\text{Rn}A = Y$). Thus, if there are left and right inverse, the operator is a bijection; moreover, left and right inverse coincide ($A_\ell^{-1} = A_r^{-1}$). Let us show it: consider $A_\ell^{-1}AA_r^{-1}$. The compositions of operators are associative, so, inserting brackets in different ways, we get

$$(A_\ell^{-1}A)A_r^{-1} = A_r^{-1}, \qquad A_\ell^{-1}(AA_r^{-1}) = A_\ell^{-1}.$$

If $\exists A_\ell^{-1}$, A_r^{-1}, then it is denoted as $A^{-1}$and is unique.

Example 15.2. *In ℓ_2, consider*

$$A_\ell x = (x_2, x_3, \dots).$$

One can see that the image of $(1,0,0,\dots)$ vanishes, so A_ℓ has a nontrivial kernel, and, therefore, the operator is not injective. However, the image of A_ℓ is the entire space (one can reconstruct the preimage of any $y \in \ell_2$ by shifting it to the right), so A_ℓ is surjective. The right inverse is A_r:

$$A_\ell A_r = I.$$

It is not a left inverse:

$$A_r A_\ell = P_{e_1^\perp},$$

since the first coordinate in the image is always zero (so the composition is a projection onto $e_1^\perp$). Obviously, for the operator A_r, an operator A_ℓ is a left inverse.

For A_ℓ, there are other options of the right inverse operator. Consider, for instance, the following one:

$$B_a x = (a, x_1, x_2, \dots),$$

*which is **not even linear**. Then $A_\ell B_a = I$ for any a. We will show that a two-sided inverse cannot be nonlinear.*

Theorem 15.9. *Let $A \in \mathcal{L}(X,Y)$, where X, Y are linear spaces. If $\exists A^{-1}$, then $A^{-1} \in \mathcal{L}(Y,X)$.*

Proof. Let us apply the inverse to a linear combination:

$$A^{-1}(\alpha y_1 + \beta y_2) = A^{-1}(\alpha A x_1 + \beta A x_2),$$

where $y_j = Ax_j$, $\exists! x_j$, since A is bijective. A is linear, so one can rewrite it as

$$A^{-1}(\alpha A x_1 + \beta A x_2) = A^{-1}A(\alpha x_1 + \beta x_2),$$

and then, collapsing $A^{-1}A = I$, we get

$$A^{-1}A(\alpha x_1 + \beta x_2) = \alpha x_1 + \beta x_2 = \alpha A^{-1} y_1 + \beta A^{-1} y_2,$$

so, by writing the beginning and the end of the chain of equalities, we obtain

$$A^{-1}(\alpha y_1 + \beta y_2) = \alpha A^{-1} y_1 + \beta A^{-1} y_2,$$

which confirms the linearity of A^{-1}. □

Lecture 16. Exercises on Compact and Inverse Operators

Discussion of Self-Study Problems form the Previous Lecture

We begin with considering some of the self-study problems from Lecture 14.

3) In $C[0,1]$, consider

$$(Af)(x) = x^2 f(0) + x\int_0^1 f(t)\,dt + f(1).$$

Find A'.

We know that $A' : (C[0,1])^* \to (C[0,1])^*$:

$$(C[0,1])^* \ni G \mapsto W \in (C[0,1])^*, \qquad A'G = W.$$

For the functionals G, W from the dual space to $C[0,1]$, there are functions $g, w \in BV_0[0,1]$ that are in one-to-one correspondence with G and W respectively. Thus, to describe the action of A', it is sufficient to construct a function w that corresponds to a given function g.

By definition,

$$(A'G)(f) = W(f) = \int_0^1 f(t)\,dw,$$

and, on the other hand, $(A'G)(f) = G(Af)$, so

$$(A'G)(f) = \int_0^1 \Big(x^2 f(0) + x\int_0^1 f(t)\,dt + f(1)\Big)\,dg(x).$$

Let us first simplify it:

$$\int_0^1 \Big(x^2 f(0) + x\int_0^1 f(t)\,dt + f(1)\Big)\,dg(x) =$$
$$= f(0)\int_0^1 x^2\,dg(x) + \int_0^1 f(t)\,dt \cdot \int_0^1 x\,dg(x) + f(1)\int_0^1 dg(x).$$

The integral of x with respect to $dg(x)$ is independent of t; thus, one can include it as a constant factor to dt:

$$f(0)\int_0^1 x^2\,dg(x) + \int_0^1 f(t)\,dt \cdot \int_0^1 x\,dg(x) + f(1)\int_0^1 dg(x) =$$
$$= f(0)\int_0^1 x^2\,dg(x) + \int_0^1 f(t)\,d\Big(\int_0^1 x\,dg(x)\cdot t\Big) + f(1)\big(g(1) - g(0)\big),$$

where $g(0) = 0$.

Now, we must establish the behavior of $w(t)$. It is equal to 0 at $t = 0$; further, as we have $f(0)$ in the expression, it must have a step at $t = 0+0$ of height $\int_0^1 x^2\,dg(x)$. Next, the function $w(t)$ is linear until $t = 1-0$. As we have the evaluation of $f(t)$ at $t = 1$ in the expression, there is a jump of height $g(1)$. See Figure 16.1.

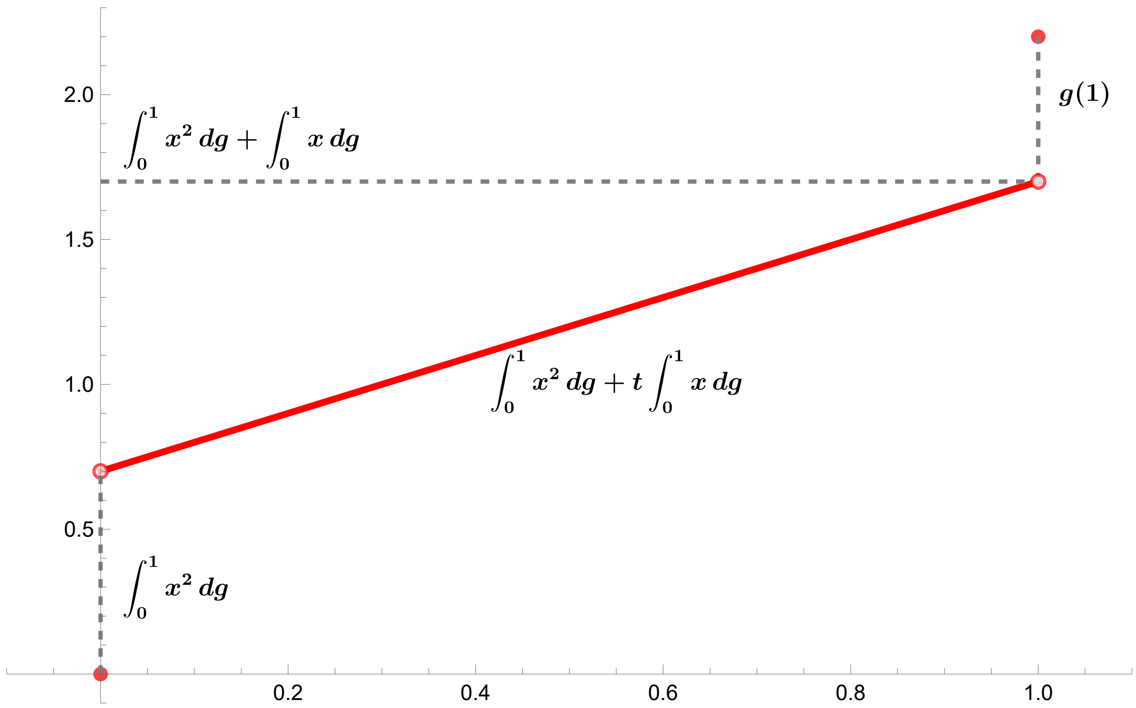

Рис. 16.1. Graph of $w(t)$.

Here is a complete description of $w(t)$:

$$w(t) = \begin{cases} 0, & t = 0, \\ \displaystyle\int_0^1 x^2\,dg + t\int_0^1 x\,dg, & t \in (0,1), \\ \displaystyle\int_0^1 x^2\,dg + t\int_0^1 x\,dg + g(1), & t = 1. \end{cases}$$

4) Let $AB - BA = I$ in a Banach space X. (Consider, e.g., $A = d/dx$, $Bf = xf$, then $AB - BA = I$.) Prove that at least one of operators A, B is unbounded.

First, consider $AB^n - B^n = nB^{n-1}$. For $n = 1$, it is the given relation $AB - BA = I$. Let us try to derive the formula for $n = 2$:

$$AB^2 - B^2A = AB\cdot B - B\cdot BA = (I + BA)B - B\cdot BA = B + B(AB - BA) = 2B.$$

Now, we will prove it by induction. Suppose the equality holds for $k = n$:

$$AB^n = nB^{n-1} + B^n A.$$

Consider it for $k = n+1$:

$$\begin{aligned} &AB^{n+1} - B^{n+1}A = AB^n B - B^{n+1}A = \\ =&(nB^{n-1} + B^n A)B - B^{n+1}A = nB^n + B^n(AB - BA) = (n+1)B^n, \end{aligned}$$

which completes the proof.

Due to this relation,

$$\|nB^{n-1}\| = \|AB^n - B^n A\| \leqslant \|AB\| \cdot \|B^{n-1}\| + \|B^{n-1}\| \cdot \|BA\|,$$

and, dividing by $\|B^{n-1}\|$, we obtain

$$n \leqslant \|AB\| + \|BA\| \leqslant 2\|A\| \cdot \|B\|$$

for any $n \in \mathbb{N}$, so at least one of these operators is unbounded.

Exercises on Compact Operators

1) In ℓ_2, consider a multiplication operator

$$A_\alpha x = (\alpha_1 x_1, \alpha_2 x_2, \dots), \qquad \alpha \in \ell_\infty.$$

We claim that

$$A_\alpha \in C(\ell_2) \quad \Leftrightarrow \quad \alpha \in c_0 \quad (\text{i.e. } \lim_{k\to\infty} \alpha_k = 0).$$

To solve exercises on compact operators, one should remember the criteria for precompactness in different spaces. In ℓ_2, the criterion is the following: $M \subset \ell_2$ is precompact iff

a) M is bounded,

b) $\forall \varepsilon > 0\ \exists n\ \forall x \in M$:

$$\Big(\sum_{k=n+1}^{\infty} |x_k|^2\Big)^{1/2} < \varepsilon.$$

The second condition means that the tails are uniformly small, or, in other words, the set is "almost finite-dimensional".

$\Rightarrow$. For the operator to be compact, we must require that the image of the unit ball is compact. Consider the basis elements

$$e_k = (0, \dots, 0, \overset{k}{1}, 0, \dots) \in B_{\ell_2}[0,1].$$

Their images $\{Ae_k\}_{k=1}^{\infty}$ must form a precompact set. One can see that

$$Ae_k = (0,\dots,0,\alpha_k,0,\dots).$$

There must exist $n \in \mathbb{N}$ such that for $k \geqslant n+1$: $|\alpha_k| < \varepsilon$, so $\alpha_k \to 0$.

$\Leftarrow$. Let $\alpha_k \to 0$. We must check that the image of $AB_{\ell_2}[0,1]$ under the action of corresponding operator is precompact.

Take $x \in \ell_2$ with $\|x\| \leqslant 1$. Then

$$\|Ax\| = \Big(\sum_{k=1}^{\infty} |\alpha_k x_k|\Big)^{1/2} \leqslant \sup_{k\geqslant 1}|\alpha_k|\Big(\sum_{k=1}^{\infty} |x_k|\Big)^{1/2},$$

and, since $\alpha \in \ell_\infty$, $AB_{\ell_2}[0,1]$ is bounded.

Now, let us verify that the tails of the elements of the image are uniformly small. Consider a partial sum

$$\Big(\sum_{k=n+1}^{\infty} |\alpha_k x_k|\Big)^{1/2}.$$

Since $\alpha \in c_0$,

$$\forall \varepsilon > 0 \quad \exists n: \quad \forall k \geqslant n+1 \quad |\alpha_k| < \varepsilon.$$

Then

$$\Big(\sum_{k=n+1}^{\infty} |\alpha_k x_k|\Big)^{1/2} < \varepsilon\Big(\sum_{k=n+1}^{\infty} |x_k|\Big)^{1/2} < \varepsilon,$$

since $\|x\| \leqslant 1$.

2) Consider

$$(Af)(x) = \int_0^x f(t)\,dt$$

a) in $C[0,1]$,

b) in $L_2[0,1]$ (later).

The operator can be written as

$$(Af)(x) = \int_0^x f(t)\,dt = \int_0^1 K(x,t)f(t)\,dt,$$

where

$$K(x,t) = \begin{cases} 1, & t < x, \\ 0, & t > x \end{cases} \quad \in \quad L_2[0,1]^2.$$

We will prove that $A \in C(L_2[0,1])$ using a theorem from the previous lecture. For $C[0,1]$, we cannot use the corresponding theorem, since $K(x,t)$ is discontinuous. However, one can show it in a straightforward way.

Let $f \in C[0,1]$, $\|f\| \leqslant 1$:

$$\|Af\| = \max_{x\in[0,1]} \left| \int_0^x f(t)\,dt \right| \leqslant \max_{x\in[0,1]} \int_0^x |f(t)|\,dt \leqslant \int_0^1 \|f\|\,dt = 1,$$

so the image of the unit ball is bounded. Now, we will check the equicontinuity:

$$\big|(Af)(x) - (Af)(y)\big| = \left| \int_y^x f(t)\,dt \right| \leqslant \left| \int_y^x |f(t)|\,dt \right| \leqslant |y-x|,$$

since $|f(t)| \leqslant \|f\|$. So, for $|y-x| < \varepsilon$, it is sufficient to take $\delta = \varepsilon$.

3) Consider A_ℓ, A_r in ℓ_2. Are these operators compact?

These operators are not compact. Let us prove it. Take the standard basis $\{e_k\}_{k=1}^\infty$. Then $A_r\{e_k\}_{k=1}^\infty = \{e_k\}_{k=2}^\infty$, and $\|e_k - e_m\| = \sqrt{2}$, $k \neq m$; therefore, there is no Cauchy subsequence. For A_ℓ, the situation is similar: $A_r\{e_k\}_{k=1}^\infty = \{e_k\}_{k=1}^\infty$. Recall that these operators are adjoint to each other. In the next section, we will consider the relation between the notions of compactness and adjointness.

Relation Between Notions of Compact and Adjoint Operators

Theorem 16.1 (without a proof)**.** *Let X, Y be Banach Spaces. Then*

$$A \in C(X,Y) \quad \Leftrightarrow \quad A' \in C(Y^*, X^*).$$

The idea of the proof is to use the Arzelà–Ascoli theorem.

The following theorem on Hilbert adjoint operators is not as difficult to prove as the previous one:

Theorem 16.2. *Let $A \in B(H)$, where H is a Hilbert space.*
If A^*A is compact, then A is compact.
If AA^* is compact, then A^* is compact.

Proof. Since the statements of the theorem are symmetric, we will prove only the first one. We must show that $AB_H[0,1]$ is precompact.

Take a sequence $\{y_k\}_{k=1}^\infty$ in $AB_H[0,1]$. By the definition of $AB_H[0,1]$,

$$\forall k \quad \exists x_k, \quad \|x_k\| \leqslant 1: \quad y_k = Ax_k.$$

Consider the set $\{A^*Ax_k\}_{k=1}^{\infty}$. It is precompact, since A^*A is a compact operator, therefore, there exists a Cauchy subsequence $\{A^*Ax_{k_n}\}_{n=1}^{\infty}$. Now, let us use these indices for the image of A:

$$\|y_{k_n} - y_{k_m}\|^2 = \big(A(x_{k_n} - x_{k_m}), A(x_{k_n} - x_{k_m})\big) = \big(A^*A(x_{k_n} - x_{k_m}), x_{k_n} - x_{k_m}\big) \leqslant \\ \leqslant \|A^*A(x_{k_n} - x_{k_m})\| \cdot \|x_{k_n} - x_{k_m}\|,$$

where $\|x_{k_n} - x_{k_m}\| \leqslant 2$ and $\|A^*A(x_{k_n} - x_{k_m})\| \to 0$ as $k_n, k_m \to \infty$. Thus, we have found a Cauchy subsequence, so the operator is compact. □

Corollary 16.1. $A \in C(H) \Leftrightarrow A^* \in C(H)$.

Proof. The composition of a bounded operator and a compact operator is compact. Suppose that A is compact. Then AA^* is compact, and, due to the theorem, A^* is compact. If A^* is compact, then we take a compact A^*A, so A is compact. □

Let us continue solving the exercises.

3) Let X be a Banach space, $\dim X = \infty$, and A be a compact operator Then there is no bounded A^{-1}.

   We will prove it by contradiction. Let there exists $A^{-1} \in B(X)$; then $AA^{-1} = I$. Since A is compact and A^{-1} is bounded, AA^{-1} is a compact operator; but the identity operator in an infinite-dimensional space is not compact since the unit ball is not a precompact space.

4) Let $\varphi \in C[a,b]$ be some certain function. Consider

$$(A_\varphi f)(x) = \varphi(x) f(x).$$

   Then

$$A_\varphi \in C(C[a,b]) \quad \Leftrightarrow \quad \varphi(x) \equiv 0.$$

   This is the simplest example of a compact operator. The proof in $\Leftarrow$ is obvious. Let us prove the inverse by contradiction using the Arzelà–Ascoli theorem.

   Let $\exists x_0$: $\varphi(x_0) \neq 0$; without loss of generality, suppose $\varphi(x_0) > 0$. Then $\exists \delta > 0$, $c > 0$: $\varphi(x) > c$ for $x \in (x_0, x_0 + \delta)$ or $x \in (x_0 - \delta, x_0)$. Let $x_0 \neq b$, and take $1/n < \delta$. Consider a sequence f_n, $\|f_n\| = 1$, see Fig. 16.2.

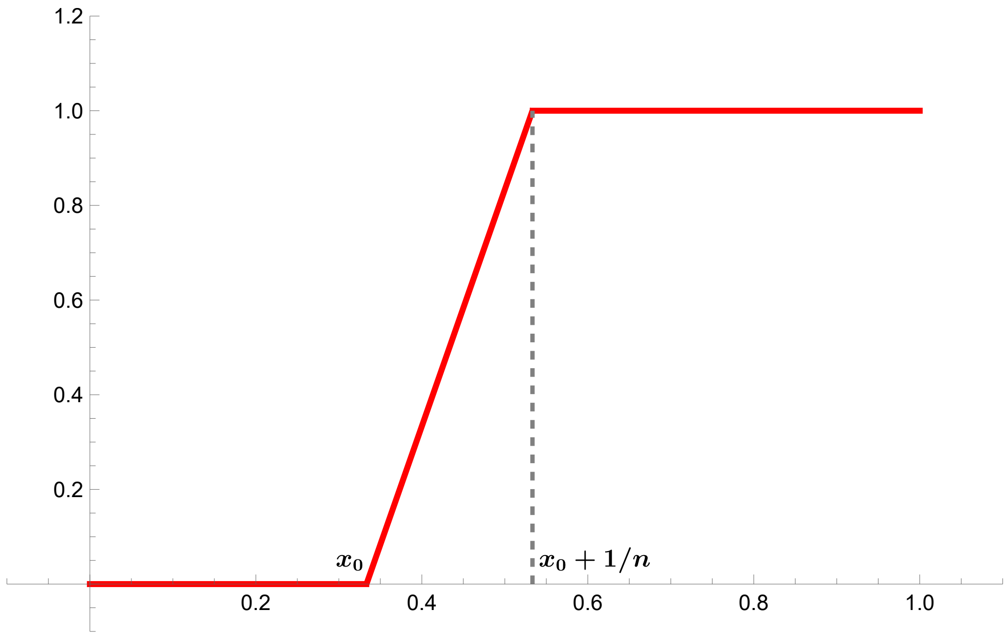

Рис. 16.2. Graph of $f_n(x)$.

$\{Af_n\}_{n=1}^{\infty}$ is precompact, therefore, it is equicontinuous; let us estimate

$$\left|\left(Af_n\right)\left(x_0+\frac{1}{n}\right)-\left(Af_n\right)(x_0)\right| \geqslant c,$$

since $\left(Af_n\right)(x_0)=0$ and $\left(Af_n\right)\left(x_0+\frac{1}{n}\right)>c$, which contradicts to the equicontinuity. Note that in $L_2[a,b]$ we will prove the same (more precisely, that multiplication operator is compact iff the corresponding function vanishes almost everywhere) later using the properties of spectrum.

Exercises on Inverse Operators

1) In $C[0,1]$, consider
$$(Af)(x)=\int_0^x f(t)\,dt.$$
Is there a right or a left inverse?

Consider the operator B, $Bf=f'$. It is obvious that $BA=I$, so $A_\ell^{-1}=B$.

Is there a right inverse? If there exists a right inverse C, $AC=I$, then A must be surjective. One can see that

$$\mathrm{Rn}A=\{g\in C^1[0,1],\ g(0)=0\},$$

so the operator is not surjective, since $\mathrm{Rn}A\neq C[0,1]$.

2) Let X be a Banach space. Prove that if $C: X \to X$, $\|C\| < 1$, then $\exists (I \pm C)^{-1}$.

If we imagine that C is just a number, not an operator, then

$$\frac{1}{1-C} = \sum_{k=0}^{\infty} C^k.$$

We claim that

$$(I-C)^{-1} = I + C + C^2 + C^3 + \dots.$$

First, we have to explain why this sum converges. Consider, for $n > m$,

$$S_n = \sum_{k=0}^{n} C^k, \quad \|S_n - S_m\| = \Big\| \sum_{k=m+1}^{n} C^k \Big\| \leqslant \sum_{k=m+1}^{n} \|C^k\| \leqslant \frac{\|C\|^{m+1}}{1-\|C\|}.$$

As $m \to \infty$, it decreases to 0; therefore, S_n is a Cauchy sequence. Thus, since $B(X,Y)$ is a Banach space when Y is Banach, there exists a limit element

$$S = \lim_{n\to\infty} S_n.$$

Let us expand the expression for S_n in $(I-C)S_n$:

$$(I-C)S_n = I + C + C^2 + \dots + C^n - C - C^2 - \dots - C^n - C^{n+1} = I - C^{n+1} \to I \quad \text{as} \quad n \to \infty,$$

Similarly, $(I+C)^{-1} = I - C + C^2 - \dots + (-1)^n C^n + \dots$.

Self-Study Exercises

1) In $L_2[0,1]$, consider the Hardy operator

$$(Af)(x) = \frac{1}{x} \int_0^x f(t)\,dt.$$

a) Prove that A is bounded.

b) Prove that A is not compact.

Hint: item a) can be solved by definition. To solve item b), one can use the property of compact operator from Lecture 15: a compact operator maps a weakly converging sequence to a sequence converging with respect to norm. So, the aim is to find an appropriate weakly converging sequence. Note that the operator seems to be bad at $x = 0$.

2) In some space, construct an operator A such that $A^2 = 0$ and A is not compact.

3) Consider A in ℓ_2 defined as an infinite matrix $A \sim (a_{ij})_{i,j=1}^{\infty}$, $(Ax)_i = \sum_{j=1}^{\infty} a_{ij}x_j$. Prove that

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}|a_{ij}|^2 < \infty \quad \Rightarrow \quad A \in C(\ell_2).$$

4) Consider the differential operator $Af = f'$ in $C[0,1]$ with domain $\mathcal{D}(A) = C^1[0,1]$. Prove that there exists a right inverse, but it is not unique.

5) Consider

$$(Af)(x) = f(x) - \int_0^x f(t)\,dt$$

   a) in $C[0,1]$.

   b) in $L_2[0,1]$.

Find the inverse operator. The answer must not involve infinite sums.

Lecture 17. Spectrum of a Bounded Operator. Classification of Points in the Spectrum

Banach Bounded Inverse Theorem

Let us continue to discuss inverse operators and property of invertibility. We begin with the Banach Bounded Inverse theorem:

Theorem 17.1 (Banach Bounded Inverse Theorem, without a proof)**.** *Let X, Y be Banach spaces, $A \in B(X,Y)$. Then*

$$\exists A^{-1} \in B(Y,X) \quad \Leftrightarrow \quad A \text{ is a bijection.}$$

It is clear that a bijection has an inverse map; it is also clear that an invertible map is a bijection. The most difficult part of this theorem is that the inverse is bounded. Moreover, under weaker assumptions, i.e. that X and Y are just some normed space (not complete), one can construct counterexamples.

Spectrum, Resolvent Set, and Resolvent

Definition 17.1. *Let $\lambda \in \mathbb{C}$, $A \in B(X)$, where X is a Banach space. We say that λ is a **point of spectrum** of the operator A ($\lambda \in \sigma(A)$) if $A - \lambda I$ is not a bijection.*

The study of operator spectra is crucial for numerous applications. In particular, in Quantum Mechanics, to each observable there corresponds a self-adjoint operator, and any measured value of the observable in an experiment must lie within the spectrum of that operator.

The complement to $\sigma(A)$ is resolvent set:

Definition 17.2. $\rho(A) = \mathbb{C}\backslash\sigma(A)$ *is called a **resolvent set**.*

If $\lambda \in \rho(A)$, there exists a bounded inverse $R_\lambda(A) = (A - \lambda I)^{-1} \in B(X)$ (called a **resolvent**).

If A is not bijective, there are two possibilities; it can be not injective or not surjective. Thus, there are different points in the spectrum.

Classification of Points in the Spectrum

Let us consider the following possibilities for $\lambda \in \sigma(A)$:

1) A is not an injection: $\mathrm{Ker}(A-\lambda I)\neq\{0\}$:

$$\exists x\neq 0: \quad (A-\lambda I)x=0 \quad\Leftrightarrow\quad Ax=\lambda x.$$

Such λ and x are called an **eigenvalue** and an **eigenvector** of A respectively. All eigenvalues form a **point spectrum**, which we denote by $\sigma_p(A)$.

2) A is an injection but not a surjection: $\mathrm{Ker}(A-\lambda I)=0$ and $\mathrm{Rn}(A-\lambda I)\neq X$.

   a) $\overline{\mathrm{Rn}(A-\lambda I)}=X$ (the image is dense). Such λ is called a **point of the continuous spectrum**; we denote $\lambda\in\sigma_c(A)$.

   b) $\overline{\mathrm{Rn}(A-\lambda I)}\neq X$ (the image is not dense). Such λ is called a **point of the residual spectrum**; we denote $\lambda\in\sigma_r(A)$.

Thus, the whole complex plane is decomposed into two disjoint sets, $\mathbb{C}=\sigma(A)\sqcup\rho(A)$, and the spectrum is decomposed into three components: $\sigma(A)=\sigma_p(A)\sqcup\sigma_c(A)\sqcup\sigma_r(A)$.

Properties of the Spectrum

Prior to studying the properties of the spectrum, we shall present the theorem on the stability of invertibility.

Theorem 17.2. *Let X be a Banach space, $A\in B(X)$, and $\exists A^{-1}\in B(X)$. Let $B\in B(X)$ such that*

$$\|B\|<\frac{1}{\|A^{-1}\|}.$$

Then $\exists(A+B)^{-1}\in B(X)$.

This means that a small (in some sense) perturbation does not affect the invertibility of an operator.

Proof. Let us recall that if $\|C\|<1$ then $\exists(I\pm C)^{-1}$.

Now, consider $A+B=A(I+A^{-1}B)$; this representation is valid since A is invertible. The inverse operator to a composition is a composition of inverse in the inverse order, i.e., $(A_1A_2)^{-1}=A_2^{-1}A_1^{-1}$, so $\big(A(I+A^{-1}B)\big)^{-1}=(I+A^{-1}B)^{-1}A^{-1}$. There exists an inverse to A, so we have to prove that there exists an inverse to $(I+A^{-1}B)$. Due to

$$\|B\|<\frac{1}{\|A^{-1}\|},$$

$\|A^{-1}B\|\leqslant\|A^{-1}\|\cdot\|B\|<1$, therefore, there exists an inverse to $A+B$ of the form

$$(A+B)^{-1}=(I+A^{-1}B)^{-1}A^{-1}=(I-A^{-1}B+A^{-1}BA^{-1}B-\dots)A^{-1}. \qquad \square$$

Theorem 17.3. *$\sigma(A)$ is a closed set ($\rho(A)$ is open).*

Proof. We will prove the second statement, so $\sigma(A)$, being a complement to $\rho(A)$, would be automatically closed.

Let $\lambda_0 \in \rho(A)$, so $A-\lambda_0 I$ is invertible, and suppose λ belongs to some neighborhood of λ_0:

$$|\lambda-\lambda_0| < \frac{1}{\|(A-\lambda_0 I)^{-1}\|}.$$

We are to prove that $A-\lambda I$ is invertible as well.

First, decompose the operator:

$$A-\lambda I = (A-\lambda_0 I)-(\lambda-\lambda_0)I,$$

where the first one is invertible and the second one is a small perturbation:

$$\|(\lambda-\lambda_0)I\| = |\lambda-\lambda_0| < \frac{1}{\|(A-\lambda_0 I)^{-1}\|}.$$

Then, due to the theorem above, there exists an inverse to $A-\lambda I$, so $\lambda \in \rho(A)$. □

As a side result, let us write the following representation for the inverse to $A-\lambda I = (A-\lambda_0 I)\big(I-(\lambda-\lambda_0)R_{\lambda_0}(A)\big)$:

$$(A-\lambda I)^{-1} = \big(I-(\lambda-\lambda_0)R_{\lambda_0}(A)\big)^{-1}R_{\lambda_0}(A) = \sum_{k=0}^{\infty}(\lambda-\lambda_0)^k R_{\lambda_0}^{k+1}(A);$$

this expression defines an analytic function of λ (an operator-valued geometric series), which converges for

$$|\lambda-\lambda_0| < \frac{1}{\|(A-\lambda_0 I)^{-1}\|}.$$

Theorem 17.4 (Spectrum Localization)**.** *Let X be a Banach space and $A \in B(X)$. Then*

$$\sigma(A) \subset \{\lambda \in \mathbb{C}:\ |\lambda| \leqslant \|A\|\}.$$

Thus, spectrum of A lies within a disk of radius $\|A\|$.

Proof. An equivalent formulation of the theorem is the following: if $|\lambda| > \|A\|$ then $\lambda \in \rho(A)$. We will prove exactly this statement.

Suppose $|\lambda| > \|A\|$. Then

$$A-\lambda I = -\lambda\left(I-\frac{1}{\lambda}A\right); \tag{17.1}$$

denote $C := A/\lambda$, and calculate its norm:

$$\|C\| = \left\|\frac{1}{\lambda}A\right\| = \frac{\|A\|}{|\lambda|} < 1.$$

Thus, (17.1) is invertible, and the inverse has the form

$$(A-\lambda I)^{-1} = -\frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{1}{\lambda^k}A^k, \tag{17.2}$$

which completes the proof. □

Note that representation (17.2) looks similar to the Laurent series. In fact, it is a well-known formula called a **Neumann series** for the resolvent.

Thus, for $\lambda \in \mathbb{C}$ such that

$$|\lambda-\lambda_0| < \frac{1}{\|R_{\lambda_0}(A)\|}, \quad \lambda_0 \in \rho(A),$$

we have the following representation for the resolvent:

$$R_\lambda(A) = \sum_{k=0}^{\infty}(\lambda-\lambda_0)^k R_{\lambda_0}^{k+1}(A),$$

and, for large λ, i.e., when $|\lambda| > \|A\|$, the Neumann series (17.2) becomes valid.

Theorem 17.5. *Let X be a Banach space, $A \in B(X)$. Then $\sigma(A) \neq \varnothing$.*

Note that the assumption that A is bounded is crucial: an unbounded operator may have an empty spectrum. However, there are examples of bounded operators, spectrum of which consists of a single point (for instance, $A=0$ and $A=I$).

Proof by contradiction. Suppose that $\sigma(A) = \varnothing$; then $\rho(A) = \mathbb{C}$. Thus, the resolvent $R_\lambda(A)$ is an analytic function on entire $\mathbb{C}$. One can see that

$$\|R_\lambda(A)\| \to 0 \quad \text{as} \quad |\lambda| \to \infty$$

due to the expansion into Neumann series. Therefore, it is bounded. Then, by Liouville's theorem from the course of Complex Analysis, $R_\lambda(A)$ is constant. Moreover, due to the estimation above, R_λ vanishes at infinity, so $R_\lambda(A) = 0$, which is a contradiction to the invertibility of $A-\lambda I$ (note that the inverse must be invertible as well). □

There is another way to demostrate that the resolvent is continuous and analytic.

Theorem 17.6 (The First Hilbert Identity)**.** *Resolvent of an operator A satisfies the relation*

$$R_\mu(A) - R_\lambda(A) = (\mu-\lambda)R_\lambda(A)R_\mu(A), \tag{17.3}$$

where $\lambda,\ \mu \in \rho(A)$.

Proof. Consider the equality

$$(A-\lambda I)-(A-\mu I)=(\mu-\lambda)I$$

and multiply it by $R_\lambda(A)$ on the left and $R_\mu(A)$ on the right. Then we obtain equality (17.3). □

This identity has a profound corollaries. For instance, it is clear that the resolvent of the same operator taken at different points of the resolvent space commute:

$$R_\lambda(A)R_\mu(A)=R_\mu(A)R_\lambda(A),$$

since, when swapping λ and μ in (17.3), one must change the signs on the left- and on right-hand sides, so the identity preserves; it can be seen clearly from the symmetry of the following expression

$$R_\lambda(A)R_\mu(A)=\frac{R_\mu(A)-R_\lambda(A)}{\mu-\lambda},$$

$\mu\neq\lambda$, with respect to transposition $\lambda\leftrightarrow\mu$.

Let us consider the limit $\mu\to\lambda$ in (17.3); then, since $(\mu-\lambda)\to 0$ on the right-hand side, the left-hand side approaches zero as well:

$$R_\mu(A)-R_\lambda(A),\quad \mu\to\lambda,$$

which means that the resolvent $R_\lambda(A)$ is *continuous* with respect to λ. Considering the limit

$$\lim_{\mu\to\lambda}\frac{R_\mu(A)-R_\lambda(A)}{\mu-\lambda}=R_\lambda^2(A),$$

we get the resolvent has a complex derivative (independent of the direction on the complex plane).

Spectrum of the Adjoint Operator

At times, determining the spectrum of an operator proves to be a difficult task, while the spectrum of its adjoint can be described with relative ease. Hence, it becomes essential to understand the relationship between the spectrum of an operator and that of its adjoint. For applications, the relationship between the spectra of Hilbert adjoint operators is of greater importance; however, we will also discuss the situation involving Banach adjoint operators.

Theorem 17.7. *Let H be a Hilbert space, $A\in B(H)$. Then*

$$\lambda\in\sigma(A)\quad\Leftrightarrow\quad\overline{\lambda}\in\sigma(A^*).$$

In a Banach space, the relation is somewhat different:

Theorem 17.8 (without a proof)**.** *Let X be a Banach space, $A \in B(X)$. Then*

$$\lambda \in \sigma(A) \quad \Leftrightarrow \quad \lambda \in \sigma(A').$$

Proof of Theorem 17.7.

Let us first note that the operations of taking the inverse and taking the adjoint commute:

$$(A^{-1})^* = (A^*)^{-1},$$

if the inverse exists (which is not always true, as opposed to the existence of the adjoint): consider

$$(AA^{-1})^* = (A^{-1})^* A^* = I, \quad (A^{-1}A)^* = A^*(A^{-1})^* = I,$$

then we see that $(A^{-1})^* = (A^*)^{-1}$.

Further, let us formulate the statement of the theorem in the equivalent form:

$$\lambda \in \rho(A) \quad \Leftrightarrow \quad \overline{\lambda} \in \rho(A^*).$$

Suppose that $\lambda \in \rho(A)$; then $\exists (A - \lambda I)^{-1}$. Moreover, there exists

$$\Big((A - \lambda I)^{-1}\Big)^* = (A^* - \overline{\lambda} I)^{-1},$$

which means that $\overline{\lambda} \in \rho(A^*)$. □

Now, recall that we have the classification of points in the spectrum. Let us find out what happens to this classification when taking the adjoint.

Theorem 17.9. *Let H be a Hilbert space, $A \in B(H)$. If $\lambda \in \sigma_r(A)$, then $\overline{\lambda} \in \sigma_p(A^*)$.*

Remark 17.1. *In Banach spaces, $\lambda \in \sigma_r(A) \Rightarrow \lambda \in \sigma_p(A')$.*

Proof. Suppose that $\lambda \in \sigma_r(A)$. Then, by definition of the residual spectrum, the image of the operator is not dense in H:

$$\overline{\mathrm{Rn}(A - \lambda I)} \subsetneq H.$$

This space is nontrivial; thus, there exists a nonzero vector that is orthogonal to it:

$$\exists x \neq 0 : \quad x \perp \mathrm{Rn}(A - \lambda I),$$

which means that

$$\forall y \in H : \quad \big(x, (A - \lambda I) y\big) = 0.$$

Using the definition of the adjoint operator, we rewrite it as

$$\big((A^* - \overline{\lambda} I)x, y\big) = 0 \quad \forall y \in H.$$

Since the vector $(A^* - \overline{\lambda} I)x$ is orthogonal to each $y \in H$, it is zero, therefore,

$$A^* x = \overline{\lambda} x,$$

so $\overline{\lambda} \in \sigma_p(A^*)$. □

Theorem 17.10. *Let H be a Hilbert space, $A \in B(H)$. If $\lambda \in \sigma_p(A)$, then $\overline{\lambda} \in \sigma_p(A^*) \cup \sigma_r(A^*)$.*

Remark 17.2. *In Banach spaces, $\lambda \in \sigma_p(A) \Rightarrow \lambda \in \sigma_p(A') \cup \sigma_r(A')$.*

Proof. First, note that due to Theorem 17.7, if $\lambda \in \sigma_p(A)$ then $\overline{\lambda} \in \sigma(A^*)$. Hence, it is sufficient to prove that $\overline{\lambda}$ does not belong to the continuous spectrum of A^*. By definition, if $\lambda \in \sigma_p(A)$ then $\exists x \neq 0$: $Ax = \lambda x$, therefore,

$$\forall y \in H : \quad \big((A - \lambda I)x, y\big) = 0.$$

Then, by the definition of the adjoint operator,

$$\forall y \in H : \quad \big(x, (A^* - \overline{\lambda} I)y\big) = 0,$$

which means that $\exists x \neq 0$: $x \perp \mathrm{Rn}(A^* - \overline{\lambda} I)$, therefore, $x \perp \overline{\mathrm{Rn}(A^* - \overline{\lambda} I)}$, so the image of $A^* - \overline{\lambda} I$ is not dense in H; that is, $\overline{\lambda} \notin \sigma_c(A^*)$. □

Example 17.1. *In ℓ_2, consider the left- and right-shift operators:*

$$A_r x = (0, x_1, x_2, \dots), \qquad A_\ell x = (x_2, x_3, \dots).$$

What are the spectra of A_r, A_ℓ? These operators are adjoint to each other; it is more convenient to study their spectra simultaneously.

First, let us try to find the point spectrum of A_r:

$$A_r x = \lambda x \quad \Leftrightarrow \quad \begin{cases} 0 = \lambda x_1, \\ x_1 = \lambda x_2, \\ \quad \dots \\ x_n = \lambda x_{n+1}, \\ \quad \dots. \end{cases}$$

In the first row, we have the product of two numbers that is equal to zero. This means that either λ or x_1 is equal to 0.

1) *Suppose* $\lambda = 0$. *Then, the entire column of right-hand sides is zero, therefore, each coordinate is equal to zero:* $x_k = 0$ $\forall k = 1, 2, \dots$. *Therefore,* x *is not an eigenvector, since an eigenvector must be nonzero.*

2) *Suppose* $x_1 = 0$, $\lambda \neq 0$. *Then, solving each equation one by one, we obtain* $x_2 = 0$, $x_3 = 0$, $\dots$, *so* x *is not an eigenvector again; thus,* $\sigma_p(A_r) = \varnothing$.

Now, consider the eigenequation for A_ℓ*:*

$$A_\ell x = \lambda x \quad \Leftrightarrow \quad \begin{cases} x_2 = \lambda x_1, \\ x_3 = \lambda x_2, \\ \dots \\ x_{n+1} = \lambda x_n, \\ \dots . \end{cases}$$

Note that, since the operator is linear, one can seek for solutions (eigenvectors) up to a constant factor. As above, condidering $x_1 = 0$, *we obtain that* $x_2 = x_3 = \dots = 0$. *However, e.g., for* $\lambda = 0$, *the eigenequation for* A_ℓ *has a solution:*

$$A_\ell e_1 = 0.$$

Let us proceed as follows: setting $x_1 = 1$, *we obtain* $x_2 = \lambda$, $x_3 = \lambda^2$, $\dots$; *since* x *must belong to* ℓ_2, *we must reqire that*

$$\sum_{k=1}^{\infty} |\lambda|^{2(k-1)} < \infty.$$

Thus, $\{|\lambda| < 1\} \subset \sigma_p(A_\ell)$.

What do we know about the norms of these operators? Since $\|A\| = \|A^*\|$, *the norms of* A_ℓ *and* A_r *coincide. The norm of* A_r *is equal to* 1, *therefore, the same is true for* A_ℓ*:*

$$\|A_r\| = \|A_\ell\| = 1.$$

The spectrum belongs to the disk of radius equal to the norm of the operator (which is 1 *in our case). Since the spectrum is a closed set, we obtain*

$$\sigma(A_r) = \sigma(A_\ell) = \{|\lambda| \leqslant 1\}.$$

Further, due to Theorem 17.9, the residual spectrum of A_ℓ *is empty:* $\sigma_r(A_\ell) = \varnothing$. *Using Theorem 17.10 and the facts that* $\sigma_p(A_\ell) = \{|\lambda| < 1\}$, $\sigma_p(A_r) = \varnothing$, *we establish that* $\sigma_r(A_r) = \{|\lambda| < 1\}$.

The spectrum is closed; therefore, the only option for the boundary of the unit disk is to belong to the continuous spectrum:

$$\sigma_c(A_r) = \sigma_c(A_\ell) = \{|\lambda| = 1\}.$$

The results can be summarized in a table:

	A_r	A_ℓ
σ_p	$\varnothing$	$\lvert\lambda\rvert < 1$
σ_c	$\lvert\lambda\rvert = 1$	$\lvert\lambda\rvert = 1$
σ_r	$\lvert\lambda\rvert < 1$	$\varnothing$

Spectrum of a Normal Operator

Recall that a normal operator is an operator that commutes with its adjoint; A_ℓ and A_r above serve as examples of nonnormal ones.

Let us formulate the following theorem regarding the structure of spectrum of a normal operator:

Theorem 17.11. *Let A be a normal operator in a Hilbert space H. Then $\sigma_r(A) = \varnothing$.*

Proof by contradiction. Suppose that $\lambda \in \sigma_r(A)$. Then $\overline{\lambda} \in \sigma_p(A^*)$, therefore,

$$\exists x \neq 0: \quad A^*x = \overline{\lambda}x.$$

Recall that for a normal A, $A - \lambda I$ is also normal; further,

$$\|Ax\| = \|A^*x\|.$$

Let us take the vector apply this operatot to x:

$$\|(A - \lambda I)x\| = \|(A^* - \overline{\lambda} I)x\|,$$

where the right-hand side is zero, since x is an eigenvector of A^* corresponding to $\overline{\lambda}$. Thus,

$$\|(A - \lambda I)x\| = 0,$$

therefore, $\lambda \in \sigma_p(A)$, which is a contradiction to our assumption $\lambda \in \sigma_r(A)$ (note that the discrete and residual spectrum do not intersect). □

Recall that self-adjoint, unitary, and multiplication operators are normal. Therefore, they all have empty residual spectrum.

Spectrum of a Self-Adjoint Operator

We already know that for $A = A^*$, the residual specrum is empty: $\sigma_r(A) = \varnothing$.

In Linear Algebra, all symmetric operators have purely real (discrete) spectrum. In the infinite-dimensional setting, for self-adjoint operators, the spectrum is also real, however, it may be a disjoint union of the point and continuous spectra.

Lemma 17.1. *Let X be a Banach space, Y be a normed space, $A \in B(X,Y)$, and*

$$\exists c > 0 \quad \forall x \in X : \quad \|Ax\| \geqslant c\|x\|.$$

Then $\mathrm{Rn}A$ *is closed.*

Remark 17.3. *Why is it important to study the spectrum? Assume that for some λ, we have proved*

$$\|(A-\lambda I)x\| \geqslant c\|x\|. \tag{17.4}$$

Therefore, λ cannot belong to the continuous spectrum, since due to the lemma the image of $A-\lambda I$ is closed (while, for λ to belong to the continuous spectrum, the image and its closure must be different sets).

Note also that bound (17.4) *implies that A is injective.*

Proof. Suppose that y is a limit point of $\mathrm{Rn}A$; then

$$\exists y_n \in \mathrm{Rn}A, \quad y_n \to y.$$

By definition, $\exists x_n$: $Ax_n = y_n$. Let us rewrite inequality (17.4) in the following way:

$$\|x_n - x_m\| \leqslant \frac{1}{c}\|y_n - y_m\|.$$

$y_n \to y$, so it is a Cauchy sequence, therefore, x_n is also Cauchy. Since X is Banach, the limit point belongs to X: $x_n \to x \in X$. Since the operator is continuous (which is equivalent to that it is bounded), $Ax_n \to Ax$. Thus, $Ax = y$, so $y \in \mathrm{Rn}A$, which means that the image is closed. □

Theorem 17.12. *Let $A = A^* \in B(H)$, where H is a Hilbert space. Then*

$$\sigma(A) \subset \mathbb{R}.$$

Proof.

1) Suppose that $\lambda \in \sigma_p(A)$. Then

$$\exists x \neq 0: \quad Ax = \lambda x.$$

Let us take the inner product of this equality with the same vector:

$$(Ax, x) = (\lambda x, x) = \lambda \|x\|^2.$$

Rewriting the left-hand side, we obtain

$$(x, Ax) = (x, \lambda x) = \overline{\lambda} \|x\|^2,$$

therefore, $\lambda \|x\|^2 = \overline{\lambda} \|x\|^2$, $\|x\| \neq 0$, so $\lambda = \overline{\lambda}$.

2) Suppose that $\lambda \in \sigma_c(A)$, $\lambda = \alpha + i\beta$, $\beta \neq 0$, and consider

$$\|(A-\lambda I)\|^2 = \big((A-\alpha I - i\beta I)x, (A-\alpha I - i\beta I)x\big) = \|(A-\alpha I)x\|^2 + i\beta\big((A-\alpha I)x, x\big) - i\beta\big(x, (A-\alpha I)x\big) + \|$$

Since $(A-\alpha I)^* = A^* - \alpha I = A - \alpha I$, the second and the third term cancel each other. Thus, we arrive at the bound

$$\|(A-\lambda I)\| \geqslant |\beta| \|x\|.$$

Due to the lemma above, the image of $A - \lambda I$ is closed, therefore, $\lambda \notin \sigma_c(A)$, which is a contradiction to our assumption. Therefore, $\sigma_c(A) \subset \mathbb{R}$.

3) For $A = A^*$, $\sigma_r(A) = \varnothing$. □

Spectral Radius

Furthermore, we can say that for $A = A^*$, the spectrum belongs to the interval: $\sigma(A) \subset [-\|A\|, \|A\|]$. However, this estimation is not quite sharp: e.g., consider $A = I$; for this operator, we obtain $\sigma(I) \subset [-1, 1]$, while in fact $\sigma(I) = \{1\}$. To resolve this issue, we will use a new notion.

Definition 17.3. *Let X be a Banach space and $A \in B(X)$. The* ***spectral radius*** *of A is defined as*

$$r(A) = \max_{\lambda \in \sigma(A)} |\lambda|.$$

The theorem on spectrum localization implies the following inequality:

$$r(A) \leqslant \|A\|.$$

One can see that this inequality is not sharp. Consider, for instance, a Jordan matrix of the form

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

For this operator, we have $r(A) = 0$, since $\sigma(A) = \{0\}$, while $\|A\| > 0$, since the operator is nonzero.

However, for normal (and, therefore, for self-adjoint) operators the spectral radius coincides with the norm:

Theorem 17.13 (Gelfand's Spectral Radius Formula, without proof)**.**

$$r(A) = \lim_{n\to\infty} \|A^n\|^{1/n}.$$

Remark 17.4. *Applying the Cauchy–Hadamard theorem, which determines the radius of convergence for power series, to the resolvent in the form of Neumann series*

$$R_\lambda(A) = -\frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{A^k}{\lambda^k}, \quad |\lambda| \geqslant \|A\|,$$

like for numerical series $\sum_k a_k z^k$, *for which*

$$\frac{1}{R} = \overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|},$$

then we obtain the statement of the theorem. While for numerical series, the upper limit is considered, for the operator-valued series the limit always exists due to the submultiplicative inequality $\|A^{n+k}\| \leqslant \|A^n\| \cdot \|A^k\|$.

Lecture 18. Exercises on Spectra of Operators

Discussion of Self-Study Problems form the Previous Lecture

We begin with considering some of the self-study problems from Lecture 16.

5) Consider

$$(Af)(x) = f(x) - \int_0^x f(t)\,dt = (I-C)f$$

a) in $C[0,1]$.

b) in $L_2[0,1]$.

Find the inverse operator. The answer must not involve infinite sums.

First, if $\|C\|$, we can write out the inverse operator in the form

$$(I-C)^{-1} = \sum_{k=0}^{\infty} C^k.$$

For $L_2[0,1]$, $\|C\| \leqslant \|K\|_{L_2[a,b]^2}$, where, in out problem,

$$K(x,t) = \begin{cases} 1, & t < x, \\ 0, & t > x, \end{cases}$$

therefore,

$$\|K(x,t)\| = 1/\sqrt{2}.$$

However, in $C[0,1]$ (as has been previously proved), the norm of C is not small:

$$\|Cf\| = \max_{x\in[0,1]}\left|\int_0^x f(t)\,dt\right| \leqslant \max_{x\in[0,1]}\left|\int_0^x |f(t)|\,dt\right|,$$

where $|f(t)| \leqslant \|f\| = 1$, thus, $\|Cf\| \leqslant 1$; for $f(t) \equiv 1$, we have

$$\|Cf\| = \max_{x\in[0,1]}\left|\int_0^x f(t)\,dt\right| = \max_{x\in[0,1]} x = 1,$$

so the bound is sharp, and $\|C\| = 1$.

Further, even though the bound $\|C\| < 1$ does not hold, we can employ the expansion of $(I-C)^{-1}$ into series, since it is fine if the bound holds for some power of C, meaning that the series converges if

$$\exists n_0 \quad \forall n \geqslant n_0: \quad \|C^n\| < 1.$$

Let us estimate the norms of powers of C. Beginning with the second power, we get

$$(C^2 f)(x) = \int_0^x \Big(\int_0^t f(s)\,ds \Big)\,dt;$$

here, we integrate with respect to s and t such that $0 < s < t < x$. Let us change the order of integration, so that we would integrate with respect to t first:

$$\int_0^x \Big(\int_0^t f(s)\,ds \Big)\,dt = \int_0^t f(s) \int_s^x dt\,ds = \int_0^x (x-s) f(s)\,ds.$$

Furthermore,

$$(C^n f)(x) = \int_0^x \frac{(x-t)^{n-1}}{(n-1)!} f(t)\,dt,$$

which can be easily proved by virtue of mathematical induction. It is evident that the norms decay rapidly as a result of the division by the factorial: $\|C^n\| \to 0$ as $n \to \infty$.

Now, let us find the inverse to A:

$$(A^{-1} f)(x) = \big((I-C)^{-1} f(x)\big)(x) = \sum_{k=0}^{\infty} C^k f(x) = f(x) + \sum_{k=0}^{\infty} \int_0^x \frac{(x-t)^k}{k!} f(t)\,dt,$$

where we will swap the order of summation and integration (it is totally legal since the sum converges uniformly):

$$f(x) + \sum_{k=0}^{\infty} \int_0^x \frac{(x-t)^k}{k!} f(t)\,dt = f(x) + \int_0^x e^{x-t} f(t)\,dt.$$

Another approach to solve the problem is the following. Constructing the inverse is equivalent to solving the equation

$$f(x) - \int_0^x f(t)\,dt = g(x)$$

for $f(x)$ for a given $g(x)$. Suppose the functions in question have derivatives of higher order (note that $C^1[0,1]$ is dense in both $C[0,1]$ and $L_2[0,1]$). Then we can differentiate the equation and solve the ordinary differential equation obtained.

4) Consider the differential operator $Af = f'$ in $C[0,1]$ with domain $\mathcal{D}(A) = C^1[0,1]$. Prove that there exists a right inverse, but it is not unique.

Let us note that this operator is not invertible since it has a nontrivial kernel:

$$\operatorname{Ker} A = \langle 1 \rangle.$$

However, there is a right inverse:

$$A_r^{-1}f = \int_0^x f(t)\,dt + C, \quad AA_r^{-1} = I.$$

For $C \neq 0$, the operator A_r^{-1} is nonlinear, and, of course, it is not unique.

1) In $L_2[0,1]$, consider the Hardy operator

$$(Af)(x) = \frac{1}{x}\int_0^x f(t)\,dt.$$

   a) Prove that A is bounded.

   b) Prove that A is not compact.

It is quite simple to obtain the bound $\|A\| \leqslant 2$. In further, we will see that the point spectrum of the Hardy operator consists of points $\{|z-1|<1\}$, so $\|A\| = 2$.

Let us begin with the estimation:

$$\|Af\|^2 = \int_0^1 \Big|\frac{1}{x}\int_0^x f(t)\,dt\Big|^2 dx \leqslant \int_0^1 \Big(\int_0^x |f(t)|\,dt\Big)^2 \Big(-d\frac{1}{x}\Big),$$

which can be integrated by parts:

$$\int_0^1 \Big(\int_0^x |f(t)|\,dt\Big)^2 \Big(-d\frac{1}{x}\Big) = -\Big(\int_0^x |f(t)|\,dt\Big)\frac{1}{x}\Big|_0^1 + \int_0^1 2\Big(\int_0^x |f(t)|\,dt\Big)\cdot|f(x)|\frac{1}{x}\,dx.$$

At point $x=1$, the first term is negative, so by excluding it, we obtain an upper bound; at point $x=0$, this term must be carefully calculated, since there is a possible singularity due to the x-inverse factor. Let us use the Cauchy–Bunyakovsky–Schwarz inequality:

$$\frac{1}{x}\Big(\int_0^x |f(t)|\,dt\Big)^2 \leqslant \frac{1}{x}\Big(\int_0^x 1\,dt \times \int_0^x |f(t)|^2\,dt\Big) = \frac{1}{x}x\int_0^x |f(t)|^2 \to 0 \quad \text{as} \quad x \to 0,$$

so, in fact, there is no singularity. For $\|Af\|^2$, we obtain

$$\|Af\|^2 \leqslant 2\int_0^1 \Big(\int_0^x |f(t)|\,dt\Big)\cdot|f(x)|\frac{1}{x}\,dx.$$

Let us use the Cauchy–Bunyakovsky–Schwarz inequality again (with $|f(x)|$ as one of the integrand functions):

$$2\int_0^1 \Big(\int_0^x |f(t)|\,dt\Big)\cdot|f(x)|\frac{1}{x}\,dx \leqslant 2\Big(\int_0^1 \frac{1}{x}\Big(\int_0^x |f(t)|\,dt\Big)^2\Big)^{1/2} \times \Big(\int_0^1 |f(x)|^2\,dx\Big)^{1/2},$$

where the second factor is $\|f\|$. Let us denote

$$M := \Big(\int_0^1 \frac{1}{x}\Big(\int_0^x |f(t)|\,dt\Big)^2\Big)^{1/2}.$$

Recalling the beginning of our estimation, we obtain the bounds

$$\|Af\|^2 \leqslant M^2 \leqslant 2M\|f\|,$$

therefore, $M \leqslant 2\|f\|$, and

$$\|Af\|^2 \leqslant 4\|f\| \quad \Rightarrow \quad \|A\| \leqslant 2.$$

Let us also try to solve the eigenequation for A in the form of a power function x^α, $\alpha \in \mathbb{R}$:

$$Ax^\alpha = \lambda x^\alpha, \quad \frac{1}{\lambda}\int_0^x t^\alpha\,dt = \frac{x^\alpha}{\alpha+1},$$

so $\lambda = 1/(\alpha+1))$. Note that $x^\alpha \in L_2[0,1]$ if

$$\int_0^1 x^{2\alpha}\,dx < \infty,$$

whence $2\alpha > -1$, or, equivalently, $\alpha > -1/2$. Taking

$$\alpha_n = -\frac{1}{2} + \frac{1}{n},$$

we see that

$$\lambda_n = \frac{1}{-\frac{1}{2}+\frac{1}{n}+1} \to 2 \quad \text{as} \quad n \to \infty,$$

so the spectral radius is at least 2, and, therefore, the norm is at least 2 as well.

Let us prove that this operator is not compact. We will demonstrate it using the property of compact operators: a compact operator maps a weakly converging sequence to a converging one.

First, let us point out that the Hardy operator seems to be bad near $x = 0$. We will construct a sequence of functions that concentrate at $x = 0$:

$$f_n(x) = \sqrt{n}\chi_{\left[0,\frac{1}{n}\right]}(x), \quad \|f_n\| = 1.$$

We claim that $f_n \rightharpoonup 0$. Why is that? We must show that

$$\forall F \in \big(L_2[0,1]\big)^*: \ F(f_n) \to 0.$$

By Riesz's theorem,

$$F(f_n) = (f_n, g) \equiv \int_0^{1/n} \sqrt{n}\,\overline{g(x)}\,dx,$$

to which we apply the Cauchy–Bunyakovsky–Schwarz inequality:

$$\int_0^{1/n} \sqrt{n}\,\overline{g(x)}\,dx \leqslant \Big(\int_0^{1/n} n\,dx\Big)^{1/2}\Big(\int_0^{1/n} |g(x)|^2\,dx\Big)^{1/2} = 1\cdot\Big(\int_0^{1/n} |g(x)|^2\,dx\Big)^{1/2} \to 0$$

as $n\to\infty$, since $g\in L_2[0,1]$ and the integration interval shrinks to zero (to a set of measure zero). Further,

$$\|Af_n\|^2 = \int_0^1 \Big(\frac{1}{x}\int_0^x \sqrt{n}\chi_{\left[0,\frac{1}{n}\right]}(t)\,dt\Big)^2 dx = \int_0^{1/n} \Big(\frac{1}{x}\int_0^x \sqrt{n}\,dt\Big)^2 dx + \int_{1/n}^1 \ldots\,dx,$$

where the second term is nonnegative, so

$$\|Af_n\|^2 \geqslant \int_0^{1/n} \Big(\frac{1}{x}\int_0^x \sqrt{n}\,dt\Big)^2 dx = 1 \not\to 0,$$

therefore, A is not compact.

Later, we will show that the spectrum of a compact operator, except for $\lambda=0$, is purely discrete and consists of isolated points. As can be seen, the spectrum of the Hardy operator is not of this form.

Exercises on Spectra and Spectral Radii. Spectrum of a Self-Adjoint Operator

1) Prove that for $A=A^*\in B(H)$, where H is a Hilbert space,

$$r(A) = \|A\|.$$

   a) First, we will show that $\|A^*A\| = \|A\|^2$ for any $A\in B(H)$. In one direction, the estimation is obvious:

$$\|A^*A\| \leqslant \|A^*\|\cdot\|A\| = \|A\|^2,$$

   since $\|A^*\| = \|A\|$. We know that A^*A is self-adjoint. In Lecture 13, we proved that the norm of a self-adjoint operator can be computed as the supremum of the associated quadratic form:

$$\|A^*A\| = \sup_{\|x\|=1} |(A^*Ax,x)|,$$

   so

$$\sup_{\|x\|=1} |(A^*Ax,x)| = \sup_{\|x\|=1} |(Ax,Ax)| = \sup_{\|x\|=1} \|Ax\|^2 = \|A\|^2.$$

b) If $A = A^*$, then $\|A^2\| = \|A\|^2$, thus, $\|A^{2^n}\| = \|A\|^{2^n}$; let us prove it by mathematical induction:

$$\|A^{2^{n+1}}\| = \|(A^{2^n})^2\| = \|A^{2^n}\|^2,$$

which is equal to $\|A\|^{2^{n+1}}$ by the induction hypothesis.

Next, using this in the formula for the spectral radius, we obtain

$$r(A) = \lim_{n\to\infty} \|A^n\|^{1/n} = \lim_{k\to\infty} \|A^{2^k}\|^{1/2^k} = \|A\|.$$

In Lecture 17, that we obtained the following: for $A = A^*$,

$$\sigma(A) \subset [-\|A\|, \|A\|].$$

This bound is not quite good, since, for instance, the spectrum of the identity operator is $\sigma(I) = \{1\}$, while the inclusion above gives us $\sigma(I) \subset [-1, 1]$. Now, we will try to make the bound sharper.

2) Let $A = A^*$. Define

$$m = \inf_{\|x\|=1} (Ax, x), \quad M = \sup_{\|x\|=1} (Ax, x).$$

Then, we claim that $\sigma(A) \subset [m, M]$, moreover, both endpoints belong to the spectrum, i.e., $m, M \in \sigma(A)$, and $\max(|m|, |M|) = \|A\|$.

For example, for $A = I$ we have $m = M = 1$, and this is precisely the spectrum of I.

a) For any $x \in H$, consider

$$m\|x\|^2 \leqslant (Ax, x) \leqslant M\|x\|^2.$$

For $x = 0$, it is the equality; for $x \neq 0$, we will divide it by $\|x\|^2$:

$$m \leqslant \left(A\frac{x}{\|x\|}, \frac{x}{\|x\|}\right) \leqslant M,$$

which follows from the definition of m and M.

Further, we should discuss the localization of spectrum.

b) Let $\lambda \in \sigma_p(A)$: $\exists x$, $x \neq 0$, $Ax = \lambda x$. Then,

$$(Ax, x) = \lambda\|x\|^2, \quad m\|x\|^2 \leqslant \lambda\|x\|^2 \leqslant M\|x\|^2,$$

therefore, $m \leqslant \lambda \leqslant M$.

c) The residual spectrum is empty.

d) We must show that if $\lambda > M$ (and, similarly, $\lambda < m$), then $\lambda \notin \sigma_c(A)$.

Let $\lambda = M + \delta$, $\delta > 0$. Then, consider

$$\|(A-\lambda I)x\|^2 = \Big((A-MI-\delta I)x, (A-MI-\delta I)x\Big) =$$
$$= \|(A-MI)x\|^2 - \delta\Big((A-MI)x, x\Big) - \delta\Big(x, (A-MI)x\Big) + \delta^2\|x\|^2,$$

where, due to the self-adjointness of A,

$$-\delta\Big((A-MI)x, x\Big) - \delta\Big(x, (A-MI)x\Big) = -2\delta\Big((A-MI)x, x\Big) =$$
$$= -2\delta(Ax, x) + 2\delta M\|x\|^2 \geqslant 0,$$

so, excluding these terms from the equality above, we obtain the bound

$$\|(A-\lambda I)x\|^2 \geqslant \delta\|x\|^2.$$

Thus, due to the theorem from Lecture 17, the image is closed, which implies that $\lambda \notin \sigma_c(A)$. The proof for $\lambda < m$ is similar.

e) Now, let us show that the endpoints of the interval belong to the spectrum: m, $M \in \sigma(A)$. Let us consider

$$\widetilde{A} = A - mI.$$

This operator is self-adjoint as well, and $\widetilde{m} = 0$, $\widetilde{M} = M - m$:

$$\sigma(\widetilde{}) \subset [0, M-m];$$

moreover, $\|\widetilde{A}\| = M - m$, and $\|\widetilde{A}\| = r(\widetilde{A}) \Rightarrow M - m = r(\widetilde{A})$. Therefore, there exists $\lambda \in \sigma(\widetilde{A})$:

$$\lambda = M - m.$$

Thus, shifting it back, we obtain

$$M \in \sigma(A).$$

For $\widetilde{A} = A - MI$, we obtain

$$\widetilde{m} = m - M \leqslant 0,$$

and the further proof for $m \in \sigma(A)$ is similar.

Spectra of Similar Operators

The problem of finding the spectrum of an operator is often quite challenging. Next, we will consider an approach that simplifies it in certain cases.

Definition 18.1. *Let X, Y be Banach spaces, and $A \in B(X)$. Let there exist a bijective operator S, $S \in B(X,Y)$, and an operator $B \in B(Y)$ such that the diagram is commutative*

$$\begin{array}{ccc} X & \xrightarrow{A} & X \\ {\scriptstyle S}\big\downarrow & & \big\downarrow{\scriptstyle S} \\ Y & \xrightarrow{B} & Y, \end{array}$$

i.e., $SA = BS$. Then we say that the operator A is ***similar to the operator*** *B and denote*

$$A \sim B.$$

Note that since S is bijective, due to the Banach bounded inverse theorem, there exists S^{-1} so that

$$SAS^{-1} = B.$$

In finite-dimensional spaces, we can fix a basis, and the operator takes the form of a matrix in that basis. Under a change of basis with a transition matrix S, the matrix of the operator transforms according to the same rule. It is a known result in Linear Algebra that the characteristic polynomial of a matrix is unchanged under a change of basis. Therefore, the eigenvalues of the operator, which are the roots of the characteristic polynomial, also remain invariant. The same is true in Banach spaces: the spectra of similar operator coincide.

Theorem 18.1. *Let $A \in B(X)$, $B \in B(Y)$, where X, Y are some Banach spaces. Let $A \sim B$. Then*

$$\sigma(A) = \sigma(B),$$

moreover, the classification of points in spectra coincide.

Proof.

1) Let $\lambda \in \rho(A) \Leftrightarrow \exists (A - \lambda I)^{-1} \in B(X)$,

$$(A - \lambda I)^{-1} = (S^{-1}BS - \lambda I)^{-1} = \left(S^{-1}(B - \lambda I)S\right)^{-1} = S^{-1}(B - \lambda I)^{-1}S,$$

therefore, $\lambda \in \rho(B)$, so the resolvent sets of A and B coincide, which means that the spectra coincide as well.

2) Let $\lambda \in \sigma_p(A)$. Then $\exists x \neq 0$:

$$Ax = \lambda x.$$

Since $A = S^{-1}BS$, we have

$$S^{-1}BS = \lambda x \quad \Leftrightarrow \quad BS = \lambda Sx,$$

and $Sx \neq 0$, since S is injective; therefore, Sx is an eigenvector of B corresponding to an eigenvalue λ.

3) The continuous and residual spectra of A are related to the properties of image of A (more precisely, to whether the imeage is dense in the entire space). Consider

$$(A-\lambda I) = S^{-1}(B-\lambda I)S.$$

Thus, if the image of $(A-\lambda I)$ is dense in X, then the image of $(B-\lambda I)$ is dense in Y, and vice versa. Therefore,

$$\sigma_c(A) = \sigma_c(B) \quad \text{and} \quad \sigma_r(A) = \sigma_r(B). \qquad \square$$

Example 18.1. *Consider A_ℓ, A_r in two-sided ℓ_2: $\ell_2(\mathbb{Z})$, where*

$$\ell_2(\mathbb{Z}) \ni x = (\dots, x_{-1}, x_0, x_1, \dots)$$

with the condition

$$\sum_{k=-\infty}^{\infty} |x_k|^2 < \infty.$$

(For instance, the discrete Schrödinger operator is usually considered in this space).

In $\ell_2(\mathbb{Z})$, for $x = \big(\dots, x_{-1}, (x_0), x_1, \dots\big)$,

$$A_r x = \big(\dots, x_{-2}, (x_{-1}), x_0, \dots\big), \quad A_r e_n = e_{n+1},$$

and

$$A_\ell x = \big(\dots, x_0, (x_1), x_2, \dots\big), \quad A_r e_n = e_{n-1}.$$

It is known that all separable Hilbert spaces are isometrically isomorphic. Thus, there are a bijection S and an operator B_r such that

$$\begin{array}{ccc} \ell_2(\mathbb{Z}) & \xrightarrow{A_r} & \ell_2(\mathbb{Z}) \\ {\scriptstyle S}\big\downarrow & & \big\downarrow{\scriptstyle S} \\ L_2[0,2\pi] & \xrightarrow{B_r} & L_2[0,2\pi], \end{array}$$

where B_r acts on the basis elements in the same way as A_r does.

Operator S must map a basis into a basis; in $\ell_2(\mathbb{Z})$, a basis can be chosen in the form

$$e_n = (\dots, 0, \overset{n}{1}, 0, \dots);$$

in $L_2[0,\pi]$, let us fix a complex exponential basis:

$$E_n(t) = \frac{1}{\sqrt{2\pi}} e^{int}, \quad n \in \mathbb{Z}.$$

Since $e_n \mapsto e_{n+1}$ under A_r, we have for B_r

$$B_r E_n = E_{n+1},$$

so B_r is a multiplication operator:

$$B_r f(t) = e^{it} f(t), \quad f \in L_2[0,2\pi].$$

Similarly, the operator A_ℓ is similar to the multiplication operator B_ℓ such that

$$B_\ell f(t) = e^{-it} f(t), \quad f \in L_2[0,2\pi].$$

It is clear that A_ℓ and A_r are adjoint and inverse to each other in $\ell_2(\mathbb{Z})$, and the same holds for B_ℓ, B_r. Thus, these operators are unitary. The spectrum of a unitary operator lies on the unit circle.

In further lectures, we will study the multiplication operators in more detail. For now, we will formulate the following theorem:

Theorem 18.2. *Let $\varphi \in L_\infty[a,b]$. Then, for $A_\varphi : L_2[a,b] \to L_2[a,b]$,*

$$A_\varphi f = \varphi(x) f(x),$$

the equality

$$\sigma(A_\varphi) = \operatorname{ess} E(\varphi)$$

holds, where $\operatorname{ess} E(\varphi)$ is the set of essential values of φ:

$$\operatorname{ess} E(\varphi) = \Big\{ \lambda : \ \forall \varepsilon > 0 \ \mu\big(\{x : \ |\varphi(x) - \lambda| < \varepsilon\}\big) > 0 \Big\}.$$

Note that, e.g., for $\varphi \in C[\alpha,\beta]$, the essential range is simply the range. Next, consider

$$\operatorname{sgn} t = \begin{cases} -1, & t < 0, \\ 0, & t = 0, \\ 1, & t > 0. \end{cases}$$

The values -1 and 1 of $\operatorname{sgn} t$ are essential, and the value 0 is not essential since the function takes this value on the set of measure zero.

Note also that the multiplication operators in L_2 are normal. Therefore, the residual spectra of B_ℓ and B_r are empty.

Are there eigenvalues of B_ℓ and B_r? For $\lambda \in \sigma_p(A_\varphi)$, there must exist a function $f \in L_2[a,b]$, $f \neq 0$, such that

$$A_\varphi f = \lambda f,$$

which is the same as

$$(\varphi(x)-\lambda)f(x)=0.$$

For this product to vanish on the entire $[a,b]$, either the function f must be vanishing on $[a,b]$ (in the sense of L_2), or $\varphi(x)=\lambda$ on a set of positive measure. For instance, the function

$$\varphi\big|_{(\alpha,\beta)}\equiv C$$

satisfies the condition. However, $e^{\pm it}$ is not constant on any set. Thus, the spectra of B_r and B_ℓ are purely continuous.

Self-Study Exercises

1) Prove that $\sigma(AB)\backslash\{0\}=\sigma(BA)\backslash\{0\}$.

2) If $AB-BA=I$, then at least one of the operators A, B is unbounded.

   Hint: $AB=BA+I$ implies

   $$AB-\lambda I=BA-(\lambda-1)I,$$

   so the “shifted” set must coincide with the set itself. Therefore, it is either an unbounded set, or an empty set. However, the spectrum of a bounded operator cannot be empty.

3) Let $U^*=U^{-1}$. Prove that

   $$\sigma(U)\subset\{z\in\mathbb{C}:\ |z|=1\}.$$

4) Let $\alpha=(\alpha_1,\alpha_2,\dots)\in\ell_\infty$. In ℓ_2, consider

   $$A_\alpha x=(\alpha_1x_1,\alpha_2x_2,\dots).$$

   Find $\sigma(A_\alpha)$.

5) Let X be a Banach space and $\Omega\subset\mathbb{C}$ be a nonempty compact set. Prove that

   $$\exists A\in B(X):\quad \sigma(A)=A.$$

   Hint: Use problem 4 to construct an operator A.

6) Let $U = U^* = U^{-1}$. Describe all operators of this form.

   Hint: The entire Hilbert space must be decomposed into two components $H = H_0 \oplus H_0^\perp$ such that

$$U = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix},$$

   where $U\big|_{H_0} = I$ and $U\big|_{H_0^\perp} = -I$.

7) In ℓ_2, for $a,\, b \in \mathbb{R}$, consider

$$Ae_n = be_{n-1} + ae_n + be_{n+1}, \quad n \geqslant 2, \quad Ae_1 = ae_1 + be_2,$$

$$A \sim \begin{pmatrix} a & b & 0 & 0 & \dots \\ b & a & b & 0 & \dots \\ 0 & b & a & b & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

   Find the spectrum of A by constructing a similar operator.

Lecture 19. The Hilbert–Schmidt Theorem

Weyl Sequences

We continue to study the spectrum. In this lecture, we will formulate a number of theorems that help one to find it.

Definition 19.1. *Let X be a Banach space, $A \in B(X)$. We say that for $\lambda \in \mathbb{C}$ there exists a **Weyl sequence** $\{x_n\}$ if*

$$\|x_n\| = 1, \quad (A - \lambda I)x_n \to 0 \quad \text{as} \quad n \to \infty.$$

For instance, suppose that $x \neq 0$ is an eigenvector corresponding to an eigenvalue $\lambda \in \mathbb{C}$ of an operator A; consider $x_n \equiv x$. Then $\{x_n\}$ is a Weyl sequence for λ. Thus, for clarity, it is convenient to think of a Weyl sequence as an "almost eigenvector".

Theorem 19.1. *If for $\lambda \in \mathbb{C}$ there exists a Weyl sequence $\{x_n\}$, then $\lambda \in \sigma(A)$.*

Proof by contradiction. Let $\lambda \in \rho(A)$; then there is an inverse:

$$\exists (A - \lambda I)^{-1} \in B(X).$$

Denote $y_n := (A - \lambda I)x_n$; $y_n \to 0$. Applying the inverse to y_n, we get

$$(A - \lambda I)^{-1} y_n = x_n \not\to 0,$$

which is a contradiction to the continuity of $(A - \lambda I)^{-1}$, since for a continuous (which is the same as *bounded*) operator T, $w_n \to 0 \Rightarrow Tw_n \to 0$. Thus, there is no bounded inverse, so $\lambda \in \sigma(A)$, which completes the proof. □

In the previous lecture, we formulated a theorem on the spectrum of a multiplication operator. Let us return to it:

Theorem 19.2. *Let $A_\varphi : L_2[a,b] \to L_2[a,b]$, $A_\varphi f = \varphi(x) f(x)$, where $\varphi \in L_\infty[a,b]$. Then*

$$\sigma(A_\varphi) = \operatorname{ess} E(\varphi) \equiv \left\{ \lambda \in \mathbb{C} : \ \forall \varepsilon > 0 \ \mu\{x : |\varphi(x) - \lambda| < \varepsilon\} > 0 \right\}.$$

Moreover, if there exists a measurable set Ω, $\mu(\Omega) > 0$, such that $\varphi\big|_\Omega \equiv \lambda$, then $\lambda \in \sigma_p(A_\varphi)$. The remaining essential values form the continuous spectrum $\sigma_c(A_\varphi)$.

Remark 19.1. *1) Note that A_φ is normal, so $\sigma_r(A_\varphi) = \varnothing$.*

2) For application, the most useful case is $\varphi \in C[a,b]$. For such φ, $\operatorname{ess} E(\varphi)$ is just the set of all values that φ takes.

Proof.

1) Let $\lambda \in \operatorname{ess} E(\varphi)$. Then, by the definition of essential range,
$$\forall n \in \mathbb{N} : \mu\left\{x : |\varphi(x) - \lambda < \frac{1}{n}\right\} > 0;$$
let us denote $M_n := \{x : |\varphi(x) - \lambda < 1/n\}$, and define a function f_n:
$$f_n := \frac{\chi_{M_n}(x)}{\sqrt{\mu(M_n)}}.$$
One can see that $\|f_n\|_{L_2} = 1$. Next,
$$\|(A - \lambda I)f\|^2 = \int_{M_n} \frac{|\varphi - \lambda|^2}{\mu(M_n)}\, d\mu < \frac{1}{n^2} \to 0,$$
so f_n is a Weyl sequence, and, therefore, $\lambda \in \sigma(A)$.

2) Now, we are to prove the inverse. Suppose $\lambda \notin \operatorname{ess} E(\varphi)$; we will show that $\lambda \in \rho(A_\varphi)$. Let us denote
$$M_\varepsilon := \{x : \ |\varphi(x) - \lambda| < \varepsilon\};$$
by definition,
$$\lambda \notin \operatorname{ess} E(\varphi) \quad \Leftrightarrow \quad \exists \varepsilon > 0 : \quad \mu(M_\varepsilon) = 0.$$
The problem of construction the resolvent $R_\lambda(A_\varphi) = (A - \lambda I)^{-1}$ is equivalent to solving the equation
$$(A_\varphi - \lambda I)f = g$$
for an arbitrary given $g \in L_2$. The equation can be rewritten as
$$(\varphi(x) - \lambda)f(x) = g(x),$$
so, if we seek for a solution $f(x)$, it is sufficient to divide by the first factor:
$$f(x) = \frac{1}{\varphi(x) - \lambda} g(x).$$
However, if $|\varphi(x) - \lambda|$ is “small”, the resulting function $f(x)$ can be that “large” so it would not belong to L_2. Let us exclude the small values from the result by considering
$$f(x) = \begin{cases} \dfrac{1}{\varphi(x) - \lambda} g(x), & x \notin M_\varepsilon, \\ 0, \quad x \in M_\varepsilon. \end{cases}$$
Note that since M_ε is a set of measure zero, and the space $L_2[a,b]$ is an equivalence class of functions that are equal almost everywhere (that is, they are equal except

for a set of measure zero), $f(x)$ may take any form on M_ε, and all functions that differ on M_ε are indistinguishable in the L_2-sense.

Further, we must verify that the resolvent defined by the rule above is bounded. Consider

$$\|R_\lambda(A_\varphi)g\|^2 = \int\limits_{[a,b]\setminus M_\varepsilon} \frac{|g(x)|^2}{|\varphi(x)-\lambda|^2}\,d\mu;$$

on the integration set, $|\varphi(x)-\lambda| > \varepsilon$, so

$$\|R_\lambda(A_\varphi)g\|^2 < \frac{1}{\varepsilon^2}\int_{[a,b]\setminus M_\varepsilon} |g(x)|^2\,d\mu \leqslant \frac{1}{\varepsilon^2}\|g\|^2,$$

thus, $\|R_\lambda\| < 1/\varepsilon$. Note that ε is a fixed nonzero value. Therefore, R_λ is bounded, and $\lambda \in \rho(A_\varphi)$.

Next, we must prove the statements on the classifications of points in spectrum; it is quite simple.

3) Let $\lambda \in \operatorname{ess} E(\varphi)$. When $\lambda \in \sigma_p(A_\varphi)$? For λ to belong to the discrete spectrum, the following must hold:

$$\exists f \in L_2[a,b],\ f \neq 0\ (\exists \Omega:\ \mu(\Omega) > 0,\ f\Big|_\Omega(x) \neq 0\ \forall x \in \Omega), \qquad A_\varphi f = \lambda f.$$

It means that

$$(\varphi(x)-\lambda)f(x) = 0 \quad \text{in} \quad L_2[a,b];$$

since $f(x) \neq 0$ on Ω, the first factor must vanish on this set:

$$\varphi(x) - \lambda \equiv 0 \quad \text{on} \quad \Omega,$$

where $\mu(\Omega) > 0$.

Since the residual spectrum is empty, all the other points of $\operatorname{ess} E(\varphi)$ belong to $\sigma_c(A_\varphi)$. □

Note that the essential range of a continuous function on an interval coincides with range. However, this is not true for continuous functions on $\mathbb{R}$; consider, e.g.,

$$\varphi(x) = \frac{1}{x^2+1}, \quad x \in \mathbb{R}.$$

This function takes the values $(0,1]$, see Fig. 19.1, while the essential range is $[0,1]$.

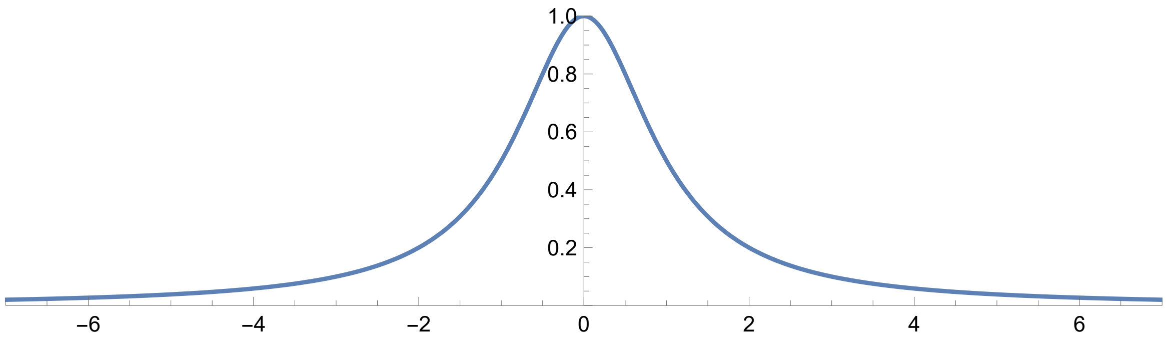

Рис. 19.1. Graph of $\varphi(x)$.

Further, let us consider a multiplication operator in $C[a,b]$. The result is similar to one in $L_2[a,b]$ with minor modifications.

Theorem 19.3. *Let $\varphi \in C[a,b]$, $A_\varphi : C[a,b] \to C[a,b]$, $A_\varphi f = \varphi(x) f(x)$. Then*

$$\sigma(A_\varphi) = \operatorname{Rn}\varphi \equiv \{\lambda : \ \exists x \in [a,b] \ \lambda = \varphi(x)\},$$

moreover,

$$\lambda \in \sigma_p(A_\varphi) \quad \Leftrightarrow \quad \exists (\alpha,\beta) \subset [a,b] : \ \varphi\big|_{(\alpha,\beta)} \equiv \lambda,$$

and other values of φ belong to $\sigma_r(A_\varphi)$.

Proof.

1) If $\lambda \in \operatorname{Rn}\varphi$, consider

$$g(x) := (\varphi(x) - \lambda) f(x) \in \operatorname{Rn}\varphi.$$

By definition, $\lambda \in \operatorname{Rn}\varphi$ means that

$$\exists x_\lambda \in [a,b] : \quad \lambda = \varphi(x_\lambda).$$

At this point, $g(x_\lambda) = 0$, thus,

$$\operatorname{Rn}(A_\varphi) \neq C[a,b]$$

(the operator is not surjective), so $\lambda \in \sigma(A)$.

2) Let $\lambda \notin \operatorname{Rn}\varphi$. $\operatorname{Rn}\varphi$, being an image of a closed set under continuous map, is a closed set. Therefore,

$$\operatorname{dist}(\lambda, \operatorname{Rn}\varphi) = d > 0,$$

where

$$\operatorname{dist}(\lambda, \operatorname{Rn}\varphi) = \min_{x \in [a,b]} |\lambda - \varphi(x)|.$$

Now, we will construct the resolvent and check its boundedness. This is equivalent to solving the equation

$$(\varphi(x)-\lambda)f(x)=g(x)$$

for an arbitrary given $g(x)$. Formally dividing by the factor $(\varphi(x)-\lambda)$, we get

$$f(x)=\frac{1}{\varphi(x)-\lambda}g(x)=A_{\frac{1}{\varphi(x)-\lambda}}g(x).$$

As we proved in previous lectures, the norm of multiplication operator $A_{\frac{1}{\varphi(x)-\lambda}}$ in $C[a,b]$ is equal to the maximum of $1/(\varphi(x)-\lambda)$:

$$\|R_\lambda(A_\varphi)\|\equiv\left\|A_{\frac{1}{\varphi(x)-\lambda}}\right\|=\max_{[a,b]}\frac{1}{|\varphi(x)-\lambda|}=\frac{1}{\min\limits_{[a,b]}|\varphi(x)-\lambda|}=\frac{1}{d}<\infty,$$

so the resolvent is bounded, and $\lambda\in\rho(A_\varphi)$, which completes the proof of the first statement.

Next, we show that the classification is as stated.

3) When $\lambda\in\sigma_p(A_\varphi)$? For this to be true, the following must hold:

$$\exists f\in C[a,b],\quad f\not\equiv 0:\quad A_\varphi f=\lambda f,$$

that is,

$$(\varphi(x)-\lambda)f(x)=0. \tag{19.1}$$

$f\not\equiv 0$ means that $\exists x_0\in[a,b]$: $f(x_0)\neq 0$. Since f is continuous,

$$\exists(\alpha,\beta)\ni x_0:\quad f\Big|_{(\alpha,\beta)}(x)\neq 0.$$

Thus, for validity of (19.1), it is necessary that

$$\varphi(x)-\lambda=0\quad\text{on}\quad(\alpha,\beta).$$

Why do other points belong to the residual spectrum? Let $\lambda\in\operatorname{Rn}\varphi$; it means that

$$\exists x_\lambda:\quad \lambda=\varphi(x_\lambda).$$

If $g\in\operatorname{Rn}(A_\varphi-\lambda I)$, then

$$g(x)=(\varphi(x)-\lambda)f(x),\qquad g(x_\lambda)=0.$$

Consider the closure of the range in $C[a,b]$; the uniform convergence preserves the values at points: if

$$h(x)\in\overline{\operatorname{Rn}(A_\varphi-\lambda I)},$$

then $h(x_\lambda)=0$. So the closure does not coincide with the entire space, and therefore, by definition, $\lambda\in\sigma_r(A_\varphi)$. □

We have considered only bounded operators, although many concepts carry over to the unbounded case as well. For instance, the position operator in quantum mechanics, i.e., the operator of multiplication by x in $L_2(\mathbb{R})$, has the entire real line as its spectrum.

The Hilbert–Schmidt Theorem: Auxiliary Propositions

The fundamental Hilbert–Schmidt theorem concerns the properties of compact self-adjoint operators. Recall that when discussing the Gram–Schmidt process, we mentioned that, at present, there are two known methods for constructing orthogonal bases in Hilbert spaces. The first method involves taking a closed linearly independent system and orthogonalizing it using the Gram–Schmidt procedure. By a well-known theorem, a closed orthogonal system forms a basis. The second method relies on the Hilbert–Schmidt theorem. Before we state this theorem, we need to establish a few auxiliary results.

Definition 19.2. *Let $A \in B(H)$, where H is a Hilbert space. A subspace $H_0 \subset H$ is an* ***invariant subspace*** *of A if*

$$\forall x_0 \in H_0: \quad Ax \in H_0.$$

Lemma 19.1. *If H_0 is an invariant subspace of A, then $H_0^{\perp}$ is invariant under the operator A^*.*

Proof. Let $x \in H_0$, $y \in H_0^{\perp}$. We must prove that $A^*y \in H_0^{\perp}$. Consider

$$(A^*y, x) = (y, Ax) = 0,$$

since $y \in H_0^{\perp}$ and $Ax \in H_0$, so $A^*y \in H_0^{\perp}$. □

This lemma has an obvious corollary:

Corollary 19.1. *If $A = A^*$ then $H_0^{\perp}$ is invariant under A.*

Recall that for $A = A*$, we know

$$\|A\| = \sup_{\|x\|=1} |(Ax, x)|.$$

Lemma 19.2. *If there exists x_0, $\|x_0\| = 1$, such that*

$$|(Ax_0, x_0)| = \|A\|,$$

then x_0 is an eigenvector of A corresponding to $\lambda = \pm\|A\|$:

$$Ax_0 = \lambda x_0.$$

Proof. Assume that $\dim H \geqslant 2$. Take $z \in H$, $\|z\| = 1$, and $z \perp x_0$. Consider

$$x(t) = x_0 \cos t + z \sin t.$$

For $t \in [0, 2\pi]$, it forms a circle in two-dimensional span of x_0, z. By the Pythagorean theorem, $\|x(t)\| = 1$. Let us plug it into the quadratic form, and consider

$$f(t) = \big(Ax(t), x(t)\big).$$

At zero, we get $f(0) = (Ax_0, x_0)$, and this is an extremum of f, so $f'(0) = 0$. Since

$$f(t) = \big(A(x_0 \cos t + z \sin t), x_0 \cos t + z \sin t\big) = \cos^2 t (Ax_0, x_0) + 2\operatorname{Re}(Ax_0, z) \sin t \cos t + \sin^2 t (Az, z),$$

we obtain

$$0 = f'(0) = 2\operatorname{Re}(Ax_0, z).$$

Changing $z \mapsto iz$, we get $\operatorname{Re}(Ax_0, iz) = -\operatorname{Im}(Ax_0, z) = 0$. Thus, $(Ax_0, z) = 0$, and therefore, $Ax_0 \in z^\perp$, so

$$Ax_0 \in (x_0^\perp)^\perp = \langle x_0 \rangle,$$

and $Ax_0 = \lambda x_0$, which means that x_0 is an eigenvector. The equality $\lambda = \pm\|A\|$ is obvious since the quadratic form equals to λ:

$$|(Ax_0, x_0)| = \lambda (x_0, x_0) = \lambda,$$

and $|(Ax_0, x_0)| = \|A\|$. □

The following is the property of compact operators.

Lemma 19.3. *Let $A \in C(H)$, where H is a Hilbert space. Let $x_n \rightharpoonup x$. Then*

$$(Ax_n, x_n) \to (Ax, x).$$

This means that quadratic form is a weakly continuous function.

Proof. Consider the difference of quadratic forms

$$\Big|(Ax_n, x_n) - (Ax, x)\Big| = \Big|(Ax_n, x_n) - (Ax, x_n) + (Ax, x_n) - (Ax, x)\Big| \leqslant \\ \leqslant \Big|\big(A(x_n - x), x_n\big)\Big| + \Big|\big(Ax, (x_n - x)\big)\Big|.$$

Each summand in this bound tends to zero: by virtue of the Cauchy–Bunyakovsky–Schwarz inequality,

$$\Big|\big(A(x_n - x), x_n\big)\Big| \leqslant \|Ax_n - Ax\| \cdot \|x_n\|,$$

where $\|x_n\|$ is bounded, since $\{x_n\}_{n=1}^{\infty}$ weakly converges to x, and therefore, is weakly bounded (which is, by the Banach–Steinhaus theorem, equivalent to being bounded):

$$\|Ax_n - Ax\| \cdot \|x_n\| \leqslant \|Ax_n - Ax\| \cdot C,$$

and, since A is compact, it makes a converging sequence out of weakly converging, thus,

$$\|Ax_n - Ax\| \cdot C \to 0, \quad \text{and so is} \quad \Big|\big(A(x_n - x), x_n\big)\Big|;$$

as for the second summand, due to Riesz's theorem, it is the evaluation of the functional F_{Ax}, which corresponds to a fixed element Ax, at the element $x_n - x$, so

$$\Big|\big(Ax, (x_n - x)\big)\Big| = \Big|F_{Ax}(x_n - x)\Big| \to 0,$$

which completes the proof. □

Theorem 19.4. *A unit ball in a Hilbert space is weakly sequentially compact. It means that*

$$\forall \{x_n\}_{n=1}^{\infty}, \quad \|x_n\| \leqslant 1,$$

there exists a weakly converging subsequence $x_{n_k} \rightharpoonup x$.

Given the difficulty of proving this theorem, we will omit the complete proof and focus on the key idea, which is the following. For a separable space (while the theorem is valid for unseparable spaces as well), in a unit ball, where $\|x_n\| \leqslant 1$,

$$x_n \rightharpoonup x \quad \Leftrightarrow \quad (x_n, e_k) \to (x, e_k)$$

$\forall k$, where $\{e_k\}_{k=1}^{\infty}$ is an orthonormal basis. This means that $\forall f \in H^*$: $f(x_n) \to f(x)$, which is equivalent to

$$\forall y \in H: \quad (x_n, y) \to (x, y)$$

by Riesz's theorem. Then, y can be expanded into the Fourier series with tail being bounded by some ε, and for (x_n, y), and for the remaining finite sum, we have the coordinate convergence.

There is an analogy for this. Consider continuous functions on a compact set. They have many remarkable properties, one of which is that a continuous function on a compact set attains its maximum and minimum. Similarly, weakly continuous functions on weakly compact sets also attain their maximum and minimum.

Theorem 19.5. *Let $(X, \|\cdot\|)$ be a normed space, and F be a weakly continuous function, i.e.,*

$$x_n \rightharpoonup x \quad \Rightarrow \quad F(x_n) \to F(x).$$

Let M be a weakly compact set. Then

$$\exists x_0 \in M: \quad F(x_0) = \sup_{x\in M} F(x).$$

Proof. By the definition of sup,

$$\exists x_n \in M: \quad F(x_n) \to C \equiv \sup_{x\in M} F(x).$$

Since M is weakly compact, $\exists \{x_{n_k}\}_{k=1}^{\infty}$: $x_{n_k} \to x_0$. Since F is weakly continuous, $F(x_{n_k}) \to F(x_0)$, and, simultaneously, $F(x_n) \to C$, so $F(x_0) = C$. □

Now, we are all set to formulate the Hilbert–Schmidt theorem.

The Hilbert–Schmidt Theorem

Theorem 19.6 (The Hilbert–Schmidt Theorem, for separable case)**.** *Let H be a separable Hilbert space, $\dim H = \infty$. Let $A = A^* \in C(H)$. Then there exists an orthonormal basis $\{e_k\}_{k=1}^{\infty}$ in H that consists of eigenvectors: $Ae_k = \lambda_k e_k$. $\lambda_k \in \mathbb{R}$. Moreover, if λ_n are enumerated such that*

$$|\lambda_1| \geqslant |\lambda_2| \geqslant \ldots |\lambda_k| \geqslant \ldots,$$

then

$$|\lambda_1| = \|A\| \quad \textit{and} \quad \lim_{n\to\infty} \lambda_n = 0.$$

Note that a basis exists only if we consider the eigenvalues with multiplicities. For each eigenvalue λ_k, there may be a set of linearly independent eigenvectors e_{k_j}, and it is necessary to choose an orthogonal basis in their span. However, if all eigenvalues are simple (i.e., to any λ_k, there correspond a unique e_k up to a constant factor), then all the eigenvectors are automatically orthogonal.

Proof.

1) Consider the supremum of the quadratic form associated to A:

$$\sup_{\|x\|=1} \big|(Ax,x)\big|.$$

The unit sphere is weakly compact; A is compact, therefore, (Ax,x) is weakly continuous, and $\exists e_1$, $\|e_1\| = 1$:

$$\big|(Ae_1,e_1)\big| = \sup_{\|x\|=1} \big|(Ax,x)\big|,$$

and it is equal to $\|A\|$, since A is self-adjoint. By Lemma 19.2, e_1 is an eigenvector:

$$Ae_1 = \lambda_1 e_1, \quad |\lambda_1| = \|A\|.$$

2) The subspace $\langle e_1 \rangle \subset H$ is invariant under A, so, due to Lemma 19.1, $H_1 = \langle e_1 \rangle^{\perp}$ is invariant under A as well. Consider the restriction of A to H_1:

$$A\Big|_{H_1} = A_1, \quad A_1 = A_1^*$$

and $A_1 \in C(H_1)$. Thus, by the same argument,

$$\exists e_2 \in H_1 : \quad \big|(A_1 e_2, e_2)\big| = \sup_{\|x\|_{H_1}=1} (A_1 x, x),$$

so

$$Ae_2 = \lambda_2 e_2, \quad |\lambda_2| = \|A_1\| \leqslant \|A\| = |\lambda_1|.$$

3) Through mathematical induction, we can construct a sequence $\{e_k\}_{k=1}^{\infty}$, which is orthogonal, and

$$Ae_k = \lambda_k e_k, \quad \text{and} \quad |\lambda_1| \geqslant |\lambda_2| \geqslant \ldots.$$

Why $\lambda_n \to 0$? Let us prove it by contradiction. Suppose that there exists $C > 0$ and a subsequence $|\lambda_{n_k}| \geqslant C$. Taking the inner product with e_k (this operation is a linear functional), we obtain the Fourier coefficients, which belong to ℓ_2. Thus, $e_{n_k} \rightharpoonup 0$, and, therefore, by the property of compact operators,

$$Ae_{n_k} \overset{\|\cdot\|}{\to} 0,$$

but

$$\|Ae_{n_k}\| = |\lambda_{n_k}| \cdot \|e_{n_k}\| = |\lambda_{n_k}| > C,$$

which leads to a contradiction.

Further, define

$$H_\infty = \langle e_1, e_2, \ldots \rangle^{\perp}.$$

There are two possibilities:

a) $H_\infty = \{0\}$. Then, $\{e_k\}$ is an ONB.

b) $H_\infty \neq \{0\}$. Then, for the restriction of A to this space, we have

$$\left\| A\Big|_{H_\infty} \right\| \leqslant \|A_n\| = |\lambda_n| \to 0,$$

so

$$A\Big|_{H_\infty} = 0.$$

This means that $H_\infty = \mathrm{Ker} A$. Let us take an orthonormal basis in the kernel:

$$\{f_k\}_{k=1}^{N}, \quad N \leqslant \infty.$$

Then $\{e_k\}_{k=1}^{\infty} \cup \{f_k\}_{k=1}^{N}$ is an orthonormal basis in H. □

Example: a Compact Operator in ℓ_2

Consider the following operators in ℓ_2:

1) $Ax = \left(x_1, \frac{x_2}{2}, \ldots, \frac{x_n}{n}, \ldots\right)$, that is,

$$Ae_k = \frac{1}{k} e_k.$$

2) $Ax = \left(x_1, 0, \frac{x_3}{3}, 0, \frac{x_5}{5}, 0, \ldots\right)$,

What is H_∞ in these cases? In case 1, it is $H_\infty = \{0\}$. One can see that $H_\infty = \langle e_2, e_4, \ldots\rangle$ in case 2.

Lecture 20. Applications of the Hilbert–Schmidt Theorem

Discussion of Self-Study Exercises from the Previous Lecture

We begin by discussing the self-study problems from Lecture 18.

1) Prove that $\sigma(AB)\backslash\{0\} = \sigma(BA)\backslash\{0\}$. Additionally, if at least one of operators has a bounded inverse, then $\sigma(AB) = \sigma(BA)$.

   Note that if, for certainty, there exists A^{-1}, then $AB \sim BA$:

$$AB = A(BA)A^{-1}.$$

   Therefore, the spectra coincide.

   Without the assumptions on invertibility of A and B, the problem is a little more difficult. Let $\lambda \neq 0$, and $\lambda \in \rho(BA)$ (for example, $|\lambda| > \|BA\|$). For $|\lambda| > \|BA\|$, let us use the Neumann series for the resolvent:

$$(AB-\lambda I)^{-1} = -\frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{(AB)^k}{\lambda^k} = -\frac{1}{\lambda}\Big(I + \frac{AB}{\lambda} + \frac{ABAB}{\lambda^2} + \dots\Big).$$

   All the summands in the brackets, except for I, have A as the first factor and B as the last one. We can write it in the form

$$-\frac{1}{\lambda}\Big(I + \frac{AB}{\lambda} + \frac{ABAB}{\lambda^2} + \dots\Big) = -\frac{1}{\lambda}\Big(I + \frac{1}{\lambda}A\Big(I + \frac{BA}{\lambda} + \frac{BABA}{\lambda^2} + \dots\Big)B\Big) = \\ = -\frac{1}{\lambda}\Big(I - AR_\lambda(BA)B\Big).$$

   Now, let us look at the formulas obtained and see that the answer has no series included. Thus, it is possible that the same equality holds for other points of the resolvent set, and not only for $|\lambda| > \|BA\|$:

$$R_\lambda(AB) = -\frac{1}{\lambda}\Big(I - AR_\lambda(BA)B, \quad \lambda \in \rho(BA)\Big),$$

   and, similarly,

$$R_\lambda(BA) = -\frac{1}{\lambda}\Big(I - BR_\lambda(AB)A\Big).$$

   It is easy to check that these are indeed resolvents to the corresponding operators by multiplying it by $(AB-\lambda I)$ and $(BA-\lambda I)$ respectively.

   Example, where the spectra are not exactly the same, can be provided by A_ℓ, A_r in ℓ_2:

$$A_\ell A_r = I, \quad \sigma(A_\ell A_r) = \{1\},$$

and

$$A_rA_\ell) = P_{e_1^\perp}, \quad \sigma(A_rA_\ell) = \{0,1\},$$

so the spectra coincide expect for 0. The fact that the spectrum of any projection operator belongs to $\{0,1\}$ will be proved later.

4) Let $\alpha = (\alpha_1, \alpha_2, \dots) \in \ell_\infty$. In ℓ_2, consider

$$A_\alpha x = (\alpha_1 x_1, \alpha_2 x_2, \dots).$$

Find $\sigma(A_\alpha)$.

The point spectrum is easy to find:

$$A_\alpha x = \lambda x \quad \Rightarrow \quad \forall k: \quad \alpha_k x_k = \lambda x_k;$$

if $x_k \neq 0$, then $\lambda = \alpha_k$. For instance,

$$A_\alpha e_k = \alpha_k e_k.$$

Thus, $\sigma_p(A_\alpha) = \{\alpha_k\}_{k=1}^\infty$. Further, note that since the sequence $\{\alpha_k\}_{k=1}^\infty$ is bounded, due to the Bolzano theorem, it has limit points. Therefore, since the spectrum is a closed set,

$$\overline{\{\alpha_k\}_{k=1}^\infty} \subset \sigma(A_\alpha).$$

For instance, consider

$$Ax = (x_1, \frac{x_2}{2}, \frac{x_3}{3}, \dots), \quad \alpha_k = \frac{1}{k},$$

so $\sigma_p(A) = \{1/k\}_{k=1}^\infty$, however, $0 \in \sigma(A)$ and $0 \in \overline{\{1/k\}_{k=1}^\infty}$.

Returning to the general case, we can claim that

$$\overline{\{\alpha_k\}_{k=1}^\infty} = \sigma(A_\alpha).$$

Let us show it. Suppose $\lambda \neq \overline{\{\alpha_k\}_{k=1}^\infty}$. Then, the distance to this set is positive:

$$\inf_{k\geqslant 1} |\alpha_k - \lambda| = d > 0,$$

and we construct a bounded resolvent $R_\lambda(A_\alpha)$, i.e., solve $(A_\alpha - \lambda I)x = y$. In coordinate form, the solution can be expressed as follows:

$$R_\lambda(A_\alpha)y = \Big(\frac{y_1}{\alpha_1 - \lambda}, \frac{y_2}{\alpha_2 - \lambda}, \dots\Big);$$

in fact, this is a multiplication operator corresponding to $\{\beta_k\}_{k=1}^\infty \equiv \{1/(\alpha_k - \lambda)\}_{k=1}^\infty$:

$$R_\lambda(A_\alpha) = A_\beta.$$

The norm of this operator is

$$\|R_\lambda(A_\alpha)\| = \sup_{k\geqslant 1}\frac{1}{|\alpha_k-\lambda|} = \frac{1}{\inf\limits_{k\geqslant 1}|\alpha_k-\lambda|} = \frac{1}{d} < \infty,$$

which completes the proof.

5) Let X be a Banach space and $\Omega\subset\mathbb{C}$ be a nonempty compact set. Prove that

$$\exists A\in B(X): \quad \sigma(A)=A.$$

In Ω, there is a countable dense set, and $\forall n\in\mathbb{N}$ there exists a finite $(1/n)$-net $y1^n$, $y_2^n,\dots,\ y_{m_n}^n$, where the superscript stands for the number of the approximation step. The union $\cup_{n\in\mathbb{N}}\{y_1^n,\dots,y_{m_n}^n\}$ of these nets is a countable set, and it is dense. Let us enumerate it like this: $(\alpha_1,\alpha_2,\dots,\alpha_n,\dots)$.

6) Let $U=U^*=U^{-1}$. Describe all operators of this form.

First, $\|U\|=1$, therefore,

$$\forall\lambda\in\mathbb{C},\ |\lambda|>1, \quad\Rightarrow\quad \lambda\in\rho(U).$$

Suppose $|\lambda|<1$. U has an inverse $U^{-1}=U^*$, $\|U^*\|=1$, and

$$\|\lambda I\| < \frac{1}{\|U^{-1}\|},$$

so the operator $U-\lambda I$ has a bounded inverse, since we can consider λI as a small perturbation of U. Therefore,

$$\sigma(U)\subset\{\lambda\in\mathbb{C}:\ |\lambda|=1\};$$

in fact, for any closed subset of the unit circle, there exists a unitary operator that has this subset as spectrum.

Further, we consider an operator that is self-adjoint and unitary at the same time. So, due to the properties of these operators, $\sigma(U)\subset\{\pm\}$. Next, let us find out whether these options are possible or not. Consider the operator

$$\frac{I-U}{2}+\frac{I+U}{2}=I.$$

Squaring the first one, we get

$$\Big(\frac{I-U}{2}\Big)^2=\frac{I-2U+U^2}{4}=\frac{I-2U}{2},$$

so this is a projection operator. The same can be verified for the second one. Further, consider a vector x from the image of the first operator; denote $\operatorname{Rn}\frac{I-U}{2} =: H_0$:

$$\frac{I-U}{2}x = x, \quad x \in H_0,$$

thus, $(I-U)x = 2x$, so $Ux = -x$. Similarly, $\forall x \in H_1$, $H_1 = \operatorname{Rn}\frac{I+U}{2}$, we have $Ux = x$. Therefore, in the decomposition

$$H = H_0 \oplus H_1,$$

the operator is of the form

$$U = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}.$$

It is possible for any of these spaces to be trivial; for instance, if $U = I$, then $H_0 = \{0\}$.

Now, we are to prove that for the projection operator P, $\sigma(P) \subset \{0,1\}$. Let X be a Banach space decomposed into $X = X_0 \oplus X_1$ with X_j being closed. Let $P: X \to X_0$ be a projection along X_1.

First, let us try to find the eigenvectors:

$$Px = \lambda x, \quad x_0 = \lambda(x_0 + x_1), \quad x_j \in X_j.$$

Rearranging this equation, we obtain

$$(1-\lambda)x_0 = \lambda x_1,$$

and, since X_0 and X_1 have a trivial intersection, it is equal to 0:

$$(1-\lambda)x_0 = 0, \quad \lambda x_1 = 0.$$

When is it true? Taking $\lambda = 0$ and an arbitrary x_1, we have $x_0 = 0$, so $X_0 = \operatorname{Ker} P$. Further, for $\lambda \neq 0$ and $x_1 = 0$, we obtain, for $x_0 \neq 0$, that $\lambda = 1$. Therefore, $\sigma_p(P) \subset \{0,1\}$.

In fact, the spectrum of P is purely discrete. One can prove it by constructing the resolvent. Suppose $\lambda \notin \{0,1\}$. Let us solve

$$(P - \lambda I)x = y;$$

decomposing it with respect to the components X_0 and X_1, we get

$$x_0 - \lambda(x_0 + x_1) = y_0 + y_1.$$

The sum is direct, therefore, the components with index 0 on the left-hand side coincide with those on the right-hand side; the same goes for the index 1:

$$\begin{aligned} (1-\lambda)x_0 &= y_0, \\ -\lambda x_1 &= y_1 \end{aligned} \quad \Rightarrow \quad \begin{aligned} x_0 &= \frac{y_0}{1-\lambda}, \\ x_1 &= -\frac{y_1}{\lambda}. \end{aligned}$$

Thus, the resolvent can be written in the form

$$R_\lambda(P) = \frac{P}{1-\lambda} - \frac{I-P}{\lambda}.$$

Exercises: Applications of the Hilbert–Schmidt Theorem

A while back, we considered the operator

$$(Af)(x) = \int_0^x f(t)\,dt$$

in $L_2[0,1]$, and obtained the bound for its norm. Recall that for any $T : L_2[a,b] \to L_2[a,b]$,

$$(Tf)(x) = \int_a^b K(x,t) f(t)\,dt, \quad K(x,t) \in L_2[a,b]^2,$$

the following bound is valid:

$$\|T\| \leqslant \|K\|_{L_2}.$$

Since for A we have

$$(Af)(x) = \int_0^1 \chi_{t\leqslant x}(t) f(t)\,dt,$$

the bound for the norm is $\|A\| \leqslant \frac{1}{2}$. In fact, the norm is less than this upper bound. Let us find it by employing the Hilbert–Schmidt theorem.

This operator is compact, but not self-adjoint. However, we know that

$$\|A^*A\| = \|A\|^2.$$

This operator is self-adjoint and compact, so one can apply the Hilbert–Schmidt theorem, which gives that the largest eigenvalue is equal to the norm:

$$\lambda_1(A^*A) = \|A^*A\|,$$

where λ_1 is taken with absolute value omitted since the operator is nonnegative:

$$(A^*Ax, x) = (Ax, Ax) = \|Ax\|^2 \geqslant 0.$$

Further,

$$\|A\| = \sqrt{\lambda_1(A^*A)}.$$

We have to find the adjoint operator at first. For the integral operator in $L_2[a,b]$,

$$(Af)(x) = \int_a^b K(x,t)f(t)\,dt,$$

the adjoint is given by

$$(A^*f)(x) = \int_a^b \overline{K(t,x)}f(t)\,dt.$$

Thus, in our case, we have

$$(Af)(x) = \int_0^1 \chi_{t\geqslant x}f(t)\,dt = \int_x^1 f(t)\,dt;$$

one can see that the operator A is not self-adjoint since the integral kernel is not symmetric.

Next, let us consider the eigenequation

$$A^*Af = \lambda f.$$

Expanding the right-hand side, we obtain

$$\int_x^1 \left(\int_0^t f(s)\,ds\right)dt = \lambda f(x). \tag{20.1}$$

To find f, we will differentiate it. Why is it legal, considering $f \in L_2[0,1]$? For an arbitrary function from $L_2[0,1]$, the derivative is not defined, however, this function is an eigenvector of our operator, and is defined by the equation above. The first integration on the left-hand side of the equation gives us a function from $AC[0,1]$, and the second one takes this function to $C^1[0,1]$; therefore, the right-hand side is from $C^1[0,1]$ as well. Repeating this argument, we integrate $f \in C^1[0,1]$ twice, and obtain that $\in C^3[0,1]$, and so on, thus, $f \in C^\infty[0,1]$.

Differentiating with respect to the lower limit of the integral, we obtain

$$-\int_0^x f(s)\,ds = \lambda f'(x), \tag{20.2}$$

and, differentiating again, we arrive at

$$-f(x) = \lambda f''(x),$$

and one can see that $\lambda = 0 \to f \equiv 0$, so λ is positive. Thus, we have

$$f''(x) = -\frac{1}{\lambda}f(x),$$

and the solution is a linear combination of sine and cosine:

$$f(x) = a\sin\frac{x}{\sqrt{\lambda}} + \cos\frac{x}{\sqrt{\lambda}}. \tag{20.3}$$

Note that the differential equation is not equivalent to the integral equation, since the boundary condition must be imposed. Note that it follows from the integral equation (20.1) that $f(1) = 0$, and equation (20.2) implies that $f'(0) = 0$. We have a second-order differential equation, so there are two boundary conditions to be imposed, and we just have found them.

It is better to begin with considering the condition for f', since it is posed at 0, where the sine vanishes. By differentiating (20.3), we obtain

$$f'(x) = \frac{a}{\sqrt{\lambda}}\cos\frac{x}{\sqrt{\lambda}} - \frac{b}{\sqrt{\lambda}}\sin\frac{x}{\sqrt{\lambda}},$$

thus,

$$f'(0) = \frac{a}{\sqrt{\lambda}},$$

so $a = 0$. Therefore,

$$f = b\cos\frac{x}{\sqrt{\lambda}} = 0,$$

and $b \neq 0$. Therefore,

$$\frac{1}{\sqrt{\lambda}} = \frac{\pi}{2} + \pi n, \quad n = 0, 1, \ldots,$$

which gives

$$\lambda_n = -\frac{4}{\pi^2(1+2n)^2}, \quad \lambda_0 = \frac{4}{\pi^2},$$

therefore, $\|A\| = \sqrt{\lambda_0} = 2/\pi$. From the Hilbert–Schmidt theorem it also follows that

$$f_n(x) = \cos\left(\frac{\pi}{2} + \pi n\right)x$$

is an orthogonal basis in $L_2[0,1]$.

Note that if we consider AA^*, we evidently obtain the same result, since nonzero eigenvalues of AA^* and A^*A coincide for a bounded operator A.

Further, let us consider self-study problem 7 from Lecture 18, where the operator of the form

$$A \sim \begin{pmatrix} a & b & 0 & 0 & \ldots \\ b & a & b & 0 & \ldots \\ 0 & b & a & b & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

acts in ℓ_2, a, $b \in \mathbb{R}$. We will construct a similar operator, and find the spectrum.

The aim is to find a unitary isomorphism U and an operator B such that

$$\begin{array}{ccc} \ell_2 & \xrightarrow{\quad A \quad} & \ell_2 \\ U\downarrow & & \downarrow U \\ L_2[0,\pi] & \xrightarrow{\ B\ } & L_2[0,\pi]. \end{array}$$

In $L_2[-\pi,\pi]$, one of the standard orthonormal bases is

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{1}{\sqrt{\pi}}\sin nx, \quad \frac{1}{\sqrt{\pi}}\cos nx, \quad n\in\mathbb{N}. \tag{20.4}$$

In $L_2[0,\pi]$, one can take as a basis either odd or even part of the basis above, so

$$\frac{\sqrt{2}}{\sqrt{\pi}}\sin nx, \quad \text{and} \quad \frac{\sqrt{2}}{\sqrt{\pi}}\cos nx,\ \frac{1}{\sqrt{\pi}}$$

are both bases. It is quite simple to demonstrate; considering the odd (or even) extension of $f\in L_2[0,\pi]$ to $L_2[-\pi,\pi]$, we see that the Fourier series with respect to basis (20.4) consists only of sines (or, respectively, cosines).

Next, let us fix the standard basis $\{e_n\}$ in ℓ_2 and the sine basis $\left\{\frac{E_n\sqrt{2}}{\sqrt{\pi}}\sin nx\right\}$ in $L_2[0,\pi]$. There is an isometric isomorphism U so that $U: e_n \to E_n$. Let us construct it. It can be easily seen that

$$Ae_1 = ae_1 + be_2,$$

and, for $n \geqslant 2$,

$$Ae_n = be_{n-1} + ae_n + be_{n+1}.$$

Therefore, for B we have

$$BE_1 = aE_1 + bE_2 = (a + 2b\cos x)E_1,$$

and, for $n \geqslant 2$,

$$BE_n = \frac{\sqrt{2}}{\sqrt{\pi}}\big(b\sin(n-1)x + a\sin nx + b\sin(n+1)x\big) = \frac{\sqrt{2}}{\sqrt{\pi}}(a+2b\cos x)\sin nx = (a+2b\cos x)E_n,$$

thus, since B acts the same way for any element of the basis,

$$Bf = (a+2b\cos x)f(x).$$

This is a multiplication operator, spectra of which are well-studied, so

$$\sigma(B) = \big[a-2|b|, a+2|b|\big],$$

and, more precisely, it is $\sigma_c(B)$, because the measure of the preimage for each value that $(a+2b\cos x)$ takes is zero. Therefore, the same goes for $\sigma(A)$.

Schatten–von Neumann Classes and Nuclear Operators

Let $A \in C(H)$. Then A^*A is self-adjoint and compact. By the Hilbert–Schmidt theorem,

$$\exists e_k : \quad A^*Ae_k = \lambda_k e_k, \quad \lambda_k \geqslant 0.$$

Define $s_k(A) := \sqrt{\lambda_k(A^*A)}$; we call them *s-numbers* of the operator A.

Definition 20.1. *If $\{c_k\}_{k=1}^{\infty} \in \ell_p$, then we denote $A \in S_p$ and say that A belongs to the Schatten–von Neumann class. The case $p = 1$, S_1, is often referred to as the nuclear class. S_2 is called a Hilbert–Schmidt class.*

For S_1-class operators, the trace is well-defined, i.e., the sum

$$\sum_k (Ae_k, e_k)$$

is independent of the choice of basis.

Note also that S_∞ is the space of all compact operators in H. Now, we remind that the space of compact operators is a closed two-sided ideal in the space of bounded operators; the classes S_p are ideals as well, however, they are not closed. Their closure is S_∞. Thus, the classes S_p can be can be regarded as a certain classification of compact operators.

In perturbation theory, compact perturbations of operators are often considered. Sometimes, stricter conditions must be imposed, such as requiring the perturbation to belong to the class S_p for some p. For example, the well-known Kato's theorem states that the absolutely continuous spectrum of a self-adjoint operator is stable under trace-class perturbations.

Self-Study Exercises

1) Consider, for $A \in C(H)$,

$$Ax = \sum_{k=1}^{N} s_k(A)(x, \varphi_k)\psi_k, \quad N \leqslant \infty,$$

where $\{\varphi_k\}$ is an orthonormal basis and $\{\psi_k\}$ is an orthogonal system. This is called the *Schmidt representation*. Prove the validity of the representation.

2) Consider

$$(Af)(x) = \int_0^1 \min(x,t) f(t)\, dt$$

in $L_2[0,1]$. Find the eigenvalues, eigenvectors, and the corresponding p for S_p class.

3) Let μ_k be solutions of

$$\tan \mu = -\frac{1}{\mu}, \quad \mu > 0.$$

Prove that $\cos \mu_k x$ forms an orthogonal system, but is not a basis. Additionally, prove that being completed by μ_0 that is a solution of

$$\coth \mu = \frac{1}{\mu},$$

$\cos \mu_k x$ forms an orthogonal basis.

This problem is equivalent to the following one. Consider

$$(Af)(x) = \int_0^1 \max(x,t) f(t)\, dt$$

in $L_2[0,1]$. Find the eigenvectors and (asymptotic) eigenvalues.

Lecture 21. Fredholm Theory

Fredholm Theory: Introduction

During this lecture, we focus on the study of Fredholm theory. Its main objective is to analyze the solvability of equations of the form

$$(I - A)x = y,$$

where A is a compact operator, in some Banach space X. The questions posed are as follows. For a given y, does a solution x exist? If not, why? If yes, is it unique?

Clearly, the case of $\dim X = \infty$ is of interest, as such problems are well-studied in linear algebra for finite-dimensional spaces.

Let us first consider the finite-dimensional analog. Suppose $\dim X = n < \infty$, $T \in \mathcal{L}(X)$. Then,

$$\dim \operatorname{Ker} T + \dim \operatorname{Rn} T = n.$$

In infinite-dimensional case, this equality makes no sense. However, we can consider it from another point of view. If $\dim \operatorname{Ker} T = 0$, then the range is the entire space X, so injectivity of T immediately implies its surjectivity, and vice versa, if $\dim \operatorname{Rn} T = n$, then the kernel is trivial, so T is surjective. For operators of the form $I - A$ with compact A, it works the same way, so injectivity and surjectivity become equivalent.

We will prove all the statements for Hilbert spaces, since the key ideas are preserved for Banach spaces, and the proof involves tedious technical work rather than conceptual difficulty. At the same time, Hilbert spaces are more natural here for applications.

In a Hilbert space H, consider the following equations:

$$(I - A)x = y, \quad (1) \qquad\qquad (I - A)x = 0, \quad (2)$$

and, for the adjoint operator,

$$(I - A^*)x = y, \quad (3) \qquad\qquad (I - A^*)x = 0. \quad (4)$$

These equations are very closely related. These equations can be also considered in a Banach space X, with A^* being replaced by A', and for the adjoint operator, the equation is given in the dual space X^*.

Auxiliary Lemmas

As a first step, it is necessary to formulate and prove several auxiliary lemmas, which will simplify the proof of the fundamental theorems in Fredholm theory. We emphasize one more time that we always assume H to be a Hilbert space and A to be a compact operator.

Lemma 21.1. $\dim\operatorname{Ker}(I-A)<\infty$.

Proof. Suppose $x \in \operatorname{Ker}(I-A)$. Then, by definition,

$$(I-A)x=0.$$

Therefore,

$$A\Big|_{\operatorname{Ker}(I-A)}=I.$$

Since A is compact, and the identity operator is compact only in a finite-dimensional space, we conclude that $\dim\operatorname{Ker}(I-A)<\infty$. □

Lemma 21.2. $\operatorname{Rn}(I-A)$ *is closed.*

Proof. Denote

$$H_0 := \operatorname{Ker}(I-A).$$

It is a finite-dimensional closed subspace of H; consider its orthogonal complement

$$H_1=H_0^{\perp}, \qquad H=H_0\oplus H_1.$$

Naturally,

$$\operatorname{Rn}(I-A)=\operatorname{Rn}(I-A)\Big|_{H_1},$$

since $I-A$ takes all the elemts of H_0 to 0.

Recall the previously proved auxiliary statement. If for a bounded operator T in a Banach space X there exists $c>0$ such that $\|Tx\|\geqslant c\|x\|$, then $\operatorname{Rn}T$ is closed.

How do we show that the range of $I-A$ is closed? We will prove a constant c so that the bound above holds for $(I-A)\big|_{H_1}$. Let us show the existence of c by contradiction. Suppose that there is no such constant. The inequality

$$\|(I-A)x\|\geqslant c\|x\|$$

means that Tx is separated from zero for $x\neq 0$; thus, the following means the inverse:

$$\exists x_n, \quad \|x_n\|=1, \quad \text{such that} \quad (I-A)x_n\to 0.$$

Further, $\{x_n\}_{n=1}^{\infty}$ is bounded and A is compact, so the set $\{Ax_n\}_{n=1}^{\infty}$ is precompact. Therefore, there exists a converging subsequence x_{n_k} such that $Ax_{n_k}\to x_0\in H_1$. At the same time, $(I-A)x_{n_k}\to 0$. Therefore,

$$x_{n_k}\to x_0$$

as well. One can see that

$$(I-A)x_0 = 0,$$

so $x_0 \in H_0$, while we supposed that $x_0 \in H_1$. Whence, $x_0 = 0$, which gives a contradiction with the continuity of the norm, since $\|x_n\| = 1$. □

Lemma 21.3. *The following equalities hold:*

$$\operatorname{Ker}(I-A) \oplus_\perp \operatorname{Rn}(I-A^*) = H, \qquad \operatorname{Ker}(I-A^*) \oplus_\perp \operatorname{Rn}(I-A) = H.$$

Remark 21.1. *If we consider $T \in B(H)$ instead of $I-A$ with $A \in C(H)$, the decomposition above becomes*

$$\operatorname{Ker} T \oplus_\perp \overline{\operatorname{Rn} T^*} = H,$$

since for an arbitrary bounded operator, the range need not form a closed subspace.

Proof. These statements are symmetric, so it is sufficient to prove only one of them.

First, we will show that these two subsets are orthogonal. Suppose $x \in \operatorname{Ker}(I-A)$ and $y \in \operatorname{Rn}(I-A^*)$, that is, $\exists z \in H$: $y = (I-A^*)z$. Then,

$$(x,y) = \big(x,(I-A^*)z\big) = \big((I-A)x,z\big) = (0,z) = 0,$$

since $x \in \operatorname{Ker}(I-A)$, so $x \perp y$.

Next, we must show that the sum of these two subspaces is the entire space. Assume that

$$\exists w \perp \Big(\operatorname{Ker}(I-A) \oplus \operatorname{Rn}(I-A^*)\Big).$$

It implies that $w \in \operatorname{Rn}(I-A^*)$, so

$$\forall y \in H: \quad 0 = \big(w,(I-A^*)y\big),$$

as $(I-A^*)y \in \operatorname{Rn}(I-A^*)$. By the definition of the adjoint operator, it can be transfered to the first argument of the inner product as

$$\big(w,(I-A^*)y\big) = \big((I-A)w,y\big),$$

therefore, since this product vanishes for all $y \in H$, we obtain $(I-A)w = 0 \Rightarrow w \in \operatorname{Ker}(I-A)$. At the same time, $w \perp \operatorname{Ker}(I-A)$, thus, $w = 0$. □

Recall that one of our aims was to show that the injectivity and surjectivity are equivalent. The following lemma is the first part of this.

Lemma 21.4. *If $\operatorname{Ker}(I-A) = \{0\}$, then $\operatorname{Rn}(I-A) = H$.*

Proof by contradiction. Suppose that

$$\mathrm{Rn}(I-A) = H_1 \subsetneq H.$$

In further, we are to consider the powers of this operator. For its powers, we have

$$H_n = (I-A)H_{n-1}, \qquad H_k = \mathrm{Rn}(I-A)^k$$

due to the injectivity of $(I-A)$; see the diagram in Fig. 21.1.

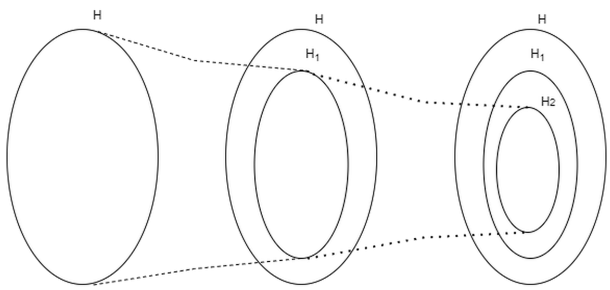

Рис. 21.1. Diagram of $H \supsetneq H_1 \supsetneq H_2$.

Thus, we obtain a chain of inclusions of subspaces

$$H \equiv H_0 \supsetneq H_1 \supsetneq H_2 \supsetneq \dots \supsetneq H_{n-1} \supsetneq H_n \supsetneq \dots .$$

For any $n \in \mathbb{N}$,

$$\exists x_{n-1} \in H_{n-1}, \quad x_{n-1} \perp H_n, \quad \|x_{n-1}\| = 1.$$

Since the set $\{x_n\}_{n=1}^{\infty}$ is bounded and the operator A is compact, the set $\{Ax_n\}_{n=1}^{\infty}$ is precompact, and, therefore, there exists a Cauchy subsequence for it. Consider, for $m > n$,

$$\|Ax_n - Ax_m\|^2 = \|(I-A)x_m - (I-A)x_n - x_m + x_n\|^2.$$

For the first term, we have $(I-A)x_m \in H_{m+1}$; the second one belongs to H_{n+1}, and the third one belongs to H_m. All three first terms together lie in H_{n+1}, while for the last one, we have $x_n \in H_n$. Therefore,

$$x_n \perp (I-A)x_m - (I-A)x_n - x_m,$$

therefore, due to the Pythagorean theorem,

$$\begin{aligned} \|(I-A)x_m - (I-A)x_n - x_m + x_n\|^2 &= \|x_n\|^2 + \|(I-A)x_m - (I-A)x_n - x_m\|^2 = \\ &= 1 + \|(I-A)x_m - (I-A)x_n - x_m\|^2 \geqslant 1, \end{aligned}$$

which contradicts to the existence of a Cauchy subsequence of $\{Ax_n\}_{n=1}^{\infty}$. $\square$

Lemma 21.5. *If* $\mathrm{Rn}(I-A)=H$*, then* $\mathrm{Ker}(I-A)=\{0\}$*.*

Let us point out that Lemmas 21.4, 21.5 together imply that an operator of the form $I-A$ with A being a compact operator is injective if and only if it is surjective.

Proof. It is sufficient to combine two previous lemmas to prove this one. If $\mathrm{Rn}(I-A)=H$, then, due to Lemma 21.3,

$$\mathrm{Ker}(I-A^*)=\{0\}.$$

Further, since A^* is compact as well, Lemma 21.4 gives

$$\mathrm{Rn}(I-A^*)=H,$$

from which, by virtue of Lemma 21.3, we obtain

$$\mathrm{Ker}(I-A)=\{0\}. \qquad \square$$

Note that the the same holds for $(I-A^*)$.

Fredholm Solvability Conditions

Recall the equations from which we started the lecture:

$$(I-A)x=y, \quad (1) \qquad\qquad (I-A)x=0, \quad (2)$$

and, for the adjoint operator,

$$(I-A^*)x=y, \quad (3) \qquad\qquad (I-A^*)x=0. \quad (4)$$

Theorem 21.1. *Equation (1) (or (3)) has a solution iff* y *is orthogonal to solutions of equation (4) (respectively, to solutions of equation (2)).*

Proof. This theorem follows from Lemma 21.3. Suppose that equation (1) has a solution; then, $y\in\mathrm{Rn}(I-A)$. According to Lemma 21.3, $y\perp\mathrm{Ker}(I-A^*)$, which is the space of solutions of equation (4). For equation (3), the proof is similar. $\square$

The Fredholm Alternative

Theorem 21.2 (The Fredholm Alternative)**.** *Either equation (1) has a unique solution for every* y*, or equation (2) admits a nontrivial solution.*

Remark 21.2. *Note that, similarly, either equation (3) has a unique solution for any* y*, or equation (4) has a nontrivial solution.*

Note also that it is possible that for some y *there is no uniqueness of solution, and for some there is no solutions at all. However, if there is a solution for every* $y\in H$*, then it is automatically a unique one.*

Proof. It is quite simple to prove the theorem by combining Lemmas 21.4 and 21.5, according to which

$$\operatorname{Ker}(I-A)=\{0\} \qquad \Leftrightarrow \qquad \operatorname{Rn}(I-A)=H.$$

Let us look at the first possibility. If we have a solution of equation (1) for any y, then $\operatorname{Rn}(I-A)=H$, then the kernel is trivial, i.e., $\operatorname{Ker}(I-A)=\{0\}$, which, in turn, means that equation (2) has only a trivial solution. Therefore, the second possibility is false. The uniqueness of the solution of (1) follows from injectivity of $I-A$.

Next, suppose that equation (2) admits a nontrivial solution. In this case, $\operatorname{Ker}(I-A)\neq\{0\}$, so $\operatorname{Rn}(I-A)\neq H$, and thus, for some $y\in H$, there are no solutions of equation (1). Therefore, the first possibility is false. □

The Third Fredholm Theorem

Note that Lemma 21.1 claims that

$$\dim\operatorname{Ker}(I-A)<\infty,$$

and, since A^* is compact as well,

$$\dim\operatorname{Ker}(I-A^*)<\infty.$$

The third theorem, in turn, claims that these dimensions are equal:

Theorem 21.3.

$$\dim\operatorname{Ker}(I-A)=\dim\operatorname{Ker}(I-A^*).$$

Proof. Denote

$$n=\dim\operatorname{Ker}(I-A), \qquad m=\dim\operatorname{Ker}(I-A^*).$$

Let $\{\varphi_1,\dots,\varphi_n\}$ be an orthonormal basis in $\operatorname{Ker}(I-A)$ and $\{\psi_1,\dots,\psi_m\}$ be an orthonormal basis in $\operatorname{Ker}(I-A^*)$. Suppose that these numbers are different, e.g., $n>m$. Consider the operator T given by

$$Tx=(I-A^*)x+\sum_{i=1}^{m}(x,\psi_i)\varphi_i.$$

Additional part has finite rank, so it is a compact operator, and Tx can be rewritten as

$$Tx=(I-B)x,$$

where $B \in C(H)$ is defined by

$$Bx = A^*x - \sum_{i=1}^{m}(x, \psi_i)\varphi_i.$$

As T is an operator of the same form as earlier, Lemmas 21.1–21.5 and Theorems 21.1–21.2 are valid. Let us show that $\mathrm{Ker}(I-B) = \{0\}$. Suppose $x \in \mathrm{Ker}(I-B)$. Then

$$(I-A^*)x - \sum_{i=1}^{m}(x, \psi_i)\varphi_i = 0.$$

By definition, $(I-A^*)x \in \mathrm{Rn}(I-A^*)$, and

$$\sum_{i=1}^{m}(x, \psi_i)\varphi_i \in \mathrm{Ker}(I-A).$$

According to Lemma 21.3, these subspaces are orthogonal to each other. Therefore, this sum vanishes if both terms vanish:

$$(I-A^*)x = 0, \qquad \sum_{i=1}^{m}(x, \psi_i)\varphi_i = 0.$$

The first equality gives $x \in \mathrm{Ker}(I-A^*)$. Next, since $\{\varphi_1, \ldots, \varphi_n\}$ is a basis, any subsystem of it is linearly independent, so

$$(x, \psi_i) = 0, \quad i = 1, \ldots, m.$$

Recalling that $\{\psi_1, \ldots, \psi_n\}$ is a basis in $\mathrm{Ker}(I-A^*)$, we conclude that $x \perp \mathrm{Ker}(I-A^*)$. Therefore, $x = 0$, which means that the kernel of $T = I-B$ is trivial, and hence, $\mathrm{Rn}(I-B) = H$. This means that the equation

$$(I-B)x = y$$

has a solution $\forall y \in H$. Let us take a look at the equation

$$Tx \equiv (I-A^*)x + \sum_{i=1}^{m}(x, \psi_i)\varphi_i = \varphi_{m+1}. \tag{21.1}$$

(Recall that we supposed that $n > m$, so we have at least one additional element of the basis in $\mathrm{Ker}(I-A)$.) Taking the inner product of this equation with φ_{m+1}, we obtain

$$\Big((I-A^*)x, \varphi_{m+1}\Big) + \sum_{i=1}^{m}(x, \psi_i)(\varphi_i, \varphi_{m+1}) = \|\varphi_{m+1}\|^2.$$

Since $\{\varphi_1, \ldots, \varphi_n\}$ is an orthonormal basis in $\mathrm{Ker}(I-A)$, the right-hand side is 1. On the left-hand side, we have $(I-A^*)x \in \mathrm{Rn}(I-A^*)$ and $\varphi_{m+1} \in \mathrm{Ker}(I-A)$, so the first summand vanishes, and since $(\varphi_i, \varphi_{m+1}) = 0$, $i = 1, \ldots, m$, all terms under the sum vanish as well. Thus, we obtain the contradiction: $0 = 1$. Therefore, $n \leqslant m$. Supposing that $n < m$, one can consider the operator S defined by

$$Sx = (I - A)x + \sum_{i=n} (x, \varphi_i)\psi_i,$$

and, using by reasoning, arrive at a similar contradiction. Thus, $n = m$, which completes the proof. □

History of the Fredholm Theory

E. Fredholm considered integral equations of the form

$$f(x) - \int_a^b K(x,t) f(t)\, dt = g(x). \tag{21.2}$$

It does not matter in which space these equations are considered – whether in Banach spaces such as $C[a,b]$ or $L_p[a,b]$, $p \neq 2$, or in Hilbert spaces such as $L_2[a,b]$. The operator A defined by

$$Af = \int_a^b K(x,t) f(t)\, dt,$$

obviously, must be compact, for the entire Fredholm theory to be applicable here. Note that equation (21.2) is called the *Fredholm equation of 2nd kind.*

It is worth noting that the Fredholm alternative does not mean "either everything is good or everything is bad". Instead, it signifies "either everything is good or almost good", where "good" corresponds to the operator $I - A$ being a bijection, where equation (21.2) has a unique solution, and "almost good" corresponds to the case where the right-hand side must be orthogonal to the kernel of the adjoint operator, which is in fact finite-dimensional, so it imposes only mild constraints on the choice of the right-hand side in the inhomogeneous equation. Additionally, in the latter case, a solution (if any) is not unique, which is a minor flaw.

The *Fredholm equation of 1st kind* is of the form

$$\int_a^b K(x,t) f(t)\, dt = g(x), \qquad Af = g.$$

To solve this equation for generic g, one must find an inverse operator to A. The problem is that, in an infinite-dimensional space, a compact operator has no bounded inverse. This leads to the following issue. Suppose that there is a solution for $g = g_0$, and consider a small perturbation of g_0: $g_0 + \varepsilon g_1$. Applying an unbounded inverse to g_1, one can make it not really small correction. Due to this fact, the problems of this kind are called sometimes *ill-posed problems.*

Corollaries: Spectra of Compact Operators in Banach Spaces

From the Fredholm theory, one can derive many valuable corollaries about the structure of the spectrum of compact operators.

Theorem 21.4. *Let X be a Banach space,* $\dim X = \infty$*, and* $A \in C(X)$*. Then*

1) $0 \in \sigma(A)$*.*

2) If $\lambda \in \sigma(A)$ *and* $\lambda \neq 0$*, then* $\lambda \in \sigma_p(A)$*, and* λ *is an isolated eigenvalue with finite multiplicity:*

$$\dim \operatorname{Ker}(A - \lambda I) < \infty.$$

3) $\forall \varepsilon > 0$ *there exists a finite number of eigenvalues* λ_k *such that* $|\lambda_k| > \varepsilon$*.*

The third property means that, outside some ball centered at 0 in $\mathbb{C}$, there is a finite number of eigenvalues of a compact operator. This implies that the only possible limit point for $\{\lambda_k\}$ is 0.

Proof.

1) We will prove this property by contradiction. Let $0 \in \rho(A)$. Then there exists $A^{-1} \in B(X)$:

$$AA^{-1} = I.$$

Since A is compact and A^{-1} is bounded, the composition is compact, but I can be compact only in X with $\dim X < \infty$. □

2) This property is an immediate corollary of the Fredholm theory. Suppose that $\lambda \neq 0$ and $\lambda \in \sigma(A)$. Constructing the resolvent is equivalent to solving the equation

$$(A - \lambda I)x = y.$$

Since $\lambda \neq 0$, this equation can be rewritten as

$$\left(I - \frac{A}{\lambda}\right)x = -\frac{y}{\lambda}.$$

Consider the possibility given by the Fredholm alternative.

a) For any right-hand side, there is a unique solution. Therefore, $(I - A/\lambda)$ is bijective, which is equivalent to that $A - \lambda I$ is bijective, so $\lambda \in \rho(A)$, which is a contradiction to the assumption $\lambda \in \sigma(A)$.

b) The homogeneous equation

$$\left(I - \frac{A}{\lambda}\right)x = 0$$

has a nontrivial solution. It is equivalent to

$$Ax = \lambda x,$$

so $\lambda \in \sigma_p(A)$.

The multiplicity of λ is finite due to Lemma 21.1. Consider $x \in \operatorname{Ker}(A - \lambda I)$, $\lambda \neq 0$. This is equivalent to $Ax = \lambda x$, that is,

$$x = \frac{A}{\lambda}x.$$

Therefore,

$$I = \frac{A}{\lambda}\Big|_{\operatorname{Ker}(A-\lambda I)},$$

and, since A is compact, and due to the fact that the identity operator is compact only in a finite-dimensional space, $\dim \operatorname{Ker}(A - \lambda I) < \infty$.

Note that there is no such restriction for $\lambda = 0$. It can belong to $\sigma_c(A)$, $\sigma_r(A)$, or $\sigma_p(A)$, and, in the latter case, it may have infinite multiplicity. □

3) To be proved in the next lecture.

Lecture 22. Fredholm Theory: Exercises

Localization of Eigenvalues of a Compact Operator

Note that the proof of the following fact was set aside for discussion in this lecture: $\forall\varepsilon > 0$ there exists a finite number of eigenvalues λ_k such that $|\lambda_k| > \varepsilon$. Now, we are to prove it by contradiction.

Suppose that there exists an infinite number of different eigenvalues, namely, $\{\lambda_k\}_{k=1}^{\infty}$, outside some ε-neighborhood of zero: $|\lambda_k| > \varepsilon$. We stress that assumption that eigenvalues are different is important due to the fact that eigenvectors correspoding to different eigenvalues are linearly independent (one can prove it through mathematical induction). Let e_k satisfy

$$Ae_k = \lambda e_k;$$

consider the linear span $X_n = \langle e_1, e_2, \dots, e_n\rangle$. These spaces are nested:

$$X_1 \subsetneq X_2 \subseteqq \dots \subsetneq X_n \subsetneq X_{n+1} \subsetneq \dots.$$

Due to Riesz's theorem, for any $n \in \mathbb{N}$, there is an element $x_n \in X_n$ such that

$$\operatorname{dist}\big(x_n, X_{n-1}\big) \geqslant 1-\delta, \quad \delta \in (0,1).$$

Since $x_n \in X_n$, one can expand it in terms of the basis

$$x_n = \sum_{k=1}^{n} a_k e_k.$$

Consider $y_n := x_n/\lambda_n$; for this element, we have $\|y_n\| < 1/\varepsilon$. Since A is compact, the set $\{Ay_n\}_{n=1}^{\infty}$ is precompact. We are going to show that it is impossible to choose a Cauchy sequence, which will lead to a contradiction. Expanding y_n and Ay_n, we obtain

$$y_n = \frac{a_n e_n}{\lambda_n} + \sum_{k=1}^{n-1} \frac{a_k e_k}{\lambda_n}, \qquad Ay_n = a_n e_n + \sum_{k=1}^{n-1} \frac{a_k \lambda_k e_k}{\lambda_n} = x_n + z_{n-1}, \qquad z_{n-1} \in X_{n-1},$$

where $z_{n-1} = Ay_n - x_n$ is indeed from X_{n-1}, since the n-th term vanishes. Let us try to choose a Cauchy subsequence in $\{Ay_n\}$; for $m > n$, consider

$$\|Ay_n - Ay_m\| = \|x_n + z_{n-1} + x_m + z_{m-1}\|.$$

Since $x_n + z_{n-1} + z_{m-1} \in X_{m-1}$, due to Riesz's theorem,

$$\|x_n + z_{n-1} + x_m + z_{m-1}\| > 1-\delta, \qquad \delta \in (0,1),$$

therefore, there is no Cauchy subsequence, hence, A is not compact, which is a contradiction. $\square$

Discussion of Self-Study Exercises from the Previous Lecture

Consider some self-study problems from Lecture 20.

1) Consider, for $A \in C(H)$,

$$Ax = \sum_{k=1}^{N} s_k(A)(x, \varphi_k)\psi_k, \quad N \leqslant \infty,$$

where $\{\varphi_k\}$ is an orthonormal basis and $\{\psi_k\}$ is an orthogonal system. This is called the *Schmidt representation*. Prove the validity of the representation.

Let us consider A^*A. This operator is compact and self-adjoint, therefore, due to the Hilbert–Schmidt theorem, there exists an orthonormal basis $\{\varphi_k\}_{k=1}^{\infty}$ such that

$$A^*A\varphi_k = \lambda_k \varphi_k.$$

Additionally, A^*A is nonnegative: $(A^*Ax, x) \geqslant 0$, therefore, $\lambda_k \geqslant 0$. By definition,

$$s_k(A) = \sqrt{\lambda_k(A^*A)}.$$

Since $\{\varphi_k\}_{k=1}^{\infty}$ is a basis, $\forall x \in H$ we have

$$x = \sum_{k=1}^{\infty}(x, \varphi_k)\varphi_k,$$

so

$$Ax = \sum_{k=1}^{N}(x, \varphi_k)A\varphi_k, \tag{22.1}$$

where we exclude numbers k such that $A\varphi_k = 0$, and $N \leqslant \infty$. For $A\varphi_k = 0$, we have $s_k(A) \neq 0$, since for φ_k we have

$$A\varphi_k = 0 \quad \Rightarrow \quad A^*A\varphi_k = 0,$$

so it is an eigenvector corresponding to the eigenvalue $\lambda = 0$. Take only φ_k, $A\varphi_k \neq 0$, and denote

$$\psi_k := \frac{A\varphi_k}{s_k(A)}.$$

Let us verify that this system is orthonormal:

$$(\psi_k, \psi_n) = \frac{(A\varphi_k, A\varphi_n)}{s_k(A)s_n(A)} = \frac{(A^*A\varphi_k, \varphi_n)}{s_k(A)s_n(A)} = \frac{\lambda_k(A^*A)(\varphi_k, \varphi_n)}{s_k(A)s_n(A)} = \delta_{kn},$$

since

$$(\varphi_k, \varphi_n) = \delta_{kn} \quad \text{and} \quad \frac{\lambda_k(A^*A)}{s_k^2(A)} = 1.$$

Thus, for (22.1), we have

$$Ax = \sum_{k=1}^{N} s_k(A)(x, \varphi_k)\psi_k,$$

where numbers k are such that $A\varphi_k \neq 0$.

3) Consider

$$(Af)(x) = \int_0^1 \max(x,t) f(t)\, dt$$

in $L_2[0,1]$. Find the eigenvectors and (asymptotic) eigenvalues.

It is clear that this operator is compact and self-adjoint (since the integral kernel is a continuous symmetric real-valued function). Due to the Hilbert–Schmidt theorem, eigenvectors of this operator form an orthogonal basis. Consider the eigenequation $Af = \lambda f$:

$$\int_0^x x f(t)\, dt + \int_x^1 t f(t)\, dt = \lambda f(x). \tag{22.2}$$

One can see that this equation implies that its solution is a differentiable function, so we can differentiate the equation with respect to x:

$$\int_0^x f(t)\, dt + x f(x) - x f(x) = \lambda f'(x). \tag{22.3}$$

Differentiating once again, we obtain

$$f(x) = \lambda f''(x).$$

Further, we must obtain the boundary conditions. Substituting $x = 0$ and $x = 1$ into (22.2) and (22.3), we get

$$\begin{aligned}
\lambda f(0) &= \int_0^1 t f(t)\, dt,\\
\lambda f(1) &= \int_0^1 f(t)\, dt,\\
\lambda f'(0) &= 0,\\
\lambda f'(1) &= \int_0^1 f(t)\, dt.
\end{aligned}$$

Thus, the following boundary conditions must be imposed:

$$f'(0) = 0, \qquad f(1) = f'(1).$$

The operator is self-adjoint, so eigenvalues are real. Consider the following possibilities.

a) $\lambda > 0$. Then

$$f(x) = ae^{x/\sqrt{\lambda}} + be^{-x/\sqrt{\lambda}},$$

and

$$f'(x) = \frac{1}{\sqrt{\lambda}}\left(ae^{x/\sqrt{\lambda}} - be^{-x/\sqrt{\lambda}}\right),$$

so, $f'(0) = 0$ gives $a = b$, therefore, $f(x) = a\cosh(x/\sqrt{\lambda})$. Further, substituting it into the second boundary condition, we obtain

$$a\cosh\frac{1}{\sqrt{\lambda}} = \frac{a}{\lambda}\sinh\frac{1}{\sqrt{\lambda}}, \quad a \neq 0,$$

which can be rewritten as

$$\coth\frac{1}{\sqrt{\lambda}} = \frac{1}{\sqrt{\lambda}}.$$

Denote $\mu = 1/\sqrt{\lambda}$. The equation $\coth\mu = \mu$ can be solved asymptotically by employing the expansion in Taylor series, however, we will omit this calculation; there exists a unique solution $\mu = \mu_0$, see Fig. 22.1.

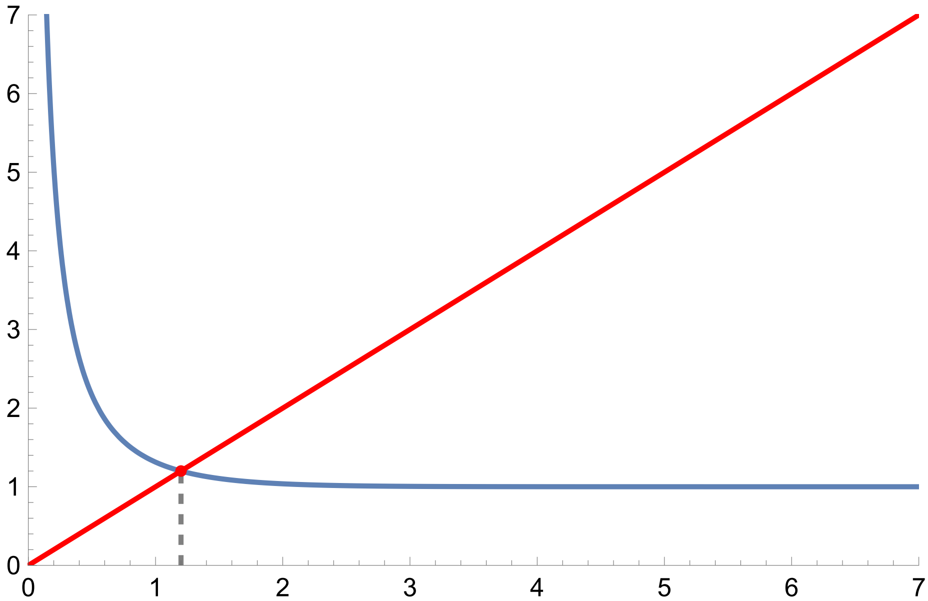

Рис. 22.1. Graphs of $u = \coth\mu$ and $u = \mu$.

Of course, there must be other eigenvectors, since they have to form a basis.

b) $\lambda < 0$. In this case,

$$f(x) = a\cos\frac{x}{\sqrt{-\lambda}} + b\sin\frac{x}{\sqrt{-\lambda}},$$

so,

$$f'(x) = \frac{1}{\sqrt{-\lambda}}\left(-a\sin\frac{x}{\sqrt{-\lambda}} + b\cos\frac{x}{\sqrt{-\lambda}}\right).$$

Therefore, the condition $f'(0) = 0$ gives $b = 0$:

$$f(x) = a\cos\frac{x}{\sqrt{-\lambda}}.$$

Substituting it into $f(1) = f'(1)$, we obtain

$$a\cos\frac{1}{\sqrt{-\lambda}} = -\frac{a}{\sqrt{-\lambda}}\sin\frac{1}{\sqrt{-\lambda}}, \quad a \neq 0.$$

Denoting $\mu := 1/\sqrt{-\lambda}$, we arrive at the equation

$$\tan\mu = -\frac{1}{\mu},$$

see Fig. 22.2.

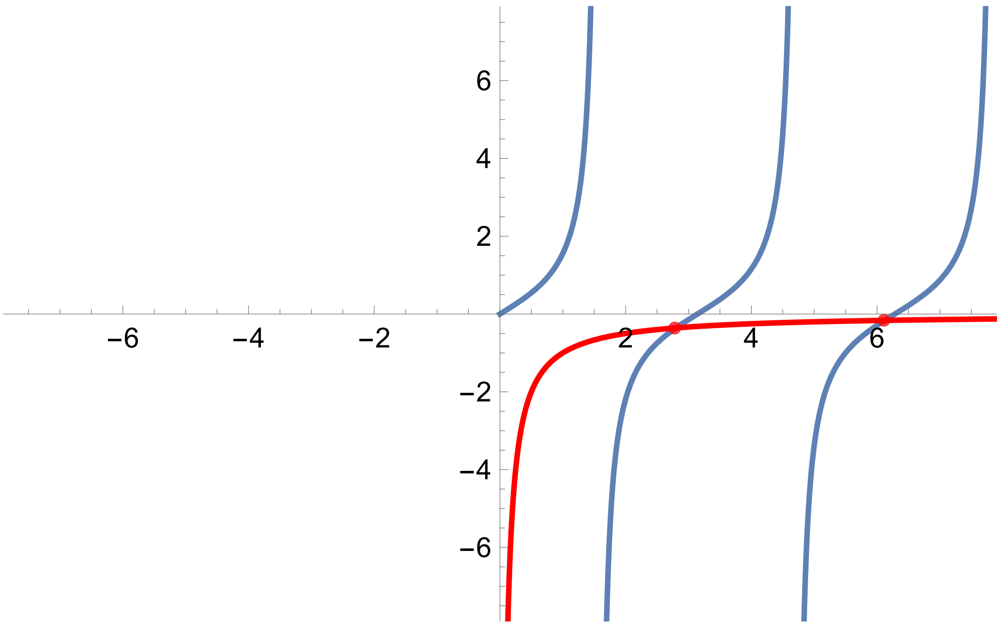

Рис. 22.2. Graphs of $u = \tan\mu$ and $u = -1/\mu$.

There are infinitely many eigenvalues $\{\mu_n\}_{n=1}^{\infty}$, and $\mu_n \sim \pi n$. Thus, our operator belongs to S_p $\forall p > 1$. The functions

$$\coth(\mu_0 x), \quad \cos(\mu_n x), \quad n \in \mathbb{N}$$

form an orthogonal basis in $L_2[0,1]$.

Fredholm Theory: Exercises

1) In $L_2[0,\pi]$, consider

$$f(x) - \lambda\int_0^{\pi}\sin(x+t)f(t)\,dt = g(x).$$

For which λ and g does a solution exist?

Due to the Fredholm Solvability theorem, there exists a solution iff g is orthogonal to the solutions of $(I - A^*)f = 0$. Since the integral kernel $\sin(x+t)$ is symmetric, the operator above is self-adjoint, therefore, g must be orthogonal to the solutions of

$$f(x) - \lambda \int_0^\pi \sin(x+t) f(t)\, dt = 0.$$

Using the sine of sum identity, we can rewrite it as

$$f(x) = \lambda \sin x \int_0^\pi \cos t f(t)\, dt + \lambda \cos x \int_0^\pi \sin t f(t)\, dt.$$

If there is a solution of this equation, it has the following form

$$f_{\text{hom}}(x) = a \sin x + b \cos x.$$

Substituting it into the homogeneous equation, we obtain

$$a \sin x + b \cos x = \lambda \sin x \int_0^\pi \cos t (a \sin x + b \cos x)\, dt + \lambda \cos x \int_0^\pi \sin x (a \sin x + b \cos x)\, dt.$$

Since $\sin x$ and $\cos x$ are linearly independent, the coefficients must match, so

$$a = \lambda \int_0^\pi b \cos^2 t\, dt, \quad a = \lambda \int_0^\pi b \sin^2 t\, dt,$$

so $a = \lambda b \pi/2$ and $b = \lambda a \pi/2$. Therefore,

$$b = \frac{\lambda^2 \pi^2 b}{4}.$$

For $\lambda = \pm 2/\pi$, this equation admits any value of b as a solution, and $a = \pm b$.

For $\lambda = 2/\pi$, we have

$$f_{\text{hom}}(x) = a(\sin x + \cos x),$$

and, for $\lambda = -2/\pi$,

$$f_{\text{hom}}(x) = a(\sin x - \cos x).$$

For $\lambda \neq 0$, the equation admits only a trivial solution $f_{\text{hom}} = 0$. Thus, in this case, there exists a unique solution $\forall g \in L_2[0, \pi]$, moreover,

$$f(x) = g(x) + a \sin x + b \cos x \tag{22.4}$$

for some certain a and b. For $\lambda = \pm 2/\pi$, g must be orthogonal to $f_{\text{hom}}(x) = \sin x \pm \cos x$, and, moreover, all solutions have the form $f(x) + C f_{\text{hom}}(x)$, where $f(x)$

has the form as in (22.4); so, there are infinitely many solutions, and they form a one-dimensional affine space.

The key here is that the integral kernel is a linear combination of two functions. In a more general setting, for

$$K(x,t) = \sum_{i=1}^{n} p_i(x) q_i(t),$$

all steps above can be repeated.

2) In $L_2[0,1]$, consider

$$f(x) - \lambda \int_0^1 K(x,t) f(t)\, dt = \sin(2024\pi x)$$

with $K(x,t)$ of the form

$$K(x,t) = \begin{cases} x(1-t), & t > x, \\ t(1-x), & t < x. \end{cases}$$

In the operator form, this equation becomes $(I - \lambda A) f = g$.

For $\lambda = 0$, we get $f = g$.

For $\lambda \neq 0$, consider first the homogeneous equation, and decompose the integral operator into the sum of two:

$$f(x) - \lambda \int_0^x t(1-x) f(t)\, dt - \lambda \int_x^1 x(1-t) f(t)\, dt = 0.$$

Differentiating this equation, we obtain

$$f'(x) - \lambda x(1-x) f(x) + \lambda \int_0^x t f(t)\, dt + \lambda x(1-x) f(x) - \lambda \int_x^1 (1-t) f(t)\, dt = 0.$$

Since the nonintegral terms cancel out, one can take the second derivative; this gives

$$f''(x) + \lambda x f(x) + \lambda (1-x) f(x) = 0,$$

and, after simplifying it, we obtain

$$f''(x) + \lambda f(x) = 0. \tag{22.5}$$

Since it is the second-order equation, we have to find two boundary conditions. One can see that $f(0) = f(1) = 0$, and that, given these boundary conditions, the operator d^2/dx^2 is negative. Let us show it. First, we take inner product of equation (22.5) with $f(x)$

$$\int f'' f\, dx + \lambda \int f^2\, dx = 0 \quad \Leftrightarrow \quad \int f'^2\, dx + \lambda \int f^2 = 0,$$

therefore, $\lambda > 0$.

Further, a solution of the homogeneous equation is of the form

$$f(x) = a\sin\sqrt{\lambda}x + b\cos\sqrt{\lambda}x.$$

The condition $f(0) = 0$ gives $b = 0$; then, substituting $x = 1$, we obtain

$$a\sin\sqrt{\lambda}x = 0.$$

For $a \neq 0$, we have

$$\sqrt{\lambda} = \pi n, \quad n \in \mathbb{N},$$

or, equivalently, $\lambda_n = \pi^2 n^2$.

Note also that the homogeneous equation $(I - \lambda A)f = 0$ with $\lambda \neq 0$ is equivalent to the following eigenproblem

$$Af = \frac{1}{\lambda} f.$$

That is, $1/(\pi^2 n^2)$ are eigenvalues of A, and the eigenvectors

$$f_n(x) = \sin(\pi n x)$$

form an orthogonal basis; for this basis to become orthonormal, one can put a normalization factor in front of sine:

$$e_n = \sqrt{2}\sin(\pi n x). \tag{22.6}$$

Further, one can try to find a solution expressed in the form of Fourier series. Expanding the right-hand side into the Fourier series, one can obtain the relation for the Fourier coefficients of the solution. Note also that $g(x) = \sin(2024\pi x)$ belongs to family (22.6).

In the case $\lambda = \pi^2 2024^2$, a solution of the homogeneous equation takes the form $f_{\text{hom}}(x) = a\sin(2024\pi x)$, so the right-hand side $g(x)$ is not orthogonal to it, therefore, due to the Fredholm theory, there is no solution for such λ. For any other λ, let us try to find a solution of the form

$$f(x) = \sum_{k=1}^{\infty} a_k e_k.$$

Substituting it into the equation, we obtain

$$\sum_{k=1}^{\infty} a_k e_k - \lambda \sum_{k=1}^{\infty} a_k \frac{1}{\pi^2 k^2} e_k = \frac{e_{2024}}{\sqrt{2}},$$

or simply

$$\sum_{k=1}^{\infty} a_k\Big(1-\frac{\lambda}{\pi^2k^2}\Big)e_k = \frac{e_{2024}}{\sqrt{2}}.$$

Upon carefully examining this equation, one can see that

a) For $\lambda \neq \pi^2k^2$, $k\in\mathbb{N}$:

$$a_{2024} = \frac{1}{\sqrt{2}\Big(1-\frac{\lambda}{\pi^2k^2}\Big)}, \qquad a_k = 0, \quad k\neq 2024.$$

b) For $\lambda = \pi^2k^2$, $k\in\mathbb{N}\backslash\{2024\}$, the coefficient a_k can be arbitrary,

$$a_{2024} = \frac{1}{\sqrt{2}\Big(1-\frac{\lambda}{\pi^2k^2}\Big)}, \qquad a_k = 0, \quad k\neq 2024,$$

and $a_n = 0$, $n\neq k, 2024$, so we have a one-dimensional affine space of solutions.

c) For $\lambda = 2024^2\pi^2$, there are no solutions.

Let us demonstrate another approach to solving problems of this kind using the following equation as an example:

$$f(x)-\lambda\int_0^1 K(x,t)f(t)\,dt = x.$$

Taking the second derivative, we obtain the equation

$$f''(x) = \lambda f(x).$$

Although this equation is the same as the homogeneous one, the boundary conditions must be modified. One can see that

$$f(0)=0, \qquad f(1)=1.$$

Substituting

$$f(x) = a\sin\sqrt{\lambda}x + b\cos\sqrt{\lambda}x,$$

into $f(0)=0$, we get $b=0$; next, substituting it into $f(1)1$, we get

$$a\sin\sqrt{\lambda} = 1,$$

so

$$a = \frac{1}{\sin\sqrt{\lambda}}$$

for $\lambda_n \neq \pi^2n^2$. In that case,

$$f(x) = \frac{\sin\sqrt{\lambda}x}{\sin\sqrt{\lambda}}, \qquad \lambda\neq\pi^2n^2.$$

For $\lambda = \pi^2n^2$, there are no solutions, since $x \not\perp \langle f_{\mathrm{hom},n}\rangle$.

3) (Weyl Theorem). Let $A \in B(X)$ and $\lambda \in \sigma(A) \backslash \sigma_p(A)$. Let $B \in C(X)$. Then $\lambda \in \sigma(A+B)$.

This can be reformulated in the following form: under a compact perturbation B of A, the continuous and residual spectra remains in the spectrum of the operator $A+B$. However, the classification may change.

The proof of this statement is quite simple. Let us prove it by contradiction.

Suppose $\lambda \notin \sigma(A+B)$. Then, there exists a bounded resolvent. Consider

$$A-\lambda I = A+B+\lambda I - B = \big(A+B-\lambda I\big)\big(I-(A+B-\lambda I)^{-1}B\big),$$

where the first factor is invertible, and $(A+B-\lambda I)^{-1}B$ is compact since B is compact and $(A+B-\lambda I)^{-1}$ is bounded, whence, the second factor is a Fredholm operator. Let us examine the possibilities for the second factor, as dictated by the Fredholm alternative. The first possibility is that the equation

$$\big(I-(A+B-\lambda I)^{-1}B\big)f = g$$

has a unique solution for any $g \in H$, that is, $\big(I-(A+B-\lambda I)^{-1}B\big)$ is invertible. Therefore, $A-\lambda I$ is invertible as well, but this is not true since $\lambda \in \sigma(A)$. Another possibility is that the homogeneous equation

$$\big(I-(A+B-\lambda I)^{-1}B\big)x = 0$$

admits a nontrivial solution, so x is an eigenvector correspoding to the eigenvalue 0; therefore, it is an eigenvector of A corresponding to λ, which is not true, since $\lambda \notin \sigma_p(A)$. □

Self-Study Exercises

1) Consider

$$f(x)-\lambda\int_a^b K(x,t)f(t)\,dt = g(x), \qquad K(x,t) = \sum_{i=1}^{n} p_i(x)q_i(x), \tag{22.7}$$

where the functions $\{p_i\}_{i=1}^n$ are linearly independent. Then a solution has the form

$$f(x) = g(x) + \sum_{i=1}^{n} c_i p_i(x),$$

where $\{c_i\}_{i=1}^n$ is a solution of the following system of equations:

$$\sum_{j=1}^{n} a_{ij}c_j = b_i, \quad i = 1,2,\ldots,n.$$

Find a_{ij}, b_i.

Note that equation (22.7) can be considered in any Banach space of functions where all the integrals and functions are well-defined.

2) Consider

$$f(x) - \lambda \int_0^\pi \cos(x-t) f(t)\, dt = g(x).$$

For which $\lambda \in \mathbb{C}$ and $g \in L_2[0,\pi]$ does a solution exist?

Lecture 23. Unbounded Operators: Introduction

Volterra Operators

Recall first what we proved in the last two lectures. Let $A \in C(X)$, where X is a Banach space, $\dim X = \infty$. Then

1) $0 \in \sigma(A)$.

2) If $\lambda \in \sigma(A)$, $\lambda \neq 0$, then $\lambda \in \sigma_p(A)$ and $\dim \operatorname{Ker} A - \lambda I < \infty$.

3) $\forall \varepsilon > 0$ there exists a finite number of eigenvalues λ such that $|\lambda| > \varepsilon$.

Now, we proceed to the following topic.

Definition 23.1. *A is called a* ***Volterra operator*** *if* $A \in C(X)$ *and* $\sigma(A) = \{0\}$.

The importance of these operators is due to the Fredholm Alternative. Consider

$$(I - A)x = y.$$

Recall that there are two possibilities: either there exists a unique solution x for any $y \in X$, or there exists a nontrivial solution x_0 to the homogeneous equation

$$(I - A)x_0 = 0.$$

If A is a Volterra operator, then for any $\lambda \in \mathbb{C}$ the equation

$$(I - \lambda A)x = y$$

has a unique solution for any $y \in X$, that is, for Volterra operators, the first possibility of the alternative always holds. To explain this, let us consider the following possibilities.

1) If $\lambda = 0$, then $x = y$, since $(I - \lambda A)$ becomes I.

2) If $\lambda \neq 0$, then
$$\Big(A - \frac{1}{\lambda} I)x = -\frac{y}{\lambda},$$
and $1/\lambda \notin \sigma(A)$, so there exists a bounded resolvent
$$R_{1/\lambda}(A) = \Big(A - \frac{1}{\lambda} I\Big)^{-1}.$$

Examples of Volterra Operators

1) Consider

$$(Af)(x) = \int_0^x f(t)\,dt$$

in $C[0,1]$ or $L_2[0,1]$.

First, let us show that the point spectrum is empty: $\sigma_p(A) = \varnothing$. Let us try to solve the eigenequation

$$Af = \lambda f, \qquad \int_0^x f(t)\,dt = \lambda f(x).$$

Note that the eigenfunction must be a differentiable function, since it is equal to the integral of itself, and the integral increases the number of derivatives. Moreover, if there is an eigenfunction f, one can see that $f \in C^\infty[0,1]$, since the aforementioned reasoning can be repeated infinitely many times. Differentiating the equation, we get

$$f(x) = \lambda f'(x) \tag{23.1}$$

with the Cauchy condition

$$f(0) = 0. \tag{23.2}$$

Thus, from (23.1), we obtain

$$f(x) = Ce^{x/\lambda}, \qquad \lambda \neq 0.$$

Substituting it into (23.2), we get $C = 0$, so $f(x) \equiv 0$, which is not an eigenfunction. Further, if $\lambda = 0$, then $f(x) \equiv 0$ as well.

Another approach is to construct the resolvent. Let $|\lambda| > \|A\|$, then the Neumann series is valid:

$$R_\lambda(A) = -\frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{A^k}{\lambda^k}. \tag{23.3}$$

Recall the expression obtained in the previous lectures:

$$A^k f = \int -0^x \frac{(x-t)^{k-1}}{(k-1)!\lambda^k} f(t)\,dt.$$

Due to the factorial in denominator, the sum in (23.3) converges, so one can interchange the summation and integration and obtain

$$R_\lambda(A)f = -\frac{1}{\lambda}\Big(f + \frac{1}{\lambda}\int_0^x e^{(x-t)/\lambda} f(t)\,dt\Big).$$

With this expression, one can drop the condition $|\lambda| > \|A\|$ since it holds for **any** nonzero λ.

2) Consider a slightly more difficult example

$$(Af)(x) = \int_a^x K(x,t)f(t)\,dt$$

in $C[a,b]$ with the condition $K \in C[a \leqslant t \leqslant x]$ (this kind of $K(x,t)$ is called a *triangle kernel*) or in $L_2[a,b]$ with the following conditions: K is measurable and $|K(x,t)| \leqslant M$. This is a Volterra operator as well, and we will consider it in detail a little later.

Unbounded Operators: Introduction

Let us recall the Hellinger–Toeplitz theorem: If $A \in \mathcal{L}(H)$, where H is a Hilbert space, and $\forall x, y \in H$

$$(Ax, y) = (x, Ay),$$

then $A \in B(H)$.

Therefore, an unbounded operator cannot be defined in the entire space H, and it must have some domain. Consider the example

$$Af = if'$$

in $L_2[0,1]$ this operator is called the momentum operator. Consider, for instance,

$$f_n(x) = \sin \pi n x, \quad \|f_n\| = \frac{1}{\sqrt{2}}.$$

For these functions,

$$\|Af\| = \frac{\pi n}{\sqrt{2}} \to \infty \quad \text{as} \quad n \to \infty.$$

One of the standard domains for this operator is

$$\mathcal{D}(A) = \{f \in W_2^1[0,1] : \; f(0) = f(1) = 0\},$$

where

$$(Af, g) = \int_0^1 if'(x)\overline{g(x)}\,dx = if(x)\overline{g(x)}\Big|_0^1 - \int_0^1 if(x)\overline{g'(x)}\,dx,$$

and the nonintegral terms vanish due to the boundary conditions in $\mathcal{D}(A)$, therefore,

$$(Af, g) = (f, Ag).$$

Note that this operator is not self-adjoint since the domain of adjoint operator is larger.

Our further studies, we will focus on the study of unbounded operators in Hilbert spaces. Let H be a Hilbert space, and $A \in \mathcal{L}(H)$ be an unbounded operator. By definition, a **domain** of the operator A is a subset $D(A) \subset H$ such that the following condition holds:

$$x \in \mathcal{D}(A) \quad \Rightarrow \quad Ax \in H.$$

The largest domain of A is called sometimes a *natural domain*; usually, this domain has no effective description.

Graph of an Operator. Graph Norm. Closed Operators

Definition 23.2. *A **graph of an operator** A is a set $\Gamma(A) \subset H \times H$ such that*

$$\Gamma(A) = \big\{\{x, Ax\} \in H \times H : \; x \in \mathcal{D}(A)\big\}.$$

Definition 23.3. $\|x\|_A = \|x\| + \|Ax\|$ *is called a **graph norm** of an operator A.*

If $A \in B(H)$, due to the fact that the boundedness is equivalent to the continuity, one can take a sequence $x_n \to x$, and then $Ax_n \to Ax$. In general, this does not work this way for unbounded operators. However, there is a class of unbounded operators, for which this property is preserved:

Definition 23.4. *A is called a **closed operator** if $\Gamma(A)$ is closed in $H \times H$ with respect to the graph norm $\|\cdot\|_A$.*

For a closed operator A, if $x_n \in \mathcal{D}(A)$ and $x_n \to x$, $Ax_n \to y$, then $x \in \mathcal{D}(A)$ and $y = Ax$.

Example of a Nonclosed Operator

Consider $A : L_2[0,1] \to L_2[0,1]$, $Af = f(0) \cdot 1$ with $\mathcal{D}(A) = C[0,1]$. This operator **is not closed**; let us show it. Consider the functions $f_n \to 0 \in L_2[0,1]$ as depicted in Fig. 23.1.

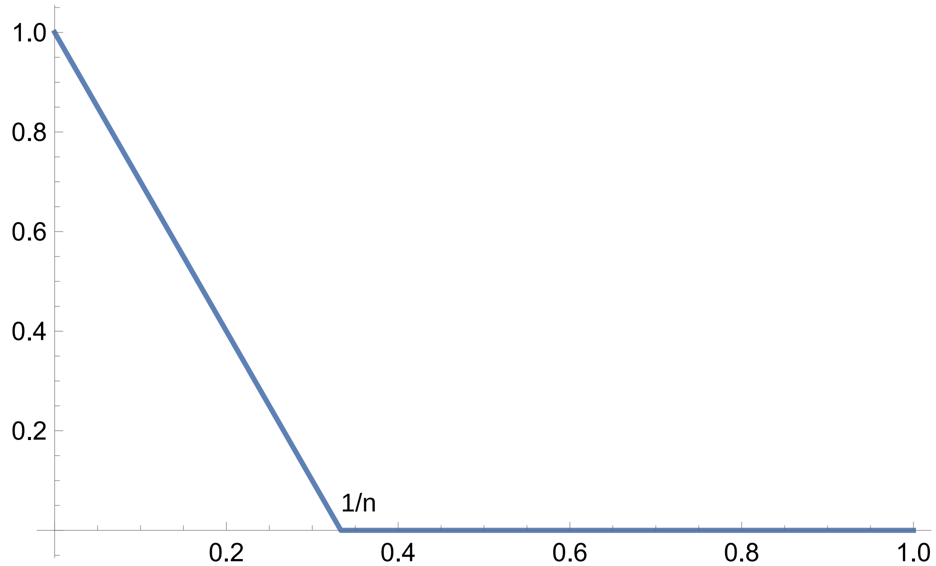

Рис. 23.1. Graph of f_n.

These functions converge to zero in $L_2[0,1]$, and we have $\{f_n, 1\} \in \Gamma(A)$, however, for the limit function, the point $\{0,1\}$ cannot belong to $\Gamma(A)$, since it is a graph of linear operator.

Note that, for a closed operator, the graph norm is equivalent to the original norm of H.

Closure of an Operator. Closable Operators

What can we do if the operator is not closed? We can consider $\overline{\Gamma(A)}$. Then

1) If $\overline{\Gamma(A)}$ is a graph of some operator B, then we call B a **closure** of A and denote $B = \overline{A}$, and A is called a **closable operator**.

2) If $\overline{\Gamma(A)}$ is mot a graph, i.e., $\{0, y\} \in \overline{\Gamma(A)}$, $y \neq 0$, then A is not closable.

In the example above, we face an nonclosable operator.

Let us also consider the operator $Af = if'$ with domain

$$\mathcal{D}(A) = \{f \in C^{\infty}[0,1],\ f(0) = f(1) = 0\}.$$

Then, for $\overline{A}$, we have

$$\mathcal{D}(\overline{A}) = \{f \in W)2^1[0,1],\ f(0) = f(1) = 0\},$$

that is, A is closable.

Definition 23.5. *If $\overline{\Gamma(A)}$ is a graph of some operator, then A is called a **closable operator**, and its closure is $\overline{A}$ with $\Gamma(\overline{A}) = \overline{\Gamma(A)}$.*

The Adjoint of an Unbounded Operator

One of the key concepts in operator theory, the adjoint operator, can be extended to the case of unbounded operators in a natural way.

Definition 23.6. *Let $A \in \mathcal{L}(H)$, $\overline{\mathcal{D}(A)} = H$. Define the domain of A^* by*

$$\mathcal{D}(A^*) = \{h \in H:\ x \mapsto (Ax, h) \text{ is a bounded functional in } H,\ x \in \mathcal{D}(A)\}.$$

By Riesz's theorem, $(Ax, h) = (x, z)$, and then we define $z := A^*h$.

It is necessary for $\mathcal{D}(A)$ to be dense in H, so for z to be unique; otherwise, the adjoint operator is not well-defined.

In the previous examples, for $Af = f(0) \cdot 1$ with $\mathcal{D}(A) = C[0,1]$, the domain is dense in $L_2[0,1]$; the same holds for $Af = if'$ with $\mathcal{D}(A) = \overset{\circ}{W}{}_2^1[0,1]$.

Theorem 23.1. *Let $A \in \mathcal{L}(H)$, $\overline{D(A)} = H$. Then A^* is closed.*

Proof. Let us consider the operator

$$W : H\times H \to H\times H, \qquad W\{x,y\} = \{-y,x\}.$$

We are going to show that $\Gamma(A^*) = \big(W\Gamma(A)\big)^{\perp}$; it is known that the orthogonal complement is closed, and, in that case, so is $\Gamma(A^*)$. Consider $(Ax,y) = (x,A^*y)$; equivalently, $(Ax,y) - (x,A^*y) = 0$. Further, it can be rewritten as

$$\{-Ax,x\} \perp \{y,A^*x\} \quad \text{in} \quad H\times H,$$

since

$$(\{x_1,y_1\},\{x_2,y_2\})_{H\times H} \overset{\text{def}}{=} (x_1,x_2)_H + (y_1,y_2)_H,$$

so

$$(\{-Ax,x\},\{y,A^*x\}) = (-Ax,y) + (x,A^*y) = 0.$$

Thus, since

$$\{-Ax,x\} = W\{x,Ax\} \quad \text{and} \quad \{y,A^*x\} \in \Gamma(A^*),$$

we see that $\Gamma(A^*) = (W\Gamma(A))^{\perp}$. □

Theorem 23.2. *Let $A \in \mathcal{L}(H)$, $\overline{D(A)} = H$. Then*

$$H = \mathrm{Ker}A^* \oplus_{\perp} \overline{\mathrm{Rn}A}.$$

Note that we proved this statement for the operators of the form $I - A$, and we did not use the boundedness of this operator.

Proof. Let us first show that $\mathrm{Ker}A^* \perp \mathrm{Rn}A$. If $x \in \mathrm{Ker}A^*$, $y \in \mathrm{Rn}A$, then $A^*x = 0$ and $y = Az$ for some $z \in \mathcal{D}(A)$. Further,

$$(x,y) = (x,Az) = (A^*x,z) = 0,$$

since $A^*x = 0$.

Next, one can verify that

$$\mathrm{Ker}A^* \perp \overline{\mathrm{Rn}A}$$

by considering the limit points of $\mathrm{Rn}A$.

Now, the only point to be proved is that $\mathrm{Ker}A^* \oplus \overline{\mathrm{Rn}A} = H$. Suppose that there exists $h \in H$ such that

$$h \perp \left(\mathrm{Ker}A^* \oplus_{\perp} \overline{\mathrm{Rn}A}\right).$$

For $x \in \mathcal{D}(A)$, consider

$$0 = (Ax,h) = (x,A^*h),$$

where $h \perp Ax$, so the first inner product vanishes, and the second one is defined for $x \in \mathcal{D}(A)$, $\overline{\mathcal{D}(A)} = H$, therefore, $A^*h = 0$, and thus, $h \in \mathrm{Ker}A^*$, which means $h = 0$, since h is orthogonal to this space. □

Closability of a Densely Defined Operator

Theorem 23.3. *Let $A \in \mathcal{L}(H)$, $\overline{D(A)} = H$. Then A is closable iff $\overline{D(A^*)} = H$.*

If A^* is densely defined, then $\overline{A} = A^{**}$ (note that it may not coincide with A for unbounded A).

Consider $Af = if'$ with $\mathcal{D}(A) = \overset{\circ}{W}{}_2^1$:

$$A^* f = if', \quad \mathcal{D}(A) = W_2^1[0,1], \tag{23.4}$$

and these operators are not self-adjoint, since the adjoint one has different domain. Both of these operators are closed, and $\overline{A} = A = A^{**}$.

Note also that there is a difference between symmetric and self-adjoint operators, and it is due to the difference in domains. However, for some symmetric operators, there exist so-called *self-adjoint extensions.* By definition, a symmetric operator satisfies

$$(Af, g) = (f, Ag) \quad \forall f, g \in \mathcal{D}(A).$$

We also know that

$$(Af, g) = (f, A^* g), \quad \forall f \in \mathcal{D}(A),\ \forall g \in \mathcal{D}(A^*),$$

so, for a symmetric operator A, the following holds: $A \subset A^*$, which means that $\mathcal{D}(A) \subset \mathcal{D}(A^*)$ and

$$A^*\big|_{\mathcal{D}(A)} = A,$$

and the closure may be non-self-adjoint. In further lectures, we will construct all self-adjoint extensions of (23.4).

Proof of Theorem 23.3. Consider the second power of W:

$$W : \{x, y\} = \{-y, x\},$$

that is, $W^2 = -I$. Since A is densely defined, for $W\Gamma(A^*)$, we have

$$W\Gamma(A^*) = W\big(W(\Gamma(A))\big)^{\perp} = \Big(W^2\Gamma(A)\Big)^{\perp} = (\Gamma(A))^{\perp},$$

and

$$\Gamma((A^*)^*) = (W\Gamma(A^*))^{\perp} = \big((\Gamma(A))^{\perp}\big)^{\perp} = \overline{\Gamma(A)} = \Gamma(\overline{A}). \qquad \square$$

Example: Nonexistence of A^{**}

Consider again the following example: in $L_2[0,1]$,

$$Af = f(0)\cdot 1, \qquad \mathcal{D}(A) = C[0,1].$$

What is A^*?

We know that

$$\operatorname{Ker} A^* \oplus_\perp \overline{\operatorname{Rn} A} = L_2[0,1]$$

and

$$\mathcal{D}(A^*) = \{g \in L_2[0,1] : \ (Af, g) \text{ is a bounded functional}\}.$$

Further,

$$\int_0^1 f(0)\cdot 1\cdot \overline{g(x)}\,dx = f(0)\int_0^1 \overline{g(x)}\,dx.$$

is the very functional that must be bounded. However, this is not a continuous functional on the domain of A; there is a way to make it continuous by restricting to the case where

$$\int_0^1 g(x)\,dx = 0.$$

Thus, $g \perp 1$, and 1 is from the range of A, therefore, $g \in \operatorname{Ker} A^*$, so $A^* = 0$. (It is not a typical situtation, however, it is quite typical for nonclosable operators.) Since A^* is not densely defined, there is no $(A^*)^*$, and, therefore, it is impossible to construct $\overline{A}$.

Inverse of an Unbounded Operator

Theorem 23.4. *Let $A \in \mathcal{L}(H)$, $\overline{\mathcal{D}(A)} = H$. Then there exists $A^{-1} \in B(H)$,*

$$A^{-1} : \operatorname{Rn} A \to H,$$

iff

$$\exists c > 0 : \quad \forall x \in \mathcal{D}(A) : \quad \|Ax\| \geqslant c\|x\|.$$

Proof. $\Rightarrow$. Since there exists A^{-1}, then

$$\forall y \in \operatorname{Rn} A : \quad \|A^{-1}y\| \leqslant \|A^{-1}\|\cdot\|y\|,$$

and $\operatorname{Ker} A = \{0\}$. There exists a unique x: $y = Ax$, and

$$\|x\| \leqslant \|A^{-1}\|\cdot\|Ax\|, \quad C = \frac{1}{\|A^{-1}\|}.$$

Further, for $\Leftarrow$, we have $\|Ax\| \geqslant c\|x\|$, which is equivalent to $\mathbf{Ker}A = \{0\}$, therefore, there exists $A^{-1} : \mathrm{Rn}A \to H$. Let us show the boundedness of A^{-1}:

$$\|y\| \geqslant c\|A^{-1}y\|,$$

so

$$\|A^{-1}y\| \leqslant \frac{1}{c}\|y\| \quad \Rightarrow \quad A^{-1} \in B(H). \qquad \square$$

Lecture 24. Symmetric Operators

Discussion of the Self-Study Problems from the Previous Lecture

We begin by considering self-study problem 3 from Lecture 22.

Consider

$$f(x) - \lambda \int_0^\pi \cos(x-t) f(t)\,dt = g(x).$$

For which $\lambda \in \mathbb{C}$ and $g \in L_2[0,\pi]$ does a solution exist?

Rewriting the equation as

$$f(x) + g(x) - \lambda \int_0^\pi \cos x \cos t f(t)\,dt - \lambda \int_0^\pi \sin x \sin t f(t)\,dt,$$

one can see that any solution takes the form

$$f(x) = g(x) + a\cos x + b\sin x.$$

Substituting it back into the equation, we get

$$g(x) + a\cos x + b\sin x = g(x) - \lambda \cos x \int_0^\pi \cos t\big(a\cos t + b\sin t + g(t)\big)\,dt$$
$$-\lambda \sin x \int_0^\pi \sin t\big(a\cos t + b\sin t + g(t)\big)\,dt,$$

so $g(x)$ cancels out on both sides. Using following equalities

$$\int_0^\pi \cos t \sin t\,dt = 0, \qquad \int_0^\pi \cos^2 t\,dt = \int_0^\pi \sin^2 t\,dt = \frac{\pi}{2},$$

we obtain

$$a\cos x + b\sin x - \lambda a\frac{\pi}{2}\cos x - \lambda \cos x \int_0^\pi a g(t)\cos t\,dt - \lambda b\frac{\pi}{2}\sin x - \lambda \sin x \int_0^\pi \sin t g(t)\,dt,$$

leading to the equation for the coefficients a and b of a solution:

$$a\Big(1 - \lambda\frac{\pi}{2}\Big) = \lambda \int_0^\pi g(t)\cos t\,dt,$$
$$b\Big(1 - \lambda\frac{\pi}{2}\Big) = \lambda \int_0^\pi g(t)\sin t\,dt.$$

For $\lambda \neq \pi/2$, we have

$$a = \frac{\lambda}{1-\lambda\pi/2}\int_0^\pi g(t)\cos t\,dt,$$
$$a = \frac{\lambda}{1-\lambda\pi/2}\int_0^\pi g(t)\sin t\,dt$$

$\forall g \in L_2[0,\pi]$. One can also see that for $\lambda = \pi/2$, a solution exists if $g(t) \perp \{\sin t, \cos t\}$.

Defect Numbers and Relatively Bounded Operators. Properties of Relatively Bounded Operators

Recall that, for a densely defined unbounded operator A, we have

$$H = \mathrm{Ker}A^* \oplus_\perp \overline{\mathrm{Rn}A}.$$

Definition 24.1. *A **defect number** of $A \in \mathcal{L}(H)$, $\overline{\mathcal{D}(A)} = H$, is defined as*

$$d_A := \dim(\mathrm{Rn}A)^\perp \equiv \dim \mathrm{Ker}A^*.$$

The defect number measures how far the operator is from being bijective. Recall that the bound

$$\|A\| \geqslant c\|x\|$$

guarantees the existence of a left inverse, which takes $y \in \mathrm{Rn}A$ to $A_\ell^{-1} y \in H$. It may be of interest to know how large this range is; the defect number indicates the dimension of its orthogonal complement.

Definition 24.2. *Let $A, B \in \mathcal{L}(H)$. The operator B is called **relatively bounded** with respect to A if*

1) $\mathcal{D}(B) \subseteq \mathcal{D}(A)$,

2) $\exists a \in [0,1)$, $b \geqslant 0$: $\forall x \in \mathcal{D}(A)$ $\|Bx\| \leqslant a\|Ax\| + b\|x\|$.

To see why this notion can be very practical, consider the differential operator

$$(A+B)y = -y'' + iy', \quad y \in \overset{circ}{W}{}^2_2[0,1],$$

where $Ay = -y''$, $By = iy'$. The operator $A+B$ retains many of the properties of A; note that $A+B$ is relatively bounded with respect to A.

Theorem 24.1 (Properties of Relatively Bounded Operators)**.** *Let A be closed and $\exists A^{-1} \in B(H)$. Suppose B is relatively bounded with respect to A. Then:*

1) $A+B$ is closed.

2) If $b = 0$, then $\exists (A+B)^{-1} \in B(H)$.

3) If $b = 0$, then $d_{A+B} = d_A$.

Recall that

$$\exists A^{-1} \in B(H) \quad \Leftrightarrow \exists c > 0\ \forall x \in \mathcal{D}(A): \ \|Ax\| \geqslant c\|x\|.$$

For $(A+B)^{-1} \in B(H)$ in point 2, one replaces the constant c with $c(1-a)$.

Proof.

1) We will show that

$$\|x\|_A \leqslant \|x\|_{A+B} \leqslant C_1\|x\|_A,$$

which means that the graph norms $\|\cdot\|_A$, $\|\cdot\|_{A+B}$. Consequently, if A is closed, it follows that $A+B$ is closed as well.

Suppose that $b \leqslant a$. Consider

$$\|x\|_{A+B} = \|x\| + \|(A+B)x\| \leqslant \|x\| + \|Ax\| + \|Bx\|;$$

since B is relatively bounded with respect to A, we have

$$\|x\| + \|Ax\| + \|Bx\| \leqslant \|x\| + \|Ax\| + a\|Ax\| + b\|x\|$$

with $b\|x\| \leqslant a\|x\|$, so

$$\|x\| + \|Ax\| + a\|Ax\| + b\|x\| \leqslant (1+a)\|x\|_A.$$

Next, for the second bound, consider

$$\|x\|_A = \|x\| + \|Ax\| = \|x\| + \|(A+B)x - Bx\| \leqslant \|x\| + \|(A+B)x\| + \|Bx\|,$$

and, using the definition of a relatively bounded operator, we get

$$\|x\| + \|(A+B)x\| + \|Bx\| \leqslant \|x\| + \|(A+B)x\| + a\|Ax\| + a\|x\|,$$

which implies

$$(1-a)\|x\|_A \leqslant \|x\|_{A+B}.$$

Now suppose $b > a$. Since B is relatively bounded with respect to A,

$$\|Bx\| \leqslant a\|Ax\| + b\|x\|;$$

multiplying this inequality by a/b, we get

$$\left\|\frac{a}{b}Bx\right\| \leqslant a\left\|\frac{a}{b}Ax\right\| + a\|x\|.$$

Define the operators $\widetilde{B} := aB/b$ and $\widetilde{A} := aA/b$ in place of B and A respectively. Then $\widetilde{B}$ is relatively bounded with respect to A with $\widetilde{b} = a$, $\widetilde{a} = a$.

Being multiplied by a constant, the new graph norm remains equivalent to the old one, so the same argument as in the first case applies to $\widetilde{a}$, $\widetilde{b}$. Hence, the conclusion follows similarly. □

2) Given $b = 0$, we have
$$\|Bx\| \leqslant a\|Ax\|.$$
By assumption, $c\|x\| \leqslant \|Ax\|$, so
$$c\|x\| \leqslant \|Ax\| = \|(A+B)x - Bx\| \leqslant \|(A+B)x\| + \|Bx\| \leqslant \|(A+B)x\| + a\|Ax\|,$$
that is,
$$(1-a)\|Ax\| \leqslant \|(A+B)x\|.$$
Thus,
$$c(1-a)\|x\| \leqslant \|(A+B)x\|,$$
so there exists a bounded operator $(A+B)^{-1}$: $\mathrm{Rn}(A+B) \to H$. □

3) We will prove this point by contradiction. Suppose $d_A < d_{A+B}$. This means $\mathrm{Rn}A \supset \mathrm{Rn}(A+B)$, i.e.m
$$\exists y \in \mathrm{Rn}A: \quad y \perp \mathrm{Rn}(A+B).$$
Since $y \in \mathrm{Rn}A$, there exists $x \in \mathcal{D}(A)$ such that $y = Ax$. Then
$$(Ax, Bx) = (Ax, Bx) - (Ax, (A+B)x),$$
since the second term vanishes due to $y = Ax \perp \mathrm{Rn}(A+B)$. Further,
$$(Ax, Bx) - (Ax, (A+B)x) = -(Ax, Ax) = -\|y\|^2,$$
thus, $y = 0$.

Now suppose $d_A > d_{A+B}$. Then there exists $y \in \mathrm{Rn}(A+B)$ such that $y \perp \mathrm{Rn}A$. Hence there is $x \in \mathrm{Rn}(A+B)$ such that $y = (A+B)x$. Consider
$$(Ax, Bx) = (Ax, Bx) - (Ax, (A+B)x),$$
where again the second term vanishes. So,
$$(Ax, Bx) - (Ax, (A+B)x) = -(Ax, Ax),$$
which implies
$$\|Ax\|^2 = |(Ax, Ax)| = |(Ax, Bx)|.$$
Using the Cauchy–Bunyakovsky inequality, we obtain
$$|(Ax, Bx)| \leqslant \|Ax\| \cdot \|Bx\| \leqslant \|Ax\| \cdot a\|Ax\|,$$
where $a < 1$, thus,
$$\|Ax\|^2 < \|Ax\|^2,$$
which is a contradiction. □

Regular Points of a Closed Operator. The Set of Regular Points

Let A be a closed operator.

Definition 24.3. $\lambda \in \mathbb{C}$ *is called a* ***regular point*** *of an operator* A *if there exists a bounded inverse* $(A-\lambda I)^{-1}$: $\mathrm{Rn}(A-\lambda I) \to H$*. The* ***set of regular points*** *is denoted by* $\widehat{\rho}(A)$.

For λ to be a point of the *resolvent set* $\rho(A)$, an additional condition must be imposed: there is an isomorphism between the range of $A-\lambda I$ and H. Formally,

$$\big(\lambda \in \widehat{\rho}(A) \text{ and } d_{A-\lambda I} = 0\big) \quad \Leftrightarrow \quad \lambda \in \rho(A).$$

If the orthogonal complement of $\mathrm{Rn}(A-\lambda I)$ is zero, then $d_{A-\lambda I} = 0$, and there is a bijection between $\mathrm{Rn}(A-\lambda I)$ and H.

In previous lectures, we established that the resolvent set is open. Here we show that the set of regular points is also open:

Theorem 24.2. $\widehat{\rho}(A)$ *is open.*

Proof. Suppose $\lambda_0 \in \widehat{\rho}(A)$, and take λ such that $|\lambda-\lambda_0| < \varepsilon$ (we will specify the value of ε during the proof). We want to show $\lambda \in \widehat{\rho}(A)$. Observe that

$$(A-\lambda I) = (A-\lambda_0 I) - (\lambda-\lambda_0)I = \widetilde{A} + \widetilde{B},$$

where $\widetilde{A} = A - \lambda_0 I$ and $\widetilde{B} = (\lambda-\lambda_0)I$. If $\widetilde{B}$ is relatively bounded with respect to $\widetilde{A}$ with $b = 0$, then Theorem 24.1 immediately shows that $\widetilde{A}+\widetilde{B}$ has a bounded inverse. In essence, we aim to prove

$$\|\widetilde{B}x\| \leqslant a\|\widetilde{A}x\|, \qquad 0 \leqslant a < 1.$$

Since $\widetilde{A} = A - \lambda_0 I$ is assumed invertible,

$$ca\|x\| \leqslant a\|(A-\lambda I)x\|,$$

so for for $|\lambda-\lambda_0| < ca$, $\widetilde{B}$ is relatively bounded with respect to $\widetilde{A} = A - \lambda I$ such that

$$\|\widetilde{B}x\| \leqslant a\|\widetilde{A}x\|.$$

Hence, $\widetilde{A}+\widetilde{B}$ has a bounded inverse whenever $\varepsilon < ca$. Therefore, $\lambda \in \widehat{\rho}(A)$, so the set $\widehat{\rho}(A)$ is open. □

Symmetric Operators. Deficiency Indices of a Symmetric Operator. Self-Adjointness of a Closed Symmetric Operator with $n_{\pm}(A) = 0$

We now introduce basic definitions related to symmetric operators, which play a central role in this theory.

Definition 24.4. *Let $A \in \mathcal{L}(H)$, $\overline{\mathcal{D}(A)} = H$. We say that A is **symmetric** if $\forall x, y \in \mathcal{D}(A)$:*

$$(Ax, y) = (x, Ay).$$

Definition 24.5. *An operator B is called an **extension** of A if*

1) $\mathcal{D}(A) \subset \mathcal{D}(B)$,

2) $B\big|_{\mathcal{D}(A)} = A$.

Building on this definition, one may also define a symmetric operator via $A \subset A^*$.

Consider the following example:

$$Af = if', \qquad \mathcal{D}(A) = \overset{\circ}{W}{}_2^1[0,1].$$

This is a symmetric operator but not self-adjoint.

Theorem 24.3. *If A is symmetric, then the sets*

$$\mathbb{C}^+ = \{\lambda \in \mathbb{C} : \ \operatorname{Im}\lambda > 0\}, \qquad \mathbb{C}^- = \{\lambda \in \mathbb{C} : \ \operatorname{Im}\lambda < 0\}$$

are both contained in $\widehat{\rho}(A)$.

Note that we established this result earlier when discussing bounded operators. However, the proof does not require that the operator considered is bounded. Nevertheless, we repeat the proof here for completeness.

Proof. Consider

$$\|(A - \lambda I)x\|^2 = \big((A - \alpha I - i\beta I)x, (A - \alpha I - i\beta I)x\big).$$

Grouping $A - \alpha I$ together, we see

$$\big((A - \alpha I - i\beta I)x, (A - \alpha I - i\beta I)x\big) = \|(A - \alpha I - i\beta I)x\|^2 + i\beta\big((A - \alpha I)x, x\big) - \\ - i\beta\big(x, (A - \alpha I)x\big) + |\beta|^2\|x\|^2.$$

Since A is symmetric, the second and the third terms, which are both imaginary, cancel. Thus,

$$\|(A - \lambda I)x\|^2 \geqslant |\beta|^2\|x\|^2.$$

Therefore, there exists a bounded inverse, so $\lambda = \alpha + i\beta \in \widehat{\rho}(A)$. □

Nothing can be concluded directly about the real axis: it may or may not be in the spectrum. This typically gives two connected components, the upper and lower half-planes. Therefore, since the defect number d_A is constant on each connected component of $\widehat{\rho}(A)$, it is sufficient to define only two defect numbers for a symmetric operator.

Definition 24.6. *Let A be a symmetric operator. Then define*

$$n_+(A) = d_A(\lambda), \qquad \operatorname{Im}\lambda > 0,$$

and

$$n_-(A) = d_A(\lambda), \qquad \operatorname{Im}\lambda < 0.$$

$n_\pm(A)$ are called ***deficiency indices****.*

Note that for $\operatorname{Im}\lambda > 0$,

$$n_+(A) = \dim\left(\operatorname{Rn}(A - \lambda I)\right)^\perp = \dim\operatorname{Ker}(A^* - \lambda I),$$

and the same holds for $n_-(A)$, $\operatorname{Im}\lambda < 0$.

Further, since $d_A(\lambda)$ is constant on each connected component of $\hat{\rho}(A)$, the following hold:

$$n_+(A) = \dim\left(\operatorname{Rn}(A - iI)\right)^\perp, \qquad n_-(A) = \dim\left(\operatorname{Rn}(A + iI)\right)^\perp.$$

Theorem 24.4. *Let A be a closed symmetric operator. Then*

$$A = A^* \quad \Leftrightarrow \quad n_-(A) = n_+(A) = 0.$$

Proof. Suppose $A = A^*$. Then $\rho(A) \supset \mathbb{C}^+ \cup \mathbb{C}^-$, therefore, there exists

$$(A \pm iI)^{-1} \in B(H): \ H \to H.$$

Hence the range of $A \pm iI$ is the whole space H, and the deficiency indices are zero.

Conversely, if A is symmetric and $n_\pm(A) = 0$, then $\operatorname{Rn}(A \pm iI) = H$. Since a symmetric operator A can be defined by $A \subset A^*$, it remains to prove that $\mathcal{D}(A) = \mathcal{D}(A^*)$. Take an arbitrary $y \in \mathcal{D}(A^*)$; consider $h := (A^* + iI)y$. Since the range of $A \pm iI$ is the whole space,

$$\exists x_0 \in D(A): \quad h = (A + iI)x_0.$$

Further, since A^* is an extension of A, we have

$$h = (A + iI)x_0 = (A^* + iI)x_0,$$

and, for our y,

$$(A^* + iI)x_0 = (A^* + iI)y.$$

Since $\mathrm{Rn}(A \pm iI) = H$, we have

$$\dim \mathrm{Ker}(A^* \pm iI) = 0.$$

Therefore, $x_0 = y$, so $y \in \mathcal{D}(A)$. □

Examples: Non-Self-Adjoint Symmetric Operators and Self-Adjoint Extensions

Consider the operator

$$Af = if', \qquad \mathcal{D}(A) = \overset{\circ}{W}{}_2^1[0,1].$$

Recall that

$$(Af, g) = \int_0^1 if'\overline{g}\,dx = if\overline{g}\Big|_0^1 + \int_0^1 f\overline{ig'}\,dx,$$

where nonintegral terms vanish due to the boundary conditions for f, while g does not necessarily meet those conditions. Moreover, if $g \in \mathcal{D}(A^*)$,

$$(Af, g) = (f, A^*g).$$

What is the domain of the adjoint operator? One can see that $\mathcal{D}(A) = W_2^1[0,1]$, which is strictly larger than $\overset{\circ}{W}{}_2^1[0,1]$, and the property $A \subset A^*$ obviously holds. Further, the deficiency indices of A are the dimensions of the solution spaces to

$$(A^* \pm iI)g = 0.$$

This equation is of the form

$$ig' = \mp ig,$$

so the solutions are

$$g(x) = e^{\mp x}.$$

Therefore, $n_\pm(A) = 1$. In the next lecture, we will show that there exists a self-adjoint extension if and only if the deficiency indices are equal; that is, the operator $A = i\,d/dx$ with domain specified above admits self-adjoint extensions. One can see that all self-adjoint extensions A_θ of A forms a one-parameter family:

$$\mathcal{D}(A_\theta) = \{y \in W_2^1[0,1] : \ y(0) = \theta y(1), \ |\theta| = 1\}.$$

Next, when considering differential operators in a higher-dimensional domains, e.g., the Laplacian in the unit disk with both Dirichlet and Neumann conditions, the deficiency indices may become infinite. Consider, for instance,

$$-\Delta u = -u_{xx} - u_{yy} \quad \text{in} \quad \Omega = \{x^2 + y^2 < 1\}$$

with the boundary conditions

$$u\big|_{\partial\Omega} = 0, \qquad \frac{\partial u}{\partial n}\big|_{\partial\Omega} = 0.$$

This operator admits self-adjoint extensions with the boundary condition of the form

$$u + h\frac{\partial u}{\partial n}\big|_{\partial\Omega} = 0,$$

where h is some real-valued function. This function is a parameter in the family of self-adjoint extensions; there are infinitely many real-valued functions, so is for extensions.

In certain situations, there can be a unique self-adjoint extension:

Definition 24.7. *A symmetric operator A is called* ***essentially self-adjoint****, if $\overline{A}$ is self-adjoint.*

In this case, $n_{\pm}(A) = 0$, and $\overline{A}$ is the unique self-adjoint extension of A.

Situation is worse when we have distinct deficiency indices for an operator. For instance, if

$$n_+ = m > n = n_-,$$

then there are symmetric extensions of the given operator. It is possible to obtain an extension with deficiency indices $(m-n, 0)$, and it is not possible to construct a self-adjoint extension.

Self-Study Exercises

1) Prove that

$$A_\theta f = if', \quad \mathcal{D}(A_\theta) = \{y \in W_2^1[0,1] : \ y(0) = \theta y(1), \ |\theta| = 1\}$$

is self-adjoint.

2) Consider

$$(Af)(x) = \int_0^x K(x,t) f(t)\, dt$$

in $L_2[0,1]$, with $|K(x,t)| \leqslant M \ \forall x,t \in [0,1]$. Prove that $r(A) = 0$ (recall that $r(A)$ denotes the spectral radius).

3) Consider

$$(Af)(x) = \frac{1}{x}\int_0^x f(t)\,dt.$$

Find $\sigma_p(A)$ and conclude that $A \notin C(L_2[0,1])$.

Lecture 25. Isometric Operators and the Cayley Transform. Self-Adjoint Extensions of Symmetric Operators

Isometric Operators. Regular Points and Deficiency Indices of an Isometric Operator

In this lecture, we continue our study of self-adjoint extensions. First, we examine some properties of isometric and unitary operators.

Definition 25.1. *An operator $V : H \to H$, where H is a Hilbert space, is called **isometry** (or **isometric operator**) if $\forall x \in \mathcal{D}(V)$:*

$$\|Vx\| = \|x\|.$$

Proposition 25.1. *If $|z| \neq 1$, then $z \in \widehat{\rho}(V)$.*

Proof. Consider

$$\|Vx\| = \|Vx - zx + zx\| \leqslant \|(V - zI)x\| + |z|\|x\|,$$

where $\|Vx\|$ equals $\|x\|$ by the definition of isometry. Therefore,

$$\|(V - zI)x\| \geqslant \big|1 - |z|\big|\|x\|.$$

Since $|z| \neq 1$, we have $\big|1 - |z|\big| > 0$. Hence there exists an inverse to $V - zI$. □

Recall that for a symmetric operator ($A \subset A^*$), we defined the deficiency indices in the previous lecture:

$$n_+(A) = \dim\big(\mathrm{Rn}(A - iI)\big)^\perp, \qquad n_-(A) = \dim\big(\mathrm{Rn}(A + iI)\big)^\perp,$$

where n_+ corresponds to the upper half-plane and n_- corresponds to the lower half-plane. Similarly, for an isometric operator, one can define a pair of deficiency indices: for the interior and exterior of the unit circle.

Definition 25.2. *Define*

$$n_i(V) = \dim\big(\mathrm{Rn}(V - zI)\big)^\perp, \qquad |z| < 1,$$

and

$$n_e(V) = \dim\big(\mathrm{Rn}(V - zI)\big)^\perp, \qquad |z| > 1.$$

Note that the deficiency indices are stable under a relatively bounded perturbation; that is,

$$n_{\pm}(A+B) = n_{\pm}(A),$$

if B is relatively bounded with respect to A. The same holds for $n_{i/e}(V)$ for an isometry V. Since $\|V\| = 1$, the operator zI with $|z| < 1$ is relatively bounded with respect to V. Therefore,

$$n_i(V) = \dim\left(\operatorname{Rn} V\right)^{\perp}.$$

One can see it through the following reasoning. Since the defect number is constant on each connected component of the set of regular points, one can take $0 \in \{|z| < 1\}$ as the value of z.

In the exterior of $\{|z| < 1\}$, the operator zI is "larger" than V; thus, one may remove V in the definition of $n_e(V)$:

$$n_e(V) = \dim\left(\operatorname{Rn}(zI)\right)^{\perp}, \qquad |z| > 1.$$

Since z is just a constant, this essentially becomes the identity map on the domain of V, so

$$n_e(V) = \dim\left(\mathcal{D}(V)\right)^{\perp}.$$

Unitary Operators

When does an isometric operator become a unitary operator?

Definition 25.3. *Let V be an isometric operator. Then V is called **unitary** if V is bijection.*

In terms of deficiency indices, one can define a unitary operator by

$$n_i(V) = n_e(V) = 0.$$

Cayley Transform

Using properties of isometric operators, we return to the theory of extensions of symmetric operators.

Definition 25.4. *Let A be a densely defined symmetric operator. Define an operator*

$$V = (A - iI)(A + iI)^{-1}.$$

*Then V is called a **Cayley transform** of A.*

Note that V is well-defined since $A+iI$ is invertible; we proved earlier that for $\lambda = \alpha + i\beta$ with $\beta \neq 0$,

$$\|(A-\lambda I)x\| \geqslant |\beta|\|x\|.$$

Also note that

$$\mathcal{D}(V) = \mathrm{Rn}(A+iI),$$

where $y = (A+iI)x$, $x \in \mathcal{D}(A)$. Regarding the action of V, we have

$$Vy = (A-iI)x.$$

Theorem 25.1. *V is isometric.*

Proof. Consider

$$\|Vy\|^2 = \|Ax\|^2 + \|x\|^2,$$

and

$$\|y\|^2 = \|Ax\|^2 + \|x\|^2.$$

Thus, $\|Vy\| = \|y\|$. □

Using the Cayley transform, one takes a symmetric operator to an isometry. However, this is not a one-to-one correspondence; not all isometries can be obtained by such a transform.

Theorem 25.2 (Properties of the Cayley Transform). *1)* $1 \notin \sigma_p(V)$.

2) $\overline{\mathrm{Rn}(V-I)} = H$.

Proof.

1) Suppose $\exists y$: $Vy = y$. Since $y = (A+iI)x \mapsto (A-iI)x$, we get

$$(A-iI)x = (A+iI)x,$$

so

$$-ix = ix.$$

Hence $x = 0$, and so $y = 0$. □

2) Suppose $\exists z \perp \mathrm{Rn}(V-I)$. Then

$$\big(z, (V-I)y\big) = 0, \qquad \forall y \in \mathcal{D}(V).$$

Decompose $(V-I)y$:

$$\big(z, (A-iI)x - (A+iI)x\big) = 0,$$

that is,

$$2i(z,x) = 0 \qquad \forall x \in \mathcal{D}(A).$$

Since A is densely defined, i.e., $\overline{\mathcal{D}(A)} = H$, we conclude $z \perp H \Rightarrow z = 0$. □

In fact, the second point implies the first one:

Proposition 25.2. *Let V be an isometric operator and $\overline{\mathrm{Rn}(V-I)} = H$. Then $1 \notin \sigma_p(V)$.*

Proof. Let $\exists x$: $Vx = x$, or, equivalently, $(V-I)x = 0$. Then $\forall z \in \mathcal{D}(V)$:

$$\big((V-I)x, z\big) = 0,$$

since the first factor vanishes. Expanding the brackets, we get

$$(Vx, z) - (x, z) = 0.$$

Since V is isometric, it preserves the inner product; thus,

$$(Vx, z) - (Vx, Vz) = 0,$$

or equivalently,

$$(Vx, z - Vz) = 0,$$

where $z - Vz \in \mathrm{Rn}(I-V)$, which is dense in H. Hence $Vx = 0$, so is x: $x = 0$. Thus, x cannot be an eigenvector. □

This means that there is a one-to-one correspondence between densely defined symmetric operators and isometric operators with dense range:

Theorem 25.3. *There is a one-to-one correspondence $A \leftrightarrow V$, where A is a symmetric operator with $\overline{\mathcal{D}(A)} = H$ and V is an isometric operator with $\overline{\mathrm{Rn}(V-I)} = H$.*

Proof. We already proved that to each A, there corresponds some V via the Cayley transform. We must now show that to any such V, one can associate an operator A.

Let V be an isometric operator with dense range. Consider $y = (A+iI)x$ and $Vy = (A-iI)x$; let us subtract one from another:

$$(V-I)y = -2ix,$$

where x belongs to the domain of A; further, one can see that $\mathcal{D}(A)$ coincides with $\mathrm{Rn}(V-I)$:

$$x \in \mathcal{D}(A) \quad \Rightarrow \quad \mathcal{D}(A) = \mathrm{Rn}(V-I).$$

Since the range of $V-I$ is dense, the domain of A is also dense.

To determine the action of the operator, let us sum the equalities $y = (A+iI)x$ and $Vy = (A-iI)x$. Specifically, observe:

$$(V+I)y = 2Ax,$$

therefore,

$$Ax = \frac{1}{2}(V+I)y, \qquad y = -2i(V-I)^{-1}x.$$

the inverse to $V - I$ exists due to the fact that $1 \notin \sigma_p(V)$. To complete the formula for A, let us substitute y of the form as above:

$$Ax = -i(V+I)(V-I)^{-1}x.$$

Finally, we need to check that the operator A, constructed in this manner, is indeed symmetric. Recall that a densely defined operator A is called symmetric if

$$(Ax,y) = (x,Ay) \qquad \forall x,y \in \mathcal{D}(A).$$

Consider the quadratic form (Ax,x), $x \in \mathcal{D}(A)$:

$$(Ax,x) = (x,Ax) = \overline{(Ax,x)},$$

so the quadratic form associated to a symmetric operator is always real-valued. Conversely, one can show that if the quadratic form associated to an operator is real-valued, then this operator is symmetric (this is a simple exercise, the idea of which is to rewrite (Ax,y) in terms of quadratic forms with $\big(A(x \pm y),(x \pm y)\big)$ and $\big(A(x \pm iy),(x \pm iy)\big)$). Therefore, it is sufficient to prove that (Ax,x) is real-valued. Rewrite it in the form

$$(Ax,x) = \frac{1}{2}\Big((V+I)y, -\frac{1}{2i}(V-I)y\Big) = \frac{1}{4i}\Big((V+I)y,(V-I)y\Big).$$

Expanding the brackets, we obtain

$$\frac{1}{4i}\Big((V+I)y,(V-I)y\Big) = \frac{1}{4i}\Big((Vy,Vy) - (Vy,y) + (y,Vy) - (y,y)\Big).$$

The first and the last terms cancel each other, so

$$(Ax,x) = \frac{1}{4i}\Big(-(Vy,y) + (y,Vy)\Big) = \frac{1}{4i}\Big(-(Vy,y) + \overline{(Vy,y)}\Big).$$

Let $(Vy,y) = \alpha + i\beta$ with α, $\beta \in \mathbb{R}$. Then we have

$$(Ax,x) = \frac{1}{4i}(-\alpha - i\beta + \alpha - i\beta) = -\frac{\beta}{2},$$

so (Ax,x) is real-valued, and, therefore, A is symmetric. □

Domain of the Adjoint of a Symmetric Operator

Theorem 25.4. *Let A be a closed symmetric operator with $\overline{\mathcal{D}(A)} = H$. Then*

$$\mathcal{D}(A^*) = \mathcal{D}(A) + \operatorname{Ker}(A^* - iI) + \operatorname{Ker}(A^* + iI).$$

Note that this sum is, in fact, direct.

Proof. The inclusion $\supset$ is evident because all spaces on the right-hand side lie in $\mathcal{D}(A^*)$. Indeed, for a symmetric operator A, we have $A \subset A^*$, so $\mathcal{D}(A) \subset \mathcal{D}(A^*)$. Likewise, kernels of $A^* \pm iI$ obviously belong to $\mathcal{D}(A^*)$ as well.

Thus, our aim is to show the converse inclusion $\subset$. Recall that the entire space H can be decomposed into

$$H = \operatorname{Rn}(A - iI) \oplus \operatorname{Ker}(A^* + iI).$$

Generally, one needs the closure of the range, but since A is closed and symmetric, and

$$\|(A - \lambda I)x\| \geqslant |\beta| \|x\|,$$

the range is already closed.

Let $y \in \mathcal{D}(A^*)$. By the decomposition, we can write

$$(A^* - iI)y = (A - iI)x - 2iu, \tag{25.1}$$

for some $x \in \mathcal{D}(A)$ and $u \in \operatorname{Ker}(A^* + iI)$. This equality uses the fact that any vector in H may be split between $\operatorname{Rn}(A - iI)$ and $\operatorname{Ker}(A^* + iI)$. By definition, $A^*u = -iu$, so

$$(A^* - iI)u = -2iu.$$

Moreover, since

$$A^*\big|_{\mathcal{D}(A)} = A,$$

equality (25.1) can be rewritten in the form

$$(A^* - iI)(y - x - u) = 0.$$

Denote $v := y - x - u$. The equality above means that $v \in \operatorname{Ker}(A^* - iI)$. Therefore, for an arbitrary $y \in \mathcal{D}(Aj)$, we arrive at the decomposition

$$y = x + v + u,$$

where $x \in \mathcal{D}(A)$, $v \in \operatorname{Ker}(A^* - iI)$, and $u \in \operatorname{Ker}(A^* + iI)$. Hence,

$$\mathcal{D}(A^*) \subset \mathcal{D}(A) + \operatorname{Ker}(A^* - iI) + \operatorname{Ker}(A^* + iI).$$

Next, let us prove that this sum is direct. To prove this, it is sufficient to demonstrate that the subspaces mentioned intersect trivially, that is, the zero element admits a unique representation via those subspaces. Suppose

$$y = x' + v' + u',$$

and, subtracting two decompositions, we get

$$0 = (x - x') + (v - v') + (u - u').$$

Suppose that

$$x + v + u = 0,$$

where $x \in \mathcal{D}(A)$, $v \in \operatorname{Ker}(A^* - iI)$, and $u \in \operatorname{Ker}(A^* + iI)$. Apply the operator $A^* - iI$. This operator sends v to zero, while x is in $\mathcal{D}(A)$, so $A^*x = Ax$. Hence,

$$(A^* - iI)(x + u + v) = 0 \quad \Leftrightarrow \quad (A - iI)x + (A^* - iI)u = 0.$$

Here, $u \in \operatorname{Ker}(A^* + iI)$, therefore, $A^*u = -iu$, so

$$(A - iI)x - 2iu = 0. \tag{25.2}$$

From our decomposition

$$H = \operatorname{Rn}(A - iI) \oplus \operatorname{Ker}(A^* + iI),$$

equality (25.2) implies

$$u = 0, \qquad (A - iI)x = 0.$$

But $(A - iI)x = 0$ would place i in the point spectrum if $x \neq 0$, which is impossible; thus, $x = 0$. Altogether, $x = 0$, $u = 0$, and $v = 0$. Therefore, the decomposition of zero is unique, so the sum is direct. □

Description of Self-Adjoint Extensions of a Symmetric Operator in Terms of Extensions of the Corresponding Isometry

We now discuss how to build self-adjoint extensions of a symmetric operator A. The core idea is as follows. Since we know

$$\mathcal{D}(A^*) \supset \mathcal{D}(A),$$

we can impose extra boundary conditions on functions in $\mathcal{D}(A^*)$. This yields an operator B^* that extends B with domain larger than $\mathcal{D}(A)$. By further, more restrictive conditions, we eventually arrive at some $\boldsymbol{C}^*$ whose domain coincides with that of $\boldsymbol{C}$, implying $\boldsymbol{C}$ is self-adjoint.

Recall the example of the operator

$$Af = if'$$

in $L_2[0,1]$ with $\mathcal{D}(A) = \overset{\circ}{W}{}_2^1[0,1]$. We demonstrated earlier that this operator is symmetric but not self-adjoint, since

$$\mathcal{D}(A^*) = W_2^1[0,1].$$

One can then impose additional constraints on functions from $\mathcal{D}(A^*)$ (for instance, boundary conditions at 0 and 1), until the domain of the resulting operator shrinks to something strictly larger than $\mathcal{D}(A)$ but still enabling a self-adjoint operator.

Note that the decomposition

$$\mathcal{D}(A^*) = \mathcal{D}(A) + \mathrm{Ker}(A^* - iI) + \mathrm{Ker}(A^* + iI),$$

and the one-to-one correspondence between symmetric operators and isometries both play an essential role. There is also a correspondence between the deficiency indices of a symmetric operator and those of the related isometry:

$$n_i(V) = \dim(\mathrm{Rn}\, V)^\perp, \quad n_e(V) = \dim\big(\mathcal{D}(V)\big)^\perp.$$

Since $\mathcal{D}(V) = \mathrm{Rn}(A + iI)$,

$$n_-(A) \equiv \dim\big(\mathrm{Rn}(A + iI)\big)^\perp = n_e(V),$$

and, similarly, since $Vy = (A - iI)x$,

$$n_+(A) \equiv \dim\big(\mathrm{Rn}(A - iI)\big)^\perp = n_e(V).$$

Thus, constructing an isometry V_0 that pairs the orthogonal complements of $\mathrm{Rn}\, V$ and $\mathcal{D}(V)$,

$$V_0 : \mathcal{D}(V) \to \mathrm{Rn}\, V,$$

yields a unitary operator $\widetilde{V} = V \oplus V_0$. This extended $\widetilde{V}$ has zero deficiency indices, so the deficiency indices of corresponding symmetric operator $\widetilde{A}$ vanish as well, making $\widetilde{A}$ a self-adjoint operator.

Next, let us consider the idea in more detail. First, write out the decomposition

$$\mathcal{D}(A^*) = \mathcal{D}(A) + \mathrm{Ker}(A^* - iI) + \mathrm{Ker}(A^* + iI).$$

Let

$$\mathcal{D}_0 \subset \operatorname{Ker}(A^* - iI), \quad R_0 \subset \operatorname{Ker}(A^* + iI)$$

such that

$$\dim \mathcal{D}_0 = \dim R_0.$$

Therefore, we can construct an isometry (a unitary operator)

$$V_0 : \mathcal{D}_0 \to R_0.$$

Then,

$$\mathcal{D}(\widetilde{A}) = \mathcal{D}(A) + (V_0 - I)\mathcal{D}_0$$

is the domain of a certain symmetric extension of A, where the second terms is due to $(V - I)y = -2ix$, $y = (A + iI)x$, where $x \in \mathcal{D}(A)$. The action of $\widetilde{A}$ is then:

$$\widetilde{A}(x + v) = Ax - iv - iV_0 v,$$

where $v \in \mathcal{D}_0 \subset \operatorname{Ker}(A^* - iI)$, so $A^* v = iv$, and $V_0 v \in R_0 \subset \operatorname{Ker}(A^* + iI)$, $A^* V_0 v = -iV_0 v$. Hence, one interlaces subspaces somehow the subspaces of $\operatorname{Ker}(A^* \pm iI)$.

Let us return to the example

$$Af = if', \qquad \mathcal{D}(A) = \overset{\circ}{W}{}_2^1[0,1]$$

with the adjoint

$$A^* f = if', \qquad \mathcal{D}(A^*) = W_2^1[0,1].$$

The deficiency indices of A both equal 1, since the equation

$$(A^* \pm iI)f = 0$$

has the solution $f = \exp\{\pm x\}$. Indeed, one finds $e^x \in \mathcal{D}_0$ and $e^{-x} \in R_0$. Let us normalize them: $e^x/\|e^x\|$, $e^{-x}/\|e^{-x}\|$. Then $\mathcal{D}_0$ is the linear span of $e^x/\|e^x\|$, and R_0 is the linear span of $e^{-x}/\|e^{-x}\|$. We can define V_0 as follows:

$$V_0 \frac{e^x}{\|e^x\|} = \gamma \frac{e^{-x}}{\|e^{-x}\|}.$$

To construct a self-adjoint extension $\widetilde{A} \equiv A_\gamma$ of A, one would use the formula for $\widetilde{A}(x + v)$ above, where $v \in \mathcal{D}_0$. This may be somewhat cumbersome to compute directly, and amore conventional approach is to impose boundary conditions.

One can see that deficiency indices both equal 1; therefore, there is a one-parameter family of self-adjoint extensions of A. We must impose some boundary conditions for the functions from $\mathcal{D}(A)$:

$$f(0) = \theta f(1), \qquad |\theta| = 1.$$

Further, let us finish with a straightforward calculation of extension through V_0. Observe that

$$\mathcal{D}_0 = \langle \frac{\sqrt{2}}{\sqrt{e^2-1}} e^x \rangle, \qquad R_0 = \langle \frac{\sqrt{2}e}{\sqrt{e^2-1}} e^{-x} \rangle,$$

and

$$V_0 \frac{\sqrt{2}}{\sqrt{e^2-1}} e^x = \gamma \frac{\sqrt{2}e}{\sqrt{e^2-1}} e^{-x}, \qquad |\gamma| = 1,$$

since any unitary operator between one-dimensional spaces is a multiplication by a number with unit norm. Thus, an additional function from the domain of the extension has the form $(V_0 - I)v$:

$$f(x) = \gamma \frac{\sqrt{2}e}{\sqrt{e^2-1}} e^{-x} - \frac{\sqrt{2}}{\sqrt{e^2-1}} e^x.$$

For this function, we have

$$f(0) = \frac{\gamma\sqrt{2}e - \sqrt{2}}{\sqrt{e^2-1}}$$

and

$$f(1) = \frac{\gamma\sqrt{2} - \sqrt{2}e}{\sqrt{e^2-1}}.$$

From these expressions, one can solve for θ in terms of γ.

Lecture 26. Functional Calculus

Spectral Theorem: Introduction

In this lecture, we discuss the so-called *spectral theorem* for self-adjoint operators. For bounded operators, we will prove all relevant statements in detail, while for unbounded operators, we will only outline and formulate the main results without fully proving them.

In short, the spectral theorem asserts that *every self-adjoint operator can be viewed as a multiplication operator*. The crucial questions, however, are: *by what function do we multiply, and in which space does this multiplication take place*?

To build intuition, let us consider some illustrative examples:

1) In finite-dimensional settings, every self-adjoint matrix can be brought to a diagonal form by choosing an appropriate orthonormal basis. In this basis, the operator simply multiplies each basis vector by a corresponding real eigenvalue.

2) By the Hilbert–Schmidt theorem, if A is a compact self-adjoint operator on an infinite-dimensional Hilbert space H, then A admits an orthonormal basis $\{e_k\}_{k=1}^{\infty}$ of eigenvectors. Specifically, for $A = A^*$ and $A \in C(H)$, there exists an orthonormal basis such that
$$Ae_k = \lambda_k e_k,$$
with each $\lambda_k \in \mathbb{R}$. Any $x \in H$ can be expanded in that basis,
$$x = \sum_{k=1}^{\infty} (x, e_k) e_k$$
and thus
$$Ax = \sum_{k=1}^{\infty} \lambda_k (x, e_k) e_k.$$
If we denote $x_k := (x, e_k)$, then A is unitary equivalent (i.e., similar) to an operator B: $\ell_2 \to \ell_2$ given by
$$Bx = (\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_n x_n, \ldots),$$
which clearly operates by multiplying each coordinate x_k by the real number λ_k. Hence, in the compact case, A is essentially "multiplication by constants" in the basis of its eigenvectors.

3) General Self-Adjoint Operators. In the most general setting (possibly unbounded, not necessarily compact), one can still show that any self-adjoint operator is *similar*

(via a suitable unitary transformation) to an operator that multiplies by a real-valued function on some space. Concretely, one may think of the example

$$Af = xf$$

which acts by multiplication by the independent variable x.

Spectral Mapping Theorem

Theorem 26.1. *Let X be a Banach space, $A \in B(X)$, and*

$$p(z) = \sum_{k=0}^{n} a_k z^k.$$

Then

$$\sigma\big(p(A)\big) = p\big(\sigma(A)\big) = \{\lambda = p(\mu),\ \mu \in \sigma(A)\}.$$

Proof. By the fundamental theorem of algebra, any nonconstant polynomial of degree n can be factored (counted with multiplicity) into n linear factors:

$$p(z) - \lambda = a_n(z - \lambda_1)(z - \lambda_2)\dots(z - \lambda_n). \tag{26.1}$$

Suppose $\lambda \in \sigma\big(p(A)\big)$. This means that

$$p(A) - \lambda I = a_n(A - \lambda_1 I)\dots(A - \lambda_n I)$$

is not invertible. Consequently, at least one factor on the right-hand side must fail to be invertible (otherwise, a composition of invertible operators would itself be invertible). Hence there is some index i such that $A - \lambda_i I$ is not invertible, i.e., $\lambda_i \in \sigma(A)$. Next, $p(\lambda_i) - \lambda = 0$, see (26.1); that is,

$$\lambda = p(\lambda_i),$$

which proves that $\sigma\big(p(A)\big) \subset p\big(\sigma(A)\big)$.

For the reverse inclusion, suppose $\lambda \in p\big(\sigma(A)\big)$. Then, by definition,

$$\lambda = p(\lambda_i), \qquad \lambda_i \in \sigma(A).$$

Thus, for some i, the operator $A - \lambda_i I$ is not invertible.

Note that sometimes the composition of noninvertible operators can sometimes be invertible (consider, for instance, $A_\ell A_r$; both have no inverse, while the composition is I). However, we will see that in our specific factorization, at least one factor must remain noninvertible in a way that forces $p(A) - \lambda I$ to be noninvertible.

If $\lambda_i \in \sigma_p(A)$ then we can write

$$p(A) - \lambda I = a_n \prod_{j \neq i} (A - \lambda_j I) \cdot (A - \lambda_i I)$$

(note that the factors $A - \lambda_k I$ so we are free to reorder them as needed). Thus, there exists $x \neq 0$: $(A - \lambda_i I)x = 0$. Therefore,

$$\big(p(A) - \lambda I\big)x = a_n \prod_{j \neq i} (A - \lambda_j I) \cdot (A - \lambda_i I)x = 0.$$

Thus, $\lambda \in \sigma_p\big(p(A)\big)$.

Further, if

$$\lambda_i \in \sigma_c(A) \cup \sigma_r(A),$$

then A is injective but its range in not the entire space. Rewrite

$$p(A) - \lambda I = (A - \lambda_i I) \cdot a_n \prod_{j \neq i} (A - \lambda_j I).$$

The composition of operators with $j \neq i$ acts somehow; the range of the last operator, $A - \lambda_i I$, is not the entire space. Then, the range of the complete composition is not the entire space as well. Hence

$$\lambda \in \sigma_c\big(p(A)\big) \cup \sigma_r\big(p(A)\big). \qquad \square$$

Let us consider some applications of this theorem. Let $A = A^*$, and $\sigma(A) = \{0, 1\}$. One can show that such an operator must be a projection.

Indeed, consider the polynomial $p(z) = z^2 - z$. By the spectral mapping theorem, we have

$$\sigma\big(p(A)\big) = p\big(\sigma(A)\big) = \{0\}.$$

Further, observe that $p(A) = A^2 - A$; the coefficients of p are real, so this operator is self-adjoint along with A. For a self-adjoint operator, the spectral radius is equal to the norm; one can see that the spectral radius of $p(A)$ is zero: $r(p(A)) = 0$. Therefore, $\|p(A)\| = 0$. Thus, A satisfies

$$A^2 = A,$$

which is the algebraic definition of a projection.

Next, for a given polynomial

$$p(z) = \sum_{k=0}^{n} a_k z^k,$$

one can define a polynomial of $A \in B(X)$ by

$$p(A) = \sum_{k=0}^{n} a_k A^k, \qquad A^0 = I.$$

However, certain applications require more general functions f, for instance analytic functions on a region containing $\sigma(A)$ or continuous functions if A is self-adjoint in a Hilbert space.

For a *compact* self-adjoint operator, one can define $f(A)$ for $f \in C\big(\sigma(A)\big)$. Due to the Hilbert–Schmidt theorem, there exists an orthonormal basis $\{e_k\}_{k=1}^{\infty}$ such that

$$Ae_k = \lambda_k e_k.$$

Expanding x, we obtain

$$x = \sum_{k=1}^{\infty} (x, e_k) e_k, \qquad Ax = \sum_{k=1}^{\infty} \lambda_k (x, e_k) e_k.$$

Then, if $f \in C\big(\{\lambda_k\}\big)$, one can define

$$f(A) = \sum_{k=1}^{\infty} f(\lambda_k)(x, e_k) e_k.$$

Continuous Functional Calculus

Theorem 26.2 (Continuous Functional Calculus)**.** *Let $A = A^* \in B(H)$, where H is a Hilbert space, and $\sigma(A) \subset [a,b]$. Then $\exists!$ homomorphism $\varphi : C[a,b] \to B[H]$ such that*

1) φ*:* $f \equiv 1 \mapsto I$,

2) φ*:* $id(x) = x \mapsto A$,

3) $\varphi(\alpha f + \beta g) = \alpha\varphi(f) + \beta\varphi(g)$, $\varphi(fg) = \varphi(f)\varphi(g)$,

4) $f_n \underset{[a,b]}{\rightrightarrows} f \Rightarrow \varphi(f_n) \overset{\|\cdot\|_{B(H)}}{\to} \varphi(f)$, *and*

$$\|\varphi(f)\| \leqslant \max_{[a,b]} |f(x)|,$$

5) $\varphi(\overline{f}) = \big(\varphi(f)\big)^*$,

6) $BA = AB \Rightarrow \varphi(f)B = B\varphi(f)$ $\forall f \in C[a,b]$.

Note that, using this theorem, one can define $\varphi(f) =: f(A)$. The only condition is that the function must be continuous on the spectrum.

Proof. Items 1, 2, and 3 define a homomorphism on polynomials: for

$$p(z) = \sum_{k=0}^{n} a_k z^k$$

we set

$$\varphi(p) = \sum_{k=0}^{n} a_k A^k.$$

For general $f \in C[a,b]$, due to the Weierstrass approximation theorem,

$$\exists p_n : \ p_n \rightrightarrows f.$$

Consider $p_n(A) - p_m(A)$. This operator is normal (in case of real coefficients of p_n, p_m, it is self-adjoint). Note that the spectral radius of a normal operator is equal to the norm (similar to the self-adjoint case). Therefore,

$$\|p_n(A) - p_m(A)\| = r\big((p_n - p_m)(A)\big) = \max_{\lambda \in \sigma'} |\lambda|,$$

where σ' stands for $\sigma\big((p_n - p_m)(A)\big)$. Thus, due to the spectral mapping theorem, we have

$$\max_{\lambda \in \sigma'} |\lambda| = \max_{\lambda \in \sigma(A)} |(p_n - p_m)(\lambda)|,$$

and, since $\sigma(A) \subset [a,b]$,

$$\max_{\lambda \in \sigma(A)} |(p_n - p_m)(\lambda)| \leqslant \max_{[a,b]} |p_n(\lambda) - p_m(\lambda)| \to 0,$$

since $p_n \rightrightarrows f$. Thus, $\{p_n(A)\}_{n=1}^{\infty}$ is a Cauchy sequence, so

$$\exists \lim_{k\to\infty} p_n(A) =: f(A).$$

Now, let us prove the uniqueness of φ. The sequence $\{p_n\}$ is not unique; let $q_n \rightrightarrows f$. Consider the sequence

$$p_1, q_1, p_2, q_2, \ldots, p_n, q_n \rightrightarrows f.$$

Thus, the sequence

$$p_1(A), q_1(A), p_2(A), q_2(A), \ldots, p_n(A), q_n(A), \ldots$$

is Cauchy, so the limit is the same:

$$\lim_{k\to\infty} p_n(A) = \lim_{k\to\infty} q_n(A).$$

Properties regarding the convergence are simple to prove; the proof is similar to reasoning above.

Next, consider

$$\overline{p} = \sum_{k=0}^{n} \overline{a}_k z^k,$$

where $z \in \mathbb{R}$ since the spectrum of $A = A^*$ is real. Then

$$\overline{p}(A) = \sum_{k=0}^{n} \overline{a}_k A^k = \Big(\sum_{k=0}^{n} a_k A^k\Big)^* = \big(p(A)\big)^*,$$

and, through the limit procedure, the same can be obtained for an arbitrary continuous function:

$$\varphi(\overline{f}) = \big(\varphi(f)\big)^*.$$

Further, suppose $BA = AB$; then

$$p(A)B = Bp(A),$$

and, through the limit procedure, the same can be concluded for an arbitrary continuous function.

Hence we obtain a unique, well-defined homomorphism φ that satisfies the above conditions. □

Borel Functional Calculus

For some applications, we need a class of functions even broader than the continuous functions. To introduce it, we begin with some preliminary definitions.

Recall that a function $f\colon \mathbb{R} \to \mathbb{R}$ is called *measurable* if $\forall c \in \mathbb{R}$

$$f^{-1}\big((c, +\infty)\big) = \{x : \ f(x) > c\}$$

is a measurable set.

Definition 26.1. *$\mathscr{B}[a,b]$, called a **Borel set** on $[a,b]$, is the minimal σ-algebra containing all open subsets of $[a,b]$.*

Definition 26.2. *A function f is called a **Borel function** if $\forall A \in \mathscr{B}[a,b]$*

$$f^{-1}(A) \in \mathscr{B}[a,b],$$

where $f^{-1}(A) = \{x : \ f(x) \in A\}$. We denote by $\mathcal{B}or[a,b]$ the set of all Borel functions on $[a,b]$.

Theorem 26.3 (Bounded Borel Functional Calculus)**.** *Let $A = A^* \in B(H)$, where H is a Hilbert space, and $\sigma(A) \subset [a,b]$. Then $\exists!$ homomorphism $\varphi\colon \mathcal{B}or[a,b] \to B(H)$ such that*

1) $\varphi\colon f \equiv 1 \mapsto I$,

2) $\varphi\colon id(x) = x \mapsto A$,

3) $\varphi(\alpha f + \beta g) = \alpha\varphi(f) + \beta\varphi(g)$, $\varphi(fg) = \varphi(f)\varphi(g)$,

4) $f_n(x) \to f(x)$ $\forall x \in [a,b]$ $\Rightarrow$ $\varphi(f_n) \xrightarrow{S} \varphi(f)$ *and*

$$\|\varphi(f)\| \leqslant \sup_{[a,b]}|f(x)|,$$

5) $\varphi(\overline{f}) = \big(\varphi(f)\big)^*$,

6) $BA = AB \Rightarrow \varphi(f)B = B\varphi(f)$ $\forall f \in \mathcal{B}or[a,b]$.

Idea of the proof. Let $f \in C[a,b]$, and take two fixed vectors $x, y \in H$. Consider the inner product

$$\big(f(A)x,y\big).$$

This is a continuous functional with respect to f: due to the Cauchy–Bunyakovsky inequality,

$$\big|\big(f(A)x,y\big)\big| \leqslant \|f(A)\| \cdot \|x\| \cdot \|y\| \leqslant \max_{[a,b]}|f(t)| \cdot \|x\| \cdot \|y\|.$$

Every continuous functional on $C[a,b]$ is a function from $BV_0[a,b]$. Hence there exists $g \in BV_0[a,b]$ such that

$$\big(f(A)x,y\big) = \int_a^b f(t)\,dg_{x,y}(t).$$

Then, one approximates a given bounded Borel function by a sequence of continuous functions: for each $f \in \mathcal{B}or[a,b]$, there exists a sequence $\{f_n\} \subset C[a,b]$:

$$f_n(t) \to f(t) \quad \forall t.$$

By Lebesgue's theorem,

$$\int_a^b f_n\,dg \to f\,dg,$$

where

$$\int_a^b f_n\,dg = \big(f_n(A)x,y\big) \to (Bx,y)$$

for some (linear) operator B. Further, $\varphi(f) =: B$ is defined as a weak limit

$$B = \lim_{n\to\infty} f_n(A),$$

i.e., formally, one obtains $f_n(A) \rightharpoonup B$. in the weak operator topology. However, since A is self-adjoint, this convergence is actually strong. To see this, consider $\overline{f}_n(t)$. Then

$$\overline{f}_n(t) f_n(t) \to |f(t)|^2,$$

which implies

$$\big(f_n(A)\big)^* f_n(A) \rightharpoonup |f(A)|^2.$$

Hence

$$\Big(\big(f_n(A)\big)^* f_n(A)x, x\Big) \to \Big(|f(A)|^2 x, x\Big).$$

By definition, we can take the adjoint to the second factor, obtaining

$$\|f_n(A)x\|^2 \to \|f(A)x\|^2,$$

which is precisely strong convergence

$$\big(f_n(A) - f(A)\big)x \to 0.$$

The rest of the properties can be verified similarly to the continuous functional calculus case. □

Example: Projection-Valued Measure

Why do we need the Borel functional calculus? One reason is to handle characteristic (indicator) functions of Borel sets. Let Ω be a Borel set, and define

$$\chi_\Omega(t) = \begin{cases} 1, & t \in \Omega, \\ 0, & t \notin \Omega. \end{cases}$$

Since it is a bounded Borel function, one can define $\chi_\Omega(A) =: E_\Omega$ by virtue of the homomorphism φ. When Ω is a disjoint union, say,

$$\Omega = \Omega_1 \sqcup \Omega_2,$$

we have

$$\chi_\Omega(t) = \chi_{\Omega_1}(t) + \chi_{\Omega_2}(t),$$

and thus, due to the properties of φ,

$$E_\Omega = E_{\Omega_1} + E_{\Omega_2}.$$

Hence $E_\Omega = \chi_\Omega(A)$ is an operator-valued measure. Moreover, since

$$\chi_\Omega^2(t) = \chi_\Omega(t),$$

we have the same equality for operators:

$$E_\Omega^2 = E_\Omega,$$

so it is a projection. Next, since $\chi_\Omega(t)$ is a real-valued function,

$$E_\Omega^* = E_\Omega,$$

so this projection is orthogonal. Altogether, E_Ω is a projection-valued measure.

Given such a measure, one can define integrals (Lebesgue, Stieltjes, etc.). For instance, if we look at $(-\infty, \lambda)$, which is a Borel set, we can set

$$E_\lambda := E_{(-\infty,\lambda)}.$$

Then an operator A can be expressed as

$$A = \int_{\sigma(A)} \lambda \, dE_\lambda,$$

so we have a representation for a Borel function of a self-adjoint operator

$$f(A) = \int_{\sigma(A)} f(\lambda) \, dE_\lambda.$$

If $f \in C[a,b]$ then with

$$\sum_i f(\xi_i)\big(E_{\lambda_i} - E_{\lambda_{i-1}}\big)$$

one defines the Riemann integral. (The set $[a,b]$ contains the spectrum of A; consider partitions of this interval with $\lambda_0 < \dots < \lambda_{i-1} < \lambda_i < \dots$, take $\xi_i \in [\lambda_{i-1}, \lambda_i]$, and take a limit.) Since E_{λ_i} are orthogonal projections, the integral converges. For Borel functions, in turn, the construction is analogous to the Lebesgue integral. Let $f \in \mathcal{B}or[a,b]$. Then

$$\sum_k f\Big(\frac{k}{2^n}\Big) E_{\left\{\frac{k}{2^n} \leqslant \lambda < \frac{k+1}{2^n}\right\}}.$$

Projection-Valued Measure for a Finite-Dimensional Operator and a Multiplication Operator

Consider a finite-dimensional example:

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & \dots & 0 \\ 0 & \dots & \ddots & 0 & \dots & 0 \\ \vdots & \dots & \dots & \ddots & \dots & \vdots \\ 0 & 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix} \quad \text{in} \quad \mathbb{C}^n,$$

with $\lambda_1 < \lambda_2 < \cdots < \lambda_n$, $\lambda_j \in \mathbb{R}$. What is E_λ in this situation?

We know that

$$A = \int_{\sigma(A)} \lambda \, dE_\lambda,$$

and

$$Ax = \sum_{k=1}^{n} \lambda_k (x, e_k) e_k,$$

where $\{e_k\}$ is an orthonormal basis. In linear algebra terms, the orthogonal projection of a vector x to the line ℓ is $(x, e_1)e_1$, where e_1 is a unit vector along ℓ.

Then E_λ can be visualized as follows. For $\lambda < \lambda_1$, the projection vanishes. Next, for $\lambda \in [\lambda_1, \lambda_2)$, E_λ is the projection P_1 onto the first eigenvector, e_1:

$$P_1 x = (x, e_1) e_1,$$

and so on, see Fig. 26.1.

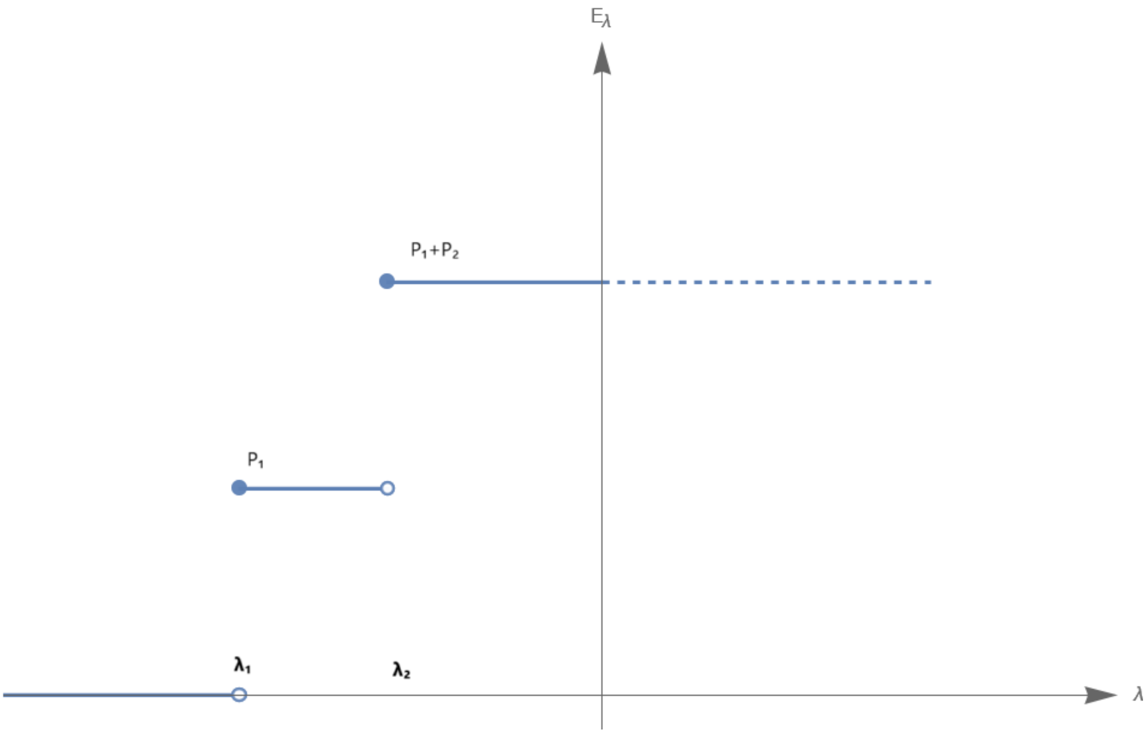

Рис. 26.1. "Graph" of E_λ.

Here, the description of E_λ is quite simple since the spectrum is purely discrete; E_λ just picks out the subspace(s) associated with eigenvalues below λ.

Consider another example. Let $\varphi \in C[a,b]$ be a real-valued function. Let

$$A_\varphi f = \varphi(x) f(x)$$

in $L_2[a,b]$. Suppose additionally that φ is strictly increasing. Then

$$\sigma(A_\varphi) = [\varphi(a), \varphi(b)],$$

and

$$E_\lambda = \chi_{(-\infty,\lambda)}(A) = \begin{cases} 0, & t > \lambda, \\ I, & t < \lambda. \end{cases}$$

The preimage of E_λ under the homomorphism, χ_λ, is of the form

$$\chi_\lambda(t) = \begin{cases} 0, & t > \lambda, \\ 1, & t < \lambda. \end{cases}$$

This is a projection on some part of $[a,b]$:

$$E_\lambda f(t) = \chi_{[a,\varphi^{-1}(\lambda)]}(t) f(t),$$

see Fig. 26.2.

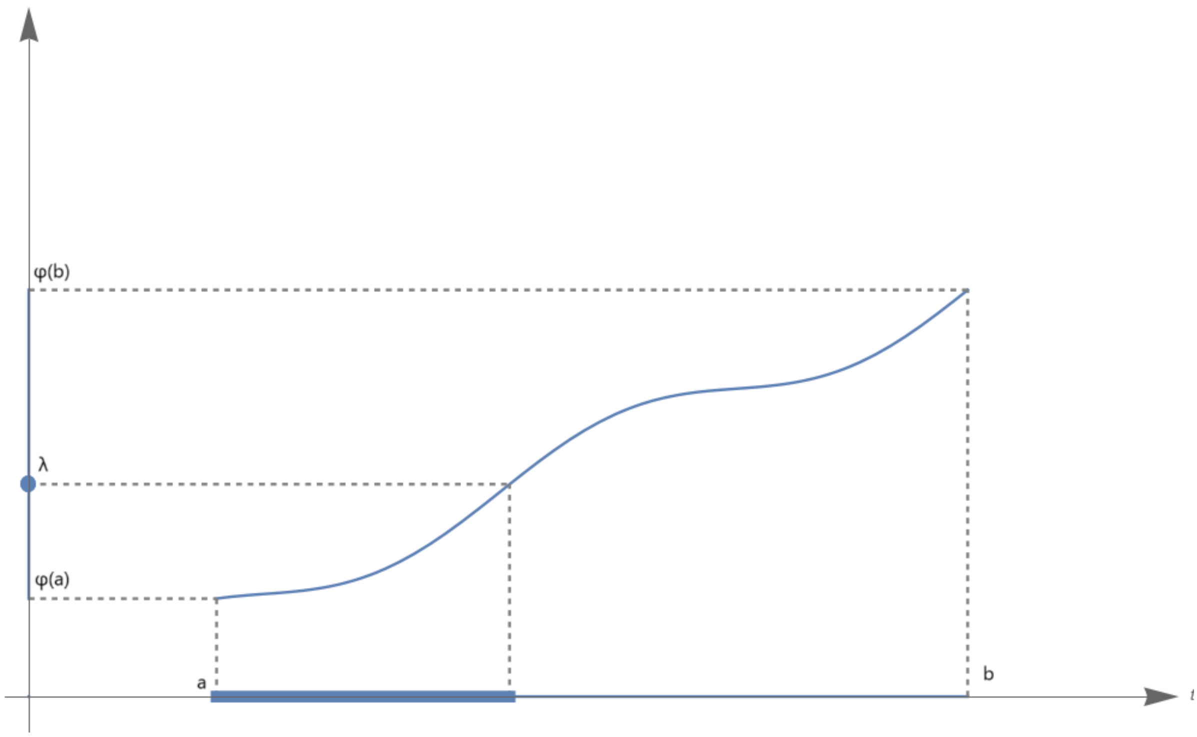

Рис. 26.2. E_λ is a projection onto the “thick” interval $[a, \varphi^{-1}(\lambda)]$.

In this example, $[\varphi(a), \varphi(b)]$ belongs to the continuous spectrum.

Self-Study Exercises

1) Let A be a normal operator. Prove the following equality for the spectral radius:

$$r(A) = \|A\|.$$

2) Consider

$$Af = x^2 f(x)$$

in $L_2[-1,1]$. Describe E_λ.

Lecture 27. Spectral Theorem for Self-Adjoint Operators. Fourier Transform in L_1

Cyclic Vectors

We now delve deeper into the spectral theorem for self-adjoint operators, which asserts that every self-adjoint operator can be represented as a multiplication operator. Before fully stating and proving the theorem, let us introduce some necessary preliminary concepts that will help clarify the reasoning.

Definition 27.1. *Let X be a Banach space, $A \in \mathcal{L}(X)$. A vector x_0 is called* ***cyclic*** *if*

$$\overline{\langle x_0, Ax_0, A^2x_0, \dots \rangle} = X.$$

Example 27.1. *Consider the matrix*

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & \dots & 0 \\ 0 & \dots & \ddots & 0 & \dots & 0 \\ \vdots & \dots & \dots & \ddots & \dots & \vdots \\ \vdots & \dots & \dots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix} \quad in \quad \mathbb{C}^n.$$

For A, there exists a cyclic vector iff $\lambda_i \neq \lambda_j$, $i \neq j$.

Let us show it. First, suppose $\lambda_i \neq \lambda_j$, $i \neq j$, and consider

$$x_0 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Then one can see that the vectors $\{A^k x_0\}_{k=0}^{n-1}$ form a Vandermonde matrix

$$\left(x_0, Ax_0, \dots, A^{n-1}x_0\right) = \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \dots & \lambda_2^{n-1} \\ 1 & \lambda_3 & \lambda_3^2 & 0 & \dots & \lambda_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \dots & \vdots \\ \vdots & \dots & \dots & \dots & \ddots & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \dots & \dots & \lambda_n^{n-1} \end{pmatrix},$$

whose determinant is nonzero; thus, these vectors form a basis, and x_0 is cyclic.

To prove the converse, suppose, without loss of generality, $\lambda_1 = \lambda_2$. Further, consider

$$p(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\dots(\lambda - \lambda_n),$$

where the factor with λ_2 is missed intentionally. Thus,

$$p(A) = 0,$$

where $p(\lambda) = \lambda^{n-1} + c_1\lambda^{n-2} + \dots$, and

$$p(A) = 0 \quad \Leftrightarrow \quad A^{n-1} = -c_1 A^{n-2} - \dots,$$

so there is a vanishing linear combination of powers of A, which means that $\{A^k x_0\}$ cannot be a basis.

Example 27.2. *In $L_2[-1,1]$, consider the operator*

$$(Af)(x) = x^2 f(x).$$

For this operator, no cyclic vectors exist.

To see why, consider two cases.

1) *Space is real. Suppose that f is a cyclic vector of A. Take*

$$h(x) = f(-x)\,\mathrm{sgn}\,x.$$

We will prove that

$$h \perp A^k f \quad \forall k = 0, 1, \dots.$$

Let us take the inner product:

$$\big(A^k f, h\big) = \int_{-1}^{1} x^{2k} f(x) f(-x)\,\mathrm{sgn}\,x\,dx,$$

One can see that the function $x^{2k} f(x) f(-x)$ is even, while $\mathrm{sgn}\,x$ is odd. Therefore, $\big(A^k f, h\big) = 0$.

2) *Space is complex. Here,*

$$\big(A^k f, h\big) = \int_{-1}^{1} x^2 f(x)\overline{h(x)}\,dx.$$

Take

$$h(x) = f(x)|f(-x)|^2\,\mathrm{sgn}\,x.$$

Then,

$$\int_{-1}^{1} x^2 f(x)\overline{h(x)}\,dx = \int_{-1}^{1} x^2 |f(x)|^2 |f(-x)|^2 \operatorname{sgn} x\,dx,$$

so, again, we have the even function $x^2|f(x)|^2|f(-x)|^2$ *times the odd function* $\operatorname{sgn} x$, *thus,* $\big(A^k f, h\big) = 0$.

Example 27.3. *Consider the same operator in* $L_2[0,1]$*:*

$$(Af)(x) = x^2 f(x).$$

Here, the constant function $f(x) \equiv 1$ *is a cyclic vector.*

One can see that $\{A^k f\}_{k=0}^{\infty} = \{x^{2k}\}_{k=0}^{\infty}$. *We know that in* $L_2[-1,1]$, *the system*

$$\sin \pi n x, \quad \cos \pi n x, \quad 1, \quad n = 1, 2, \ldots,$$

is a basis. Taking only even functions, we obtain a basis in $L_2[0,1]$. *Then, due to the fact that the Taylor expansion of cosine consists of even powers of* x, *we get that the functions* x^{2k} *form a basis. Another approach to show this is to use the Müntz theorem that claims that* $\{x^{n_k}\}_{k=0}^{\infty}$, $n_1 < n_2 < \ldots$, *is dense in* $C[0,1]$ *if* $n_1 = 0$ *and*

$$\sum_{k=1}^{\infty} \frac{1}{n_k} = \infty.$$

For $L_2[0,1]$, *one can drop the condition* $n_1 = 0$.

Spectral Theorem for Bounded Self-Adjoint Operators with Simple Spectrum

Definition 27.2. *An operator* A *is said to have a* ***simple spectrum*** *if* A *admits a cyclic vector.*

In this case, every eigenvalue of A has multiplicity one.

Theorem 27.1. *Let* H *be a separable Hilbert space, and* $A = A^* \in B(H)$. *Let* A *have a cyclic vector. Then there exist a measure* μ *on* $\sigma(A)$ *and a unitary isomorphism*

$$U : H \to L_2\big(\sigma(A), \mu\big)$$

such that the following diagram commutes:

$$\begin{array}{ccc} H & \xrightarrow{\quad A \quad} & H \\ {\scriptstyle U}\downarrow & & \downarrow{\scriptstyle U} \\ L_2\big(\sigma(A),\mu\big) & \xrightarrow{\ M\ } & L_2\big(\sigma(A),\mu\big), \end{array}$$

i.e., $UA = MU$, *where* M *is a multiplication operator*

$$Mf = t \cdot f(t).$$

The examples can be provided by a finite-dimensional self-adjoint operator (which is a multiplication operator in an appropriate basis) or by a compact self-adjoint operator (which is in fact a multiplication operator in ℓ_2). However, this theorem allows one to describe an operator with continuous spectrum.

Proof. Suppose that x_0 is a cyclic vector. For A, we have the spectral measure E_λ. Define the measure μ by

$$\mu\big([\alpha,\beta)\big) = \big(E_{[\alpha,\beta)}x_0, x_0\big)_H, \qquad E_{[\alpha,\beta)} = E_\beta - E_\alpha.$$

Further, define an isomorphism $U : H \to L_2\big(\sigma(A),\mu\big)$ by

$$U : x_0 \mapsto 1, \quad Ax_0 \mapsto t,$$

where t is an independent variable, and so on,

$$A^k x_0 \mapsto t^k.$$

We know that

$$\overline{\langle \{A^k x_0\}\rangle} = H.$$

As A is self-adjoint, the spectrum lies on some interval of real axis; the polynomial system $\langle \{t^k\}_{k=0}^{\infty}\rangle$ is dense in $L_2\big(\sigma(A),\mu\big)$.

Next, let us verify that this isomorphism is unitary. Since $\langle \{A^k x_0\}\rangle$ is dense in H, it is sufficient to establish the unitarity for this set. Consider

$$(A^k x_0, A^n x_0)_H = (A^{k+n}x_0, x_0)_H,$$

where one can substitute the integral representation for A:

$$(A^{k+n}x_0, x_0)_H = \int_{\sigma(A)} \lambda^{k+n}\,(dE_\lambda x_0, x_0); \tag{27.1}$$

recall here that for $f(A)$, we have

$$f(A) = \int_{\sigma(A)} f(\lambda)\,dE_\lambda.$$

In (27.1), we integrate with respect to the measure μ, so

$$\int_{\sigma(A)} \lambda^{k+n}\,(dE_\lambda x_0, x_0) = \int_{\sigma(A)} \lambda^{k+n}\,d\mu = (\lambda^k, \lambda^n)_{L_2\left(\sigma(A),\mu\right)},$$

thus, U is unitary. $\square$

Let us consider some examples.

Example 27.4. *In* $L_2[0,1]$*, consider*

$$(Af)(x) = x^2 \cdot f(x).$$

Let us find the measure μ *such that* $A \sim M$*:*

$$Mh(t) = t \cdot h(t)$$

in $L_2\big(\sigma(A), \mu\big)$*.*

First, we must describe the spectral measure E_λ*. On* $[0,1]$*, the values of* x^2 *are* $[0,1]$*; this is the spectrum of* A*:* $\sigma(A) = [0,1]$*. Take* $\lambda \in [0,1]$*. Then*

$$E_\lambda = \begin{cases} 0, \quad \lambda < 0, \\ \chi_{[0,\sqrt{\lambda})}(x) f(x), \quad \lambda \in [0,1), \\ f(x), \quad \lambda \geqslant 1. \end{cases}$$

For A*,* $f_0 \equiv 1$ *is a cyclic vector.*

For $[\alpha, \beta) \subset \sigma(A)$*, the measure* μ *can be defined by*

$$\mu\big([\alpha,\beta)\big) = \int_{\sigma(A)} \chi_{[\sqrt{\alpha},\sqrt{\beta})} 1\, d\mu = \sqrt{\beta} - \sqrt{\alpha}.$$

Thus, A *is unitary equivalent to the multiplication by an independent variable in* $L_2\big(\sigma(A), \mu\big)$*.*

Spectral Theorem for Bounded Self-Adjoint Operators

We now generalize the spectral theorem to the case where the spectrum of an operator is not necessarily simple.

Theorem 27.2 (w/o proof)**.** *Let* $A = A^* \in B(H)$*, where* H *is a separable Hilbert space, and* A *has no cyclic vectors. Then there exists a decomposition of* H

$$H = \oplus_{k=1}^{N} H_k, \quad N \leqslant \infty,$$

such that $A\big|_{H_k}$ *has a cyclic vector for each* H_k*.*

The idea of the proof is to use Zorn's lemma. Let us formulate the next theorem.

Theorem 27.3. *Let* H *be a separable Hilbert space and* $A = A^* \in B(H)$*. Then there exist measures* $\{\mu_k\}_{k=1}^{n}$ *and a unitary isomorphism*

$$U : \oplus_{k=1}^{N} H_k \to \oplus_{k=1}^{N} L_2\big(\sigma(A), \mu_k\big)$$

such that the following diagram commutes:

$$\begin{array}{ccc} \oplus_{k=1}^{N} H_k & \xrightarrow{\quad A \quad} & \oplus_{k=1}^{N} H_k \\ {\scriptstyle U}\downarrow & & \downarrow{\scriptstyle U} \\ \oplus_{k=1}^{N} L_2\big(\sigma(A),\mu_k\big) & \xrightarrow{\ M\ } & \oplus_{k=1}^{N} L_2\big(\sigma(A),\mu_k\big), \end{array}$$

i.e., $UA = MU$, where the operator M is defined by

$$M\Big|_{L_2\left(\sigma(A),\mu_k\right)} f = t f(t) \quad \forall f \in L_2\big(\sigma(A),\mu_k\big).$$

The proof is quite similar to given above for an operator with a cyclic vector; one must take the spectral measures of $A\big|_{H_k}$ and construct an isomorphism U for each component of the sum.

Let us consider some examples.

Example 27.5. *Consider again*

$$Af = x^2 \cdot f(x)$$

in $L_2[-1,1]$. Being defined in this space, A has no cyclic vectors. However, one can decompose $H = H_1 \oplus H_2$ with

$$H_1 = \{f \in L_2[-1,1] : \ f(-x) = f(x)\}, \qquad H_2 = \{f \in L_2[-1,1] : \ f(-x) = -f(x)\},$$

that is, H_1 is the space of even functions, and H_2 is the space of odd functions. In H_1, take $f_1 \equiv 1$ as a cyclic vector; for H_2, take $f_2(x) = x$.

One can observe that $\sigma(A) = [0,1]$, and each point of spectrum has 2 preimages: $\pm\sqrt{\lambda}$. Consequently, the spectral measure E_λ acts as follows:

$$E_\lambda f = \begin{cases} 0, & \lambda < 0, \\ \chi_{[-\sqrt{\lambda},\sqrt{\lambda}]}(x) f(x), & \lambda \in [0,1), \\ f(x), & \lambda \geqslant 1. \end{cases}$$

Next, we define the measures μ_j corresponding to H_j. One can see that

$$\mu_1([\alpha,\beta)) = 2(\sqrt{\beta} - \sqrt{\alpha}),$$

and, since in H_2 we integrate with the square of $f_2(x) = x$, we see that

$$\mu_2([\alpha,\beta)) = \frac{2}{3}\Big(\beta\sqrt{\beta} - \alpha\sqrt{\alpha}\Big).$$

A natural question arises: how can we determine whether two self-adjoint operators are similar? Let $A_1 = A_1^*$, $A_2 = A_2^*$ in H. Then, $A_1 \sim A_2$ if and only if the following conditions hold:

1) Spectra coincide: $\sigma(A_1) = \sigma(A_2)$.

2) Measures are equivalent: $\mu_1 \sim \mu_2$, which means that for any measurable set Ω,

$$\mu_1(\Omega) = 0 \quad \Leftrightarrow \quad \mu_2(\Omega) = 0.$$

3) Multiplicity functions match: $n_{A_1} = n_{A_2}$. Let the spectrum decompose as follows:

$$\sigma(A) \supset \sigma_1(A) \supset \dots \supset \sigma_j(A) \supset \sigma_{j+1}(A),$$

and define the multiplicity function n_A by

$$n_A(\lambda) = \begin{cases} 0, & \lambda \notin \sigma(A), \\ j, & \lambda \in \sigma_j \backslash \sigma_{j+1} \end{cases}$$

Example 27.6. *Consider the operator*

$$Af = \varphi(x) f(x)$$

in $L_2[-1,1]$ *with* $\varphi(x)$ *depicted in Fig. 27.1.*

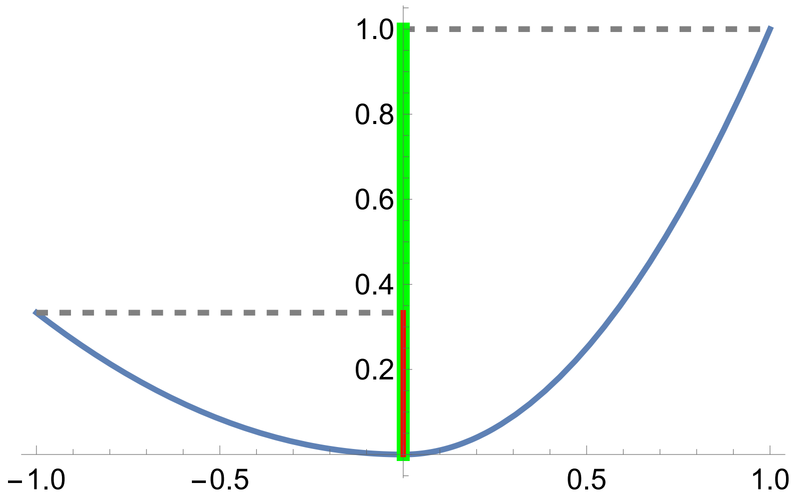

Рис. 27.1. Graph of $\phi(x)$.

One can see that the $\sigma_1(A) = \sigma(A) = [0,1]$ *and* $\sigma_2(A) = [0, \varphi(-1)]$ *(green and red lines respectively in Fig. 27.1).*

Note that conditions 1–3 listed above are sometimes referred to as the *unitary invariants* of a self-adjoint operator. If these invariants coincide for two self-adjoint operators A_1 and A_2, then A_1 and A_2 are similar (unitarily equivalent).

Fourier Transform

Now, we turn to the Fourier transform, a crucial tool for analyzing operators and functions in Hilbert spaces. First, we will define the transform itself by a formula, and then discuss for which classes of functions it is well-defined.

Define the (direct) Fourier transform by

$$\big(\widehat{F} f\big)(y) \equiv \widehat{f}(y) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ixy}\, dx$$

and the inverse Fourier transform by

$$\big(\check{F} f\big)(y) \equiv \check{f}(y) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{ixy}\, dx.$$

It is evident that this transform is well-defined for functions in $L_1(\mathbb{R})$. In further lectures, we will also consider it in $L_2(\mathbb{R})$.

Also, the Fourier transform can be considered in $\mathcal{S}(\mathbb{R})$ (the Schwartz space); however, this class of functions is quite narrow, and applications in this case are limited.

To proceed, we need to use a key result from measure theory: Lebesgue's dominated convergence theorem, stated below in a convenient form for the reasoning that follows.

Theorem 27.4. *For $f(x,y)$, where $y \in \Omega$, suppose*

1) *$\forall y \in \Omega$: $|f(x,y)| \leqslant g(x) \in L_1(\mathbb{R})$,*

2) *$\forall x \in \mathbb{R}$: $f(x,y)$ is continuous with respect to $y \in \Omega$.*

Then the function

$$F(y) := \int f(x,y)\, dx,$$

is continuous with respect to y.

We now describe the action of the Fourier transform on $L_1(\mathbb{R})$.

Theorem 27.5.

$$\widehat{F} : L_1(\mathbb{R}) \to C_0(\mathbb{R}) = \{g \in C(\mathbb{R}),\ g(y) \to 0 \text{ as } |y| \to \infty\}.$$

Proof.

1) First, we will prove the continuity. It is quite obvious: the integrand is continuous with respect to the parameter y:

$$\left|f(x)e^{-ixy}\right| \leqslant |f(x)|,$$

which is integrable, and e^{-ixy} is continuous. Thus, $\widehat{f}$ is continuous due to Lebesgue's dominated convergence theorem.

2) For $f_1,\ f_2 \in L_1(\mathbb{R})$, consider

$$\left|\widehat{f}_1(y) - \widehat{f}_2(y)\right| = \frac{1}{\sqrt{2\pi}}\left|\int_{\mathbb{R}} \big(f_1(x) - f_2(x)\big)e^{-ixy}\,dx\right| \leqslant \int_{\mathbb{R}} \left|f_1(x) - f_2(x)\right| dx = \\ = \frac{1}{\sqrt{2\pi}}\|f_1 - f_2\|_{L_1} \tag{27.2}$$

Thus, for $f_n \xrightarrow{L_1} f$, we have

$$\widehat{f}_n \rightrightarrows \widehat{f}$$

since bound (27.2) is independent of y.

3) Consider the Fourier transform of $f = \chi_{(a,b)}$:

$$\widehat{f} = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-ixy}\,dx = \frac{1}{\sqrt{2\pi}}\frac{e^{-iby} - e^{-iay}}{-iy},$$

so

$$\widehat{f}(y) \to 0 \quad \text{as} \quad |y| \to \infty.$$

Any integrable function can be approximated by a linear combination of indicators. For $f \in L_1(\mathbb{R})$, consider such a linear combination f_n; then,

$$\widehat{f}_n \rightrightarrows \widehat{f} \quad \text{and} \quad \widehat{f}_n(y) \to 0 \quad \text{as} \quad |y| \to \infty.$$

Therefore, $\widehat{f} \in C_0(\mathbb{R})$. □

Note that in the second part of the proof, we also established that the operator $\widehat{F}$ is bounded. In $C_0(\mathbb{R})$, the norm is defined by

$$\|\widehat{f}\| = \sup_{y\in\mathbb{R}}|\widehat{f}(y)|.$$

Therefore, (27.2) implies that

$$\|\widehat{F}\| \leqslant \frac{1}{\sqrt{2\pi}}.$$

It is worth mentioning that in physics, the direct Fourier transform is sometimes defined without the factor $1/\sqrt{2\pi}$, while the inverse transform has a factor of $1/(2\pi)$. We consider these transforms with equal factors to prove (in further) that $\widehat{F}$ and $\check{F}$ are unitary in L_2.

Properties of the Fourier Transform in L_1

Let us consider some properties of $\widehat{F} : L_1(\mathbb{R}) \to C_0(\mathbb{R})$:

1) $\operatorname{Ker}\widehat{F} = \{0\}$.

2) $\operatorname{Rn}\widehat{F} \neq C_0(\mathbb{R})$. Thus, we cannot consider the inverse, since it is unbounded due to the Banach bounded inverse theorem.

Recall the definition of Dini continuity at a point x:

$$\forall \varepsilon > 0 \ \exists \delta > 0: \quad \Big| \int_{-\delta}^{\delta} \frac{f(x+t) - f(x)}{t} \, dt \Big| < \varepsilon.$$

Theorem 27.6. *Let $f \in L_1(\mathbb{R})$ and f be Dini-continuous at a point x. Then*

$$f(x) = \frac{1}{\sqrt{2\pi}} \lim_{N \to \infty} \int_{-N}^{N} \widehat{f}(y) e^{ixy} \, dy.$$

This means that we can recover the value of f at the point where f is Dini-continuous via the inverse Fourier transform.

Example 27.7. *Consider $f = \chi_{(-a,a)}$. For this function, we have*

$$\widehat{f} = \frac{1}{\sqrt{2\pi}} \frac{e^{-iay} - e^{iay}}{-iy} = -\frac{1}{\sqrt{2\pi}} \frac{\sin ay}{y},$$

so $\widehat{f} \notin L_1$.

Further, $f \in L_1(\mathbb{R}) \Rightarrow \widehat{f} \in C_0(\mathbb{R})$ means that $\widehat{f} = o(1)$ as $|y| \to \infty$. Let us study two additional questions. What can be said about the differentiability of $\widehat{f}$? What is the decay rate of $\widehat{f}$?

First, let us recall the definition of $o = \overline{o}$ and $O = \underline{O}$.

Definition 27.3. *Let $f(x)$, $g(x)$ be some functions. We denote*

$$f = o(g) \quad as \quad x \to x_0$$

if

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0.$$

We denote

$$f = O(g)$$

if $\exists c$, C:

$$cg(x) \leqslant f(x) \leqslant Cg(x).$$

The rate at which $\widehat{f}$ decays depends on the smoothness of f:

Theorem 27.7. *Let $f \in L_1(\mathbb{R}) \cap C^1(\mathbb{R})$ and $f' \in L_1(\mathbb{R})$. Then*

$$\widehat{f}(y) = o\left(\frac{1}{|y|}\right) \quad as \quad |y| \to \infty.$$

Proof. Consider

$$\widehat{f}(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ixy} \, dx;$$

let us bring the exponential function inside the differential

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ixy} \, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) d \frac{e^{-ixy}}{-iy} \, dx$$

(here we cannot be concerned about the singularity at $y = 0$ since we consider $|y| \to \infty$) and integrate by parts:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) d \frac{e^{-ixy}}{-iy} \, dx = \frac{1}{\sqrt{2\pi}} f(x) \frac{e^{-ixy}}{-iy} \Big|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \frac{1}{iy} \int_{-\infty}^{\infty} f'(x) e^{-ixy} \, dx;$$

the latter integral exists due to the assumption $f' \in L_1$. For this integral, the estimate

$$\frac{1}{\sqrt{2\pi}} \frac{1}{iy} \int_{-\infty}^{\infty} f'(x) e^{-ixy} \, dx = o(1) \quad \text{as} \quad |y| \to \infty$$

holds due to $f' \in L_1$, thus,

$$\frac{1}{\sqrt{2\pi}} \frac{1}{iy} \int_{-\infty}^{\infty} f'(x) e^{-ixy} \, dx = o\left(\frac{1}{y}\right) \quad \text{as} \quad |y| \to \infty.$$

Let us prove that the same estimate holds for the first term. Since $f \in C^1$, due to the Newton–Leibniz theorem, we have

$$f(x) - f(a) = \int_a^x f'(t) \, dt;$$

since $f' \in L_1$, there exists the limit

$$f(\pm\infty) = \int_a^{\pm\infty} f'(t) \, dt + f(a).$$

The function f is integrable, so f must vanish at infinity:

$$\lim_{x \to \pm\infty} f(x) = 0;$$

otherwise, the integral

$$\int_{-\infty}^{\infty} f(x) \, dx$$

would not exist (meaning that $f \notin L_1(\mathbb{R})$). Thus,

$$f(x) \frac{e^{-ixy}}{-iy} \Big|_{x=-\infty}^{\infty} = 0 = o\left(\frac{1}{|y|}\right),$$

which completes the proof. □

Note that this theorem can be generalized to the following one.

Theorem 27.8. *Let $f \in L_1(\mathbb{R}) \cap C^k(\mathbb{R})$ and $f^{(j)} \in L_1(\mathbb{R})$ $\forall j = 1, \dots, k$. Then*

$$\widehat{f}(y) = o\left(\frac{1}{|y|^k}\right) \quad as \quad |y| \to \infty.$$

This theorem can be proved by induction.

Lecture 28. Fourier Transform in L_1, $\mathcal{S}$, and L_2

Properties of the Fourier Transform in L_1: Further Discussion

In the previous lecture, we discussed the Fourier transform in $L_1(\mathbb{R})$ and demonstrated that the smoother the function, the faster its Fourier transform decays. Now, we aim to establish the converse property, which relates the decay of the Fourier transform to the differentiability of the original function.

Theorem 28.1. *Let $f(x)$, $xf(x) \in L_1(\mathbb{R})$. Then $\widehat{f}(y) \in C^1(\mathbb{R})$.*

Proof. By definition,

$$\widehat{f}(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-ixy}\, dx.$$

Consider the formal derivative:

$$\frac{d\widehat{f}}{dy} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)(-ix)e^{-ixy}\, dx. \tag{28.1}$$

For the integrand, observe that

$$\left|f(x)(-ix)e^{-ixy}\right| \leqslant |x||f(x)|,$$

and since $|x||f(x)| \in L_1(\mathbb{R})$, the integral in (28.1) exists. Thus, by the Lebesgue dominated convergence theorem, differentiation under the integral is valid. Hence, the derivative of $\widehat{f}$ exists, and $\widehat{f}(y) \in C^1(\mathbb{R})$. □

This result can be generalized as follows:

Theorem 28.2. *Let $x^m f(x) \in L_1(\mathbb{R})$ for $m = 0, 1, \ldots, k$. Then $\widehat{f}(y) \in C^k(\mathbb{R})$.*

The proof follows by induction on k, applying the same reasoning iteratively to higher derivatives.

The Fourier Transform in Schwartz Space

Functions of the Schwartz class are both differentiable and decay rapidly. Therefore, $\mathcal{S}(\mathbb{R})$ is quite a suitable space for the Fourier transform.

We consider only the one-dimensional case, as the multi-dimensional one is essentially the same but requires more technical work.

By definition,

$$\mathcal{S} \equiv \mathcal{S}(\mathbb{R}) = \left\{ f \in C^\infty(\mathbb{R}) \ \forall p, q \in \mathbb{N} \cup \{0\} : \ \|f\|_{p,q} := \sup_{\mathbb{R}} |x^p f^{(q)}| < \infty \right\}.$$

A typical example of $f \in \mathcal{S}$ is $f(x) = e^{-x^2}$.

Definition 28.1. *A function f **decays exponentially** if*

$$\exists a > 0 : \quad e^{|a|x} f(x) \in L_1(\mathbb{R}).$$

Due to Theorems 27.8 and 28.2,

$$\widehat{F} : \mathcal{S} \to \mathcal{S},$$

as these theorems ensure the existence of all derivatives and their rapid decay. Notably, within the Schwartz space $\mathcal{S}$, the Fourier transform becomes a bijection, unlike in $L_1(\mathbb{R})$, where such a property does not hold.

Moreover, $\widehat{F}$ is a continuous operator in $\mathcal{S}$. However, the Schwartz space is not a normed one; it is equipped with countably many seminorms. Spaces of this kind are called *locally convex spaces.*

How can we define a continuous operator in such a space? Let X be a linear space with seminorms $\{p_\alpha\}_\alpha$. We say that $x_n \to x$ if $p_\alpha(x_n - x) \to 0\ \forall \alpha$; however, this is a sequential definition of convergence, while the topological one is more convenient for our purposes. The base of topology is

$$U_{x_0, p_{\alpha_1}, \ldots, p_{\alpha_n}}(\varepsilon) = \{x \in X : \ p_{\alpha_j}(x - x_0) < \varepsilon,\ i = 1, 2, \ldots, n\};$$

one can vary ε, the seminorms $p_{\alpha_1}, \ldots, p_{\alpha_n}$, and x_0.

In these terms, a continuous operator can be defined for a pair of spaces. Let $(X, \{p_\alpha\}_\alpha)$, $(Y, \{q_\beta\}_\beta)$ be locally convex spaces. Then an operator

$$A : X \to Y$$

is continuous if $\forall q_\beta\ \exists C,\ p_{\alpha_1},\ \ldots,\ p_{\alpha_n}$:

$$\forall x \in X : \quad q_\beta(Ax) \leqslant C \sum_{k=1}^{n} p_{\alpha_k}(x).$$

Let us define a pair of operators in $\mathcal{S}$:

$$Df = i\frac{d}{dx} f, \qquad Mf = xf.$$

It is easy to see that these operators are continuous in $\mathcal{S}$ since

$$\|Df\|_{p,q} = \|f\|_{p,q+1},$$

and

$$\|Mf\|_{p,q} = \sup_{\mathbb{R}} |x^p (xf)^{(q)}|.$$

Using the product rule, we obtain

$$\sup_{\mathbb{R}}|x^p(xf)^{(q)}| = \sup_{\mathbb{R}}|x^{p+1}f^{(q)} + qx^p f^{(q-1)}| \leqslant \|f\|_{p+1,q} + q\|f\|_{p,q-1}.$$

Why are these operators useful and important? These operators play a central role in Fourier analysis, as they provide a means of manipulating the Fourier transform algebraically without resorting to direct integration.

Theorem 28.3. *The following commutation relations hold in $\mathcal{S}$:*

$$\widehat{F}M = D\widehat{F}, \qquad \widehat{F}D = -M\widehat{F},$$

and

$$\check{F}M = -D\check{F}, \qquad \check{F}D = M\check{F}.$$

Proof. We will prove the statement only for the direct transform; the proof for the inverse transform is similar.

By definition,

$$\widehat{F}Mf = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} xf(x)e^{-ixy}\,dx$$

and

$$\widehat{D}\widehat{F}f = i\frac{d}{dy}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(x)e^{-ixy}\,dx.$$

Since $f \in \mathcal{S}$ decays rapidly, we can take the derivative inside the integral:

$$i\frac{d}{dy}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(x)e^{-ixy}\,dx = i\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(x)(-ix)e^{-ixy}\,dx = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(x)xe^{-ixy}\,dx.$$

Thus, it is clear that $D\widehat{F}f = \widehat{F}Mf$.

For the second relation, we have

$$-M\widehat{F}f = -\frac{y}{\sqrt{2\pi}}\int_{\mathbb{R}} f(x)e^{-ixy}\,dx \tag{28.2}$$

and

$$\widehat{F}Df = \frac{1}{\sqrt{2\pi}}\int if'(x)e^{-ixy}\,dx.$$

Integrating the latter by parts, we get

$$\frac{1}{\sqrt{2\pi}}\int if'(x)e^{-ixy}\,dx = \frac{1}{\sqrt{2\pi}}if(x)e^{-ixy}\Big|_{-\infty}^{\infty} - \frac{1}{\sqrt{2\pi}}\int i(-iy)f(x)e^{-ixy}\,dx,$$

where the boundary terms vanish due to the rapid decay of f, and the remaining integral coincides with (28.2). □

Theorem 28.4. *$\widehat{F} : \mathcal{S} \to \mathcal{S}$ is a continuous operator and a bijection.*

Proof.

1) Continuity. The proof is quite complicated. It involves constructing an equivalent system of seminorms for $\mathcal{S}$.

   Consider the same space with different systems of seminorms: $(X, \{p_\alpha\})$, $(X, \{q_\beta\})$. The equivalence of seminorms $\{p_\alpha\} \sim \{q_\beta\}$ means that the embeddings of these spaces to each other are continuous:

$$q_\beta(x) \leqslant C_1 \sum_{i=1}^{n} p_{\alpha_i}(x) \quad \forall x \in X$$

   for the embedding $(X, \{p_\alpha\}) \hookrightarrow (X, \{q_\beta\})$, and

$$p_\alpha(x) \leqslant C_2 \sum_{i=1}^{m} q_{\beta_i}(x) \quad \forall x \in X$$

   for the embedding $(X, \{p_\alpha\}) \hookleftarrow (X, \{q_\beta\})$. Given these two bounds, the systems of seminorms are equivalent: $\{p_\alpha\} \sim \{q_\beta\}$.

   Introduce

$$\|f\|_{p,q,1} = \int_{\mathbb{R}} |x^p f^{(q)}(x)|\, dx.$$

   Let us prove that the system of these seminorms is equivalent to $\|\cdot\|_{p,q}$. Due to the Newton–Leibniz theorem,

$$x^p f^{(q)}(x) = \int_{-\infty}^{x} \left(t^p f^{(q)}(t)\right)' dt = \int_{-\infty}^{x} \left(t^p f^{(q+1)}(t) + pt^{p-1} f^{(q)}(t)\right) dt;$$

   thus,

$$|x^p f^{(q)}(x)| \leqslant \int_{-\infty}^{\infty} |t^p f^{(q+1)}(t)|\, dt + p \int_{-\infty}^{\infty} |t^{p-1} f^{(q)}(t)|\, dt = \|f\|_{p,q+1,1} + p\|f\|_{p-1,q,1}.$$

   The bound is independent of x; thus, taking the supremum on the left-hand side, we obtain

$$\sup_{x\in\mathbb{R}} |x^p f^{(q)}(x)| \leqslant \|f\|_{p,q+1,1} + p\|f\|_{p-1,q,1}.$$

   Conversely,

$$\|f\|_{p,q,1} = \int_{-\infty}^{\infty} |x^p f^{(q)}|\, dx = \int_{-\infty}^{\infty} \frac{|x^p f^{(q)}|(1+x^2)}{1+x^2}\, dx$$

where we take the supremum of the numerator:

$$\int_{-\infty}^{\infty} \frac{|x^p f^{(q)}|(1+x^2)}{1+x^2}\,dx \leqslant \int_{-\infty}^{\infty} \frac{\sup|x^p f^{(q)}|(1+x^2)}{1+x^2}\,dx =$$
$$= \left(\sup|x^p f^{(q)}|(1+x^2)\right)\int_{-\infty}^{\infty}\frac{dx}{1+x^2} =$$
$$= \pi \sup|x^p f^{(q)}|(1+x^2) = \pi\left(\|f\|_{p,q}+\|f\|_{p+2,q}\right),$$

which finally gives $\|f\|_{p,q,1} \leqslant \pi\left(\|f\|_{p,q}+\|f\|_{p+2,q}\right)$, so the systems $\{\|f\|_{p,q}\}$, $\{\|f\|_{p,q,1}\}$ are equivalent.

Next, let us find the bound for $|y^p \widehat{f}^{(q)}(y)|$. Using the operators $\boldsymbol{M}$ and $\boldsymbol{D}$, we write

$$|y^p \widehat{f}^{(q)}(y)| = |M^p D^q \widehat{F} f|,$$

where we omit the imaginary unit since the absolute value is considered; next,

$$|M^p D^q \widehat{F} f| = |\widehat{D}^p M^q f| = \frac{1}{\sqrt{2\pi}}\left|\int_{\mathbb{R}} i^p (x^q f(x))^{(p)}\,dx\right| \leqslant \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\left|(x^q f(x))^{(p)}\right|dx,$$

where, due to the product rule,

$$(x^q f(x))^{(p)} = \sum_{j=0}^{\min(p,q)} C_p^j x^{q-j}\frac{q!}{j!} f^{(j)},$$

so, denoting

$$C = \max\nolimits_{j,q,p} C_p^j \frac{q!}{j!}$$

and using the triangle inequality, we get

$$\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\left|(x^q f(x))^{(p)}\right|dx \leqslant C\sum_{j=0}^{\min(p,q)} \|f\|_{q-1,j,1},$$

which gives finally

$$|y^p \widehat{f}^{(q)}(y)| \leqslant C\sum_{j=0}^{\min(p,q)} \|f\|_{q-1,j,1}.$$

The right-hand side is independent of y. Taking the supremum on the left-hand side, we obtain

$$\|\widehat{F} f\|_{p,q} \leqslant C\sum_{j=0}^{\min(p,q)} \|f\|_{q-j,j,1},$$

therefore, $\widehat{F}$ is continuous.

2) Bijectivity. Let us formulate the auxiliary statement:

Lemma 28.1. *Let* $f, g \in \mathcal{S}$, $f(x_0) = g(x_0)$, *and* $A \in \mathcal{L}(\mathcal{S})$ *commutes with* M: $AM = MA$. *Then*

$$(Af)(x_0) = (Ag)(x_0).$$

Proof. Consider $h(x) = f(x) - g(x)$; then $h(x_0) = 0$ and

$$h(x) = (x - x_0)h'(x_0) + o(x - x_0) = (M - x_0 I)h'(x_0) + o(x - x_0).$$

Next,

$$Ah(x) = A(M - x_0 I)h'(x_0)\big(1 + o(1)\big);$$

since $AM = MA$, we can rewrite it as follows:

$$A(M - x_0 I)h'(x_0)\big(1 + o(1)\big) = (M - x_0 I)Ah'(x_0)\big(1 + o(1)\big).$$

Since $A \in \mathcal{L}(\mathcal{S})$, $Ah'(x_0)\big(1 + o(1)\big) \in \mathcal{S}$; further,

$$(Ah)(x_0) = 0,$$

since $M - x_0 I$ vanishes at x_0, and, therefore, $Af(x_0) = Ag(x_0)$. □

From this lemma, one can see that A is a multiplication operator:

$$AM = MA \quad \Rightarrow \quad Af(x) = a(x)f(x).$$

Let us fix a point $x = x_0$. Thus, we live in a one-dimensional space where we can only vary the value $f(x_0)$. A linear operators in one-dimensional space is just the multiplication by a constant:

$$Af(x_0) = a(x_0)f(x_0).$$

Allowing x_0 to vary, we obtain

$$Af(x) = a(x)f(x).$$

Next, let us rewrite

$$\widehat{F}\check{F}M = \widehat{F}(-D\check{F}) = M\widehat{F}\check{F},$$

so, since the composition $\widehat{F}\check{F}$ commutes with M, this composition is a multiplication by a function:

$$\widehat{F}\check{F}f = a(x)f(x).$$

Next, since

$$\widehat{F}\check{F}D = \widehat{F}M\check{F} = D\widehat{F}\check{F},$$

i.e., the multiplication operator $\widehat{F}\check{F}$ commutes with differentiating, we conclude that $a(x) = c \equiv \text{const}$:

$$D(af) = i(af)' = iaf' + ia'f, \quad aDf = aif' \quad \Rightarrow \quad a' = 0.$$

Thus, $\widehat{F}\check{F} = cI$. To find c, let us calculate the Fourier transform of some function, e.g., e^{-ax^2}, $a > 0$. Omitting the constant factor, we have

$$\int_{\mathbb{R}} e^{-ax^2} e^{-ixy}\,dx = \int_{\mathbb{R}} e^{-\left(\sqrt{a}x+\frac{iy}{2\sqrt{a}}\right)^2} e^{-\frac{y^2}{4a}}\,dx = e^{-\frac{y^2}{4a}} \int_{\Gamma} e^{-\left(\sqrt{a}x+\frac{iy}{2\sqrt{a}}\right)^2}\,d\ell,$$

where Γ is a contour given by $\operatorname{Im}\Gamma = iy/(2\sqrt{a})$. Here, we have the integration over a line parallel to $\mathbb{R}$ in the complex plane. Let us deform the integration contour as depicted in Fig. 28.1.

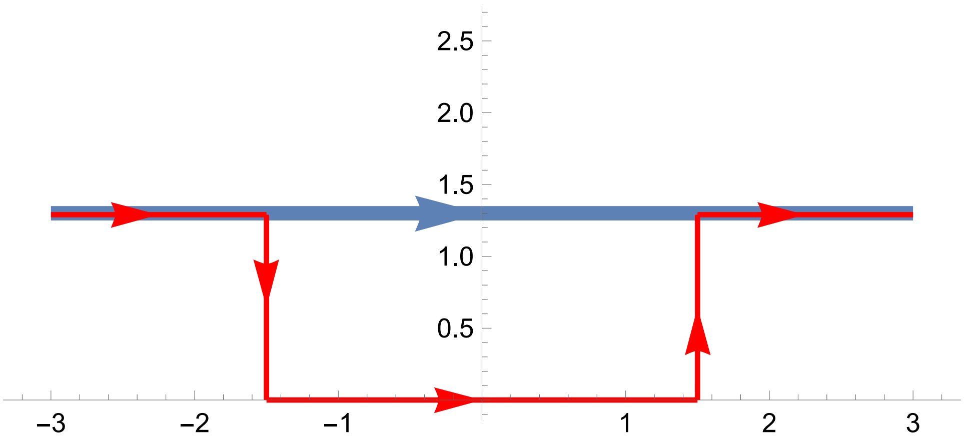

Рис. 28.1. Deformation of the blue contour to the red one.

In the closed contour (see the figure), the integrand is holomorphic; on the "vertical" parts, the integrand decays exponentially as $|x| \to \infty$. Therefore, the integration over Γ and $\mathbb{R}$ gives the same result. Then, we rewrite

$$e^{-\frac{y^2}{4a}} \int_{\Gamma} e^{-\left(\sqrt{a}x+\frac{iy}{2\sqrt{a}}\right)^2}\,dx = e^{-\frac{y^2}{4a}} \int_{\mathbb{R}} e^{-ax^2}\,dx.$$

Further, introducing $u = \sqrt{a}x$, $du = \sqrt{a}x$, we obtain

$$\frac{1}{\sqrt{a}} e^{-\frac{y^2}{4a}} \int_{\mathbb{R}} e^{-u^2}\,du = \sqrt{\frac{\pi}{a}} e^{-\frac{y^2}{4a}}.$$

Thus,

$$\widehat{F}(e^{-ax^2}) = \frac{1}{\sqrt{2a}} e^{-\frac{y^2}{4a}}.$$

Setting $a = 1/2$, we get

$$\widehat{F}(e^{-x^2/2}) = e^{-y^2/2}.$$

Therefore, $c = 1$, and

$$\widehat{F}\check{F} = I,$$

which proves that $\widehat{F}$ is a bijection. □

The Fourier Transform in L_2

Recall that

$$\|\widehat{F}\|_{L_1\to C_0} \leqslant \frac{1}{\sqrt{2\pi}}.$$

For the function $f(x) = e^{-x^2/2}$, which belongs to $L_1(\mathbb{R})$, we obtain the equality:

$$\|\widehat{F}\|_{L_1\to C_0} = \frac{1}{\sqrt{2\pi}}.$$

Let us proceed to extend the Fourier transform to $L_2(\mathbb{R})$. We already considered the Fourier transform

$$\widehat{F} : L_1 \to C_0.$$

Here, L_1 is quite a fine space, since it is normed. However, being defined on this space, $\widehat{F}$ is not surjective.

Next, we considered

$$\widehat{F} : \mathcal{S} \to \mathcal{S}.$$

Here, we have a bijection, however, the space itself is not optimal, since it is not a normed space.

But still, $\mathcal{S}$ is dense in $L_2(\mathbb{R})$; let us prove it.

First, observe that a linear combination of characteristic functions is dense in $L_2(\mathbb{R})$. By smoothing $\chi_{(a,b)}$, we ensure that it lies in $\mathcal{S}$, see Fig. 28.2; denote this function by $\varphi_{a,b,\varepsilon}$.

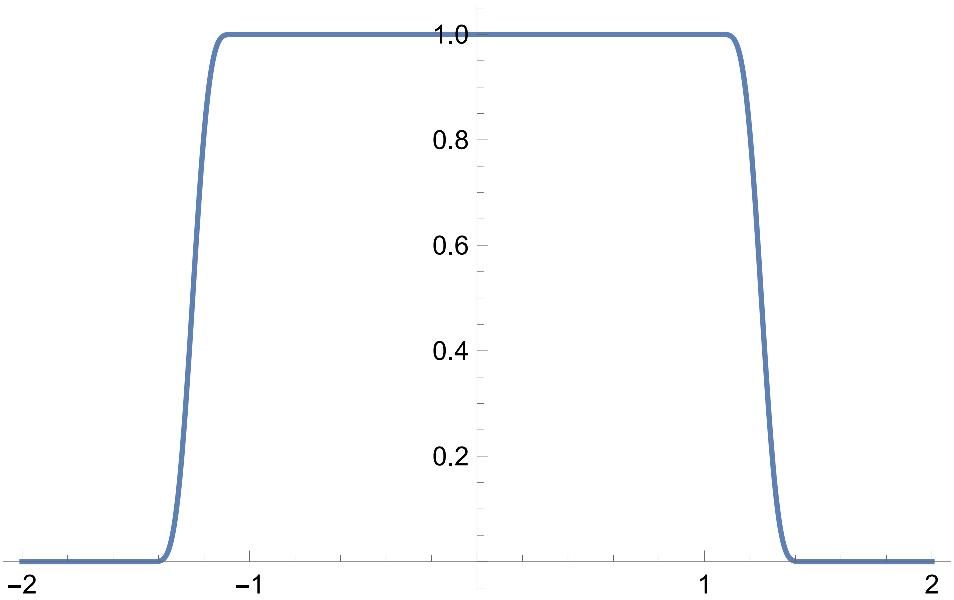

Рис. 28.2. A smooth characteristic function $\varphi_{a,b,\varepsilon}$.

If we shrink the "smoothing" parts to a and b, this function tends to $\chi_{(a,b)}$. The derivative of $\varphi_{a,b,\varepsilon}$ is the difference of two hat functions, see Fig. 28.3.

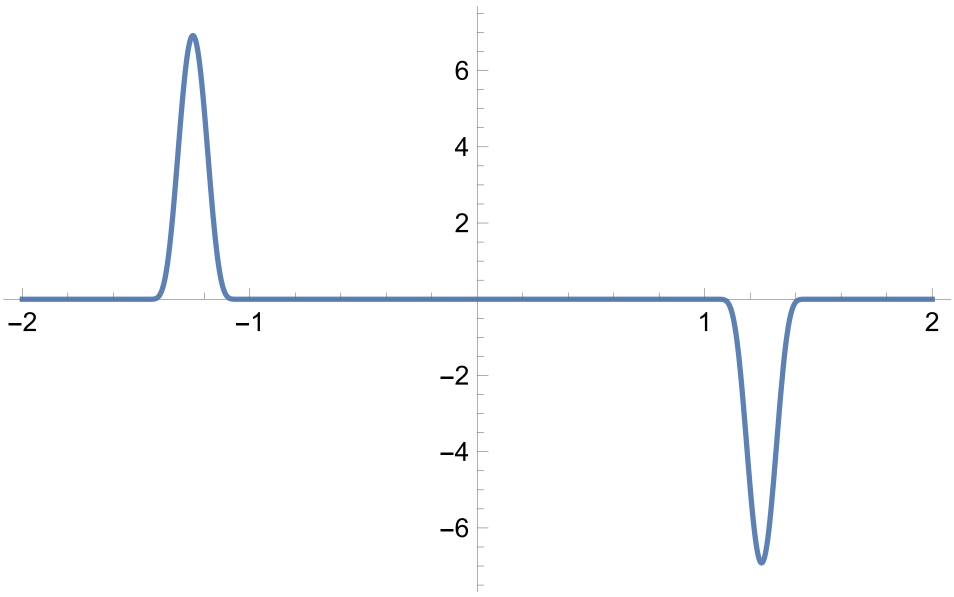

Рис. 28.3. The graph of $\varphi'_{a,b,\varepsilon}$.

The support of $\varphi'_{a,b,\varepsilon}$ (the closure of the set where the value of the function is nonzero) is $[a-\varepsilon,a]\cup[b,b+\varepsilon]$.

Define the space of *test functions*

$$\mathscr{D} = \{\varphi \in C^{\infty}(\mathbb{R}) : \ \operatorname{supp}\varphi \text{ is compact}\}.$$

Consider

$$\psi(x) = \begin{cases} e^{-1/x}, & x > 0, \\ 0, & x \leqslant 0. \end{cases}$$

This function is obviously differentiable for $x < 0$ and $x > 0$. At 0, it is also differentiable; one can prove by induction that

$$\psi^{(n)}(x) = P_{2n}\left(\frac{1}{x}\right)\psi(x),$$

where P_k is some polynomial of degree k.

Unitarity of the Fourier Transform in L_2

The following statement is the integral analog of Parseval's identity:

Statement 28.1. *Let $f, g \in \mathcal{S}$. Then*

$$\left(\widehat{f}, \widehat{g}\right)_{L_2} = \left(f, g\right)_{L_2}.$$

Proof. Consider

$$\left(\widehat{f}, \widehat{g}\right)_{L_2} = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{f}(y) \overline{\int_{\mathbb{R}} g(x) e^{-ixy}\, dx}\, dy;$$

since all functions here belong to $\mathcal{S}$, one use Fubini's theorem to switch the order of integration, and integrate with respect to y first:

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{f}(y) \overline{\int_{\mathbb{R}} g(x) e^{-ixy}\, dx}\, dy = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{f}(y) e^{ixy} \overline{\int_{\mathbb{R}} g(x)\, dx}\, dy.$$

Thus, we have the inverse Fourier transform of $\widehat{f}$:

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{f}(y) e^{ixy} \overline{\int_{\mathbb{R}} g(x)\, dx}\, dy = \int_{\mathbb{R}} f(x) \overline{g(x)}\, dx = (f, g)_{L_2},$$

so $\left(\widehat{f}, \widehat{g}\right)_{L_2} = \left(f, g\right)_{L_2}$. □

Since $\mathcal{S}$ is dense in $L_2(\mathbb{R})$, we can extend $\widehat{F}$ to the whole L_2. One can see that $\widehat{F}$ is unitary in that case.

For $f \in L_2(\mathbb{R})$,

$$\exists f_n \in \mathcal{S} : \quad f_n \xrightarrow{L_2} f, \quad \|\widehat{f_n} - \widehat{f_m}\|_{L_2} = \|f_n - f_m\|_{L_2} \to 0,$$

so $\{\widehat{f}_n\}$ is a Cauchy sequence. Thus, there exists a limit; denote

$$\widehat{f} := \lim_{n\to\infty} \widehat{f}_n.$$

This function is well-defined, since for any other sequence $\widetilde{h}_n \to f$, one can consider

$$f_1, h_1, f_2, h_2, \dots,$$

and

$$\lim_{n\to\infty} \widehat{f}_n = \lim_{n\to\infty} \widehat{h}_n.$$

Theorem 28.5 (The Plancherel Theorem)**.** *There exists a unique unitary operator $U : L_2(\mathbb{R}) \to L_2(\mathbb{R})$ such that*

$$U\big|_{L_1\cap L_2} = \widehat{F}.$$

Note that the Fourier transform in L_2 is usually defined by

$$\widehat{f} = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ixy}\, dx.$$

However, for $f \in L_2(\mathbb{R})$, this is an ill-defined operator, but it is well-defined in L_1 and $\mathcal{S}$. Nevertheless, one can prove that in the sense of the Cauchy principal value, the Fourier transform is well-defined in L_2:

$$\widehat{f}(y) = \frac{1}{\sqrt{2\pi}} \int\limits_{N\to\infty -N}^{N} f(x) e^{-ixy}\, dx.$$

Spectrum of the Fourier Transform in L_2

Since $\widehat{F}$ is unitary in L_2, it is invertible. Further, consider

$$(\widehat{F}^2 f)(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{f}(y) e^{-ixy}\, dy = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{f}(y) e^{i(-x)y}\, dy = f(-x),$$

and thus,

$$\widehat{F}^4 = I.$$

Consider $p(t) = t^4 - 1$. By the spectral mapping theorem,

$$\sigma\big(p(\widehat{F})\big) = p\big(\sigma(\widehat{F})\big);$$

on the left-hand side, we have $\sigma(0) = \{0\}$, therefore,

$$\sigma(\widehat{F}) \subset \{\pm 1, \pm i\}.$$

Consider the Hermite system, i.e., the orthogonalization of $\{e^{-x^2/2}x^n\}$: $\{p_n(x)e^{-x^2/2}\}$. We have seen already that

$$\widehat{F}e^{-x^2/2} = e^{-x^2/2};$$

for the next function, we have

$$\widehat{F}\left(xe^{-x^2/2}\right) = \widehat{F}Me^{-x^2/2} = D\widehat{F}e^{-x^2/2} = -i\frac{d}{dx}e^{-x^2/2} = -ixe^{-x^2/2}.$$

For the next two, we have

$$\widehat{F}\left(x^2 - \frac{1}{2}\right)e^{-x^2/2}, \qquad \widehat{F}\left(x^3 - \frac{3}{2}x\right)e^{-x^2/2}.$$

Exercise 28.1. *In terms of operators M and D, find out what is the action of $\widehat{F}$ here.*

In fact, for the Hermite function $f_n(x) = H_n(x)e^{-x^2/2}$, we have

$$\widehat{F}f_n = (-i)^n f_n,$$

and each of four eigenvalues is of infinite multiplicity.

Next, one can see that the Fourier transform commutes with the quantum harmonic oscillator given by

$$Af = -f'' + x^2 f = (D^2 + M^2)f,$$

since M^2 and D^2 just swap under $\widehat{F}$. Consider the action of A on $f_0(x) = e^{-x^2/2}$:

$$Af_0 = e^{-x^2/2}.$$

Example: Calculation of the Fourier Transform

Let us calculate the Fourier transform of

$$f(x) = \frac{1}{x^2 + a^2}, \quad a > 0.$$

Write out the expression for the Fourier transform:

$$\widehat{f}(y) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{1}{x^2 + a^2} e^{-ixy}\, dx.$$

Taking a semicircle in the upper half-plane for $y > 0$ (lower half-plane, for $y < 0$) connected with the interval $[-R, R]$ of the real axis as an integration contour, we see that the function has two residues: $\pm ia$. By applying the residue theorem, we compute:

$$\widehat{F}f = 2\pi i \begin{cases} \dfrac{e^{ay}}{2ia}, & y < 0, \\ -\dfrac{e^{-ay}}{-2ia}, & y > 0, \end{cases}$$

so

$$\widehat{F} f = \sqrt{\frac{\pi}{2}} \frac{e^{-a|y|}}{a}.$$

Next, for

$$g(x) = \frac{x}{x^2+1},$$

one can find the Fourier transform via residues, however, it is easier to use the representation

$$g = Mf,$$

and then,

$$\widehat{g} = \widehat{F} M f = D \widehat{F} f = -i \frac{d}{dy} \sqrt{\frac{\pi}{2}} \frac{e^{-a|y|}}{a} = -i \sqrt{\frac{\pi}{2}} e^{-a|y|} \operatorname{sgn} y.$$

One can see that this function does not belong to C_0, however, it belongs to L_2.

Lecture 29. Test Functions and Distributions

Locally Convex Spaces

Earlier, we studied linear bounded (continuous) operators and functionals in normed spaces. For a good share of applications, the normed spaces is not a convenient tool, as linear spaces without a norm often appear in applications. More precisely, we will consider *polynormed* (*locally convex*) spaces.

Definition 29.1. *A linear space X is called a* ***polynormed*** *or locally convex space if on X, there is a family of seminorms $\{p_\alpha\}_{\alpha\in\Lambda}$ (here, Λ can be of any cardinality); it is also denoted by $\big(X,\{p_\alpha\}\big)$.*

Convergence *in X is defined by the following. We say that $x_n \to x$ if $\forall\alpha$: $p_\alpha(x_n - x) = 0$.*

We say that the family of seminorms $\{p_\alpha\}$ ***distinguishes (separates) points*** *if*

$$\forall x,y,\ x\neq y, \quad \exists p_\alpha : p_\alpha(x-y)\neq 0.$$

Such a family $\{p_\alpha\}$ is called ***separating****. The seminorms in X are supposed to be separating.*

Note that the separating property of families of seminorms is quite important: it guarantees the uniqueness of the limit. For a nonseparating family of seminorms, the limit of x_n may not be unique. A little later, we will consider a topological approach for the polynormed spaces and discuss it again.

Any polynormed space $\big(X,\{p_\alpha\}\big)$ is a topological space. The base of topology is given by

$$u_{\varepsilon,p_{\alpha_1},\dots,p_{\alpha_n}}(x_0) = \{x\in X:\ \forall i=1,\dots,n\ p_{\alpha_i}(x-x_0)<\varepsilon\};$$

it is a ball of radius ε with center at x_0.

One can see that the convergence in the sense of definition above coincides with the convergence in topological space X, given the topology defined by the base above. An open set in this topology is an arbitrary union of balls $u_{\varepsilon,p_{\alpha_1},\dots,p_{\alpha_n}}$. Thus,

$$p_\alpha(x_n-x)\to 0 \quad \forall\alpha \qquad \Leftrightarrow \qquad x\in u_{\varepsilon,\alpha}(x) \quad \forall\alpha.$$

Consider the following examples:

Example 29.1. *1) $s \ni x = (x_1, x_2, \dots)$, the space of all sequences, equipped with the countable number of seminorms $p_n(x) = |x_n|$, $n = 1, 2, \dots$.*

2) $C_e[0,1]$*, the space of continuous functions on* $[0,1]$ *with pointwise convergence defined by the following family of seminorms:*

$$p_x(f) = |f(x)|, \quad x \in [0,1];$$

this family has the cardinality of the continuum.

Continuous Operators on Locally Convex Spaces

As we learned earlier, any continuous operator between normed spaces is a bounded operator. In a locally convex space, there are no tools allowing us to define a bounded operator, however, a continuous operator can be defined.

Definition 29.2. *Let* $\big(X,\{p_\alpha\}_{\alpha\in\Lambda}\big)$ *and* $\big(Y,\{q_\beta\}_{\beta\in\Gamma}\big)$ *be locally convex spaces. Let* $A \in \mathcal{L}(X,Y)$*, where* $\mathcal{L}(X,Y)$ *is the set of linear operators* $X \to Y$*. We say that* $A : X \to Y$ *is a continuous operator if*

$$\forall \beta \in \Gamma \quad \exists C > 0,\ \alpha_1,\dots,\alpha_n \in \Lambda \text{ such that } \forall x \in X: \quad q_\beta(Ax) \leqslant C\sum_{k=1}^{n} p_{\alpha_k}(x).$$

For example, given a single seminorm defined by $p_\alpha = \|\cdot\|_X$, $q_\beta = \|\cdot\|_Y$, then this definition becomes the definition of a bounded linear operator between normed spaces, since we have

$$\|Ax\| \leqslant C\|x\|.$$

In particular, f is a *continuous linear functional* on X, i.e., $f \in X'$, if

$$\exists C > 0,\ \alpha_1,\dots,\alpha_n \in \Lambda: \quad \forall x \in X \quad |f(x)| \leqslant C\sum_{j=1}^{n} p_{\alpha_j}(x);$$

note that this is due to $f : X \to \mathbb{C}$ and the fact that $\mathbb{C}$ has a single seminorm (which is actually a norm).

Spaces of Test Functions

Let us note that we will study only one-dimensional case since it allows us to omit technical difficulties while delivering the same idea. However, the generalization to the multi-dimensional case is quite straightforward.

First, consider the space

$$\mathscr{D} = \{\varphi \in C^\infty(\mathbb{R}), \operatorname{supp}\varphi \text{ is compact}\}.$$

A support of a function is the closure of a set at which the function takes nonzero values:

$$\operatorname{supp}\varphi := \overline{\{x \in \mathbb{R}: \ \varphi(x) \neq 0\}}.$$

In multi-dimensional case, one must replace $\mathbb{R}$ with $\mathbb{R}^n$ here and above; it essentially changes nothing.

Note that for $\varphi \in \mathscr{D}$,

$$\exists\, [a,b]: \quad \operatorname{supp}\varphi \subset [a,b].$$

Despite the fact that $\varphi \in \mathscr{D}$ has infinitely many derivatives, there are no analytic functions in $\mathscr{D}$ except for $f \equiv 0$. Let us show that this function is not the only one belonging to $\mathscr{D}$. Define

$$\Psi(x) = \begin{cases} e^{-1/x}, & x > 0, \\ 0, & x \leqslant 0. \end{cases}$$

One can see that Ψ has infinitely many derivatives to the left and to the right of 0. By induction, one can prove that

$$\Psi^{(n)}(x) = P_n\left(\frac{1}{x}\right)\Psi(x),$$

where P_n is a polynomial of degree $2n$. Thus, $\Psi^{(n)}(x) \to 0$ as $x \to 0$ $\forall n > 0$. Therefore, $\Psi(x) \in C^{\infty}(\mathbb{R})$.

Using $\Psi(x)$, define

$$\varphi_{a,b}(x) = \Psi(b-x)\cdot\Psi(x-a), \quad a < b.$$

This is a so-called *hat function*, see Fig. 29.1.

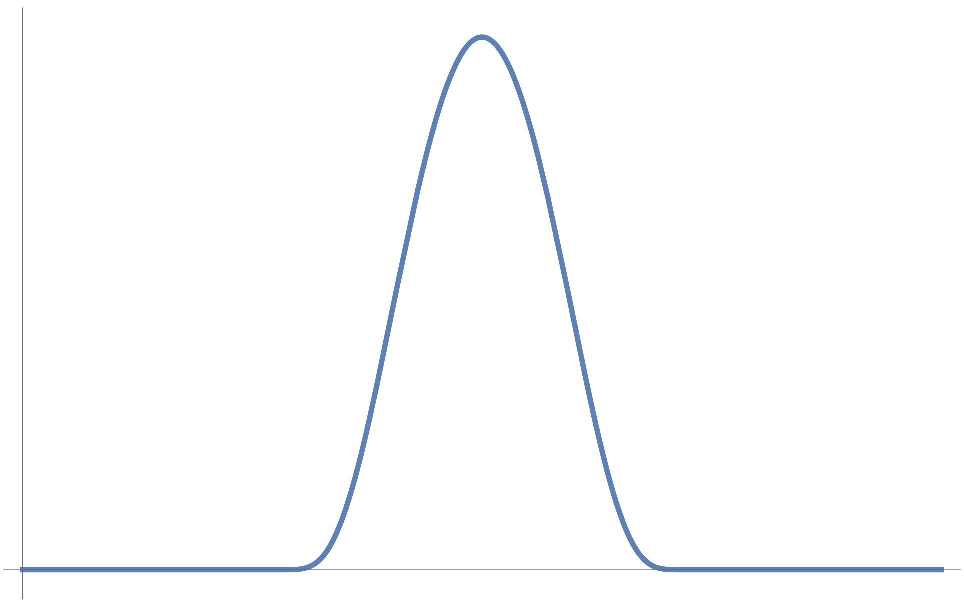

Рис. 29.1. Graph of $\varphi_{a,b}(x)$.

One can see that $\varphi_{a,b}$ is supported on $[a,b]$.

Let us make $\mathscr{D}$ a locally convex space, i.e., equip it with a family of separating seminorms.

Definition 29.3. *A seminorm p is **admissible** on $\mathscr{D}$ if $\forall a>0\ \exists C=C(a)>0\ \exists n=n(a)\in\mathbb{N}\cup\{0\}$:*

$$\forall\varphi\in\mathscr{D}(-a,a):\qquad p(\varphi)\leqslant C\cdot\|\varphi\|_{C^n[-a,a]},$$

where $\varphi\in\mathscr{D}(-a,a)$ means that $\operatorname{supp}\varphi\subset[-a,a]$.

Note that the standard norm in $C^n[-a,a]$ is defined by

$$\|\varphi\|_{C^n[-a,a]}=\sum_{k=0}^{n}\max_{[-a,a]}\left|\varphi^{(k)}(x)\right|.$$

Further, $\mathscr{D}$ is a locally convex space with *all* admissible seminorms.

Consider the following examples of admissible seminorms.

Example 29.2. *1) The standard C^n-norm is admissible.*

2) Consider

$$p(\varphi)=\sum_{k=0}^{\infty}\left|\varphi^{(k)}(k)\right|.$$

This is an admissible seminorm. Let us demonstrate it; suppose $\varphi\in\mathscr{D}(-a,a)$. Then

$$p(\varphi)\leqslant\|\varphi\|_{C^{[a]+1}[-a,a]},$$

that is, $C(a)=1$ and $n(a)=[a]+1$.

3) Let $\alpha(x)\in C(\mathbb{R})$. Define

$$p(\varphi)=\sup_{\mathbb{R}}|\alpha(x)\varphi(x)|;$$

note that this in fact is a supremum over the support of φ. It is an admissible seminorm; for $\varphi\in\mathscr{D}(-a,a)$, we have

$$p(\varphi)\leqslant\max_{[-a,a]}|\alpha(x)|\cdot\|\varphi\|_{C[-a,a]},$$

so $C(a)=\max\limits_{[-a,a]}|\alpha(x)|$ and $n(a)=0$.

Convergence in $\mathscr{D}$

Definition 29.4. *We say that $\varphi_n \to 0$ in $\mathscr{D}$ if*

1) $\exists [a,b]$: $\forall n$ $\operatorname{supp}\varphi_n \subset [a,b]$.

2) $\forall j \geqslant 0$: $\varphi_n^{(j)} \underset{[a,b]}{\rightrightarrows} 0$.

We say that $\varphi_n \to \varphi$ in $\mathscr{D}$ if $\varphi_n - \varphi \to 0$ in $\mathscr{D}$.

Theorem 29.1. *The convergence in $\mathscr{D}$ with respect to all admissible seminorms is equivalent to 1), 2) in Definition 29.4.*

Proof. Let us first prove that 1), 2) implies the convergence with respect to all admissible seminorms. Suppose p is an admissible seminorm, and $c = \max\big(|a|,|b|\big)$. Then $\forall \varphi \in \mathscr{D}(-c,c)$:

$$p(\varphi) \leqslant C\|\varphi\|_{C^n[-c,c]}.$$

Since $\operatorname{supp}\varphi_n \subset [a,b] \subset [-c,c]$, due to point 2),

$$p(\varphi) \leqslant C\|\varphi\|_{C^n[-c,c]} \to 0.$$

Thus, we have the convergence with respect to all admissible seminorms.

Next, let us prove the converse statement by contradiction. First, suppose that $\varphi_n \to 0$ with respect to all admissible seminorms and 1) does not hold. Without loss of generality, assume that $\exists x_k \to +\infty$:

$$\varphi_{n_k}(x_k) \neq 0, \quad n_k \to \infty.$$

Define the following continuous function:

$$\alpha = \begin{cases} \dfrac{1}{\varphi_{n_k}(x_k)}, & x = x_k, \\ \text{for } x \neq x_k, & \text{take a continuation as a piecewise linear function.} \end{cases}$$

We know that

$$p(\varphi) = \sup_{\mathbb{R}} |\alpha(x)\varphi(x)|$$

is an admissible seminorm, however,

$$p(\varphi_{n_k}) \geqslant 1 \neq 0,$$

which gives a contradiction.

Therefore, 1) holds. Next, there exists $[a,b]$ such that

$$\operatorname{supp}\varphi_n \subset [a,b].$$

Since the standard norm is admissible, we immediately obtain that 2) holds as well. □

Exercise 29.1. *Check that the space $\mathscr{D}$ is not metrizable: one cannot define the convergence in $\mathscr{D}$ equivalent to 1), 2) by a metric.*

Distributions

Definition 29.5. *The space $\mathscr{D}'$ of linear continuous functionals on $\mathscr{D}$ is called a **space of distributions**.*

Consider the following examples:

Example 29.3. *1) Dirac delta function $\delta(x) \in \mathscr{D}'$. It is defined as follows:*

$$\forall \varphi \in \mathscr{D}: \quad \langle \delta(x), \varphi \rangle = \varphi(0).$$

Remark 29.1. *Earlier, for the action of $f \in X^*$ at $\varphi \in X$, we used the notation $f(\varphi)$. However, in the theory of distributions, the following notation (the Dirac notation) is used:*

$$\langle f, \varphi \rangle := f(\varphi).$$

Let us verify that $\delta(x) \in \mathscr{D}'$. Take $\varphi \in \mathscr{D}(-a,a)$. Then,

$$|\langle \delta(x), \varphi \rangle| = |\varphi(0)| \leqslant \|\varphi\|_{C[-a,a]},$$

that is, the action of $\delta(x)$ is dominated by the standard norm, as any other admissible seminorm. Therefore, $\delta \in \mathscr{D}'$.

2) In prior providing the next example, let us give a definition.

Definition 29.6. *We say that $f \in L_{1,\mathrm{loc}}$ if*

$$\forall [a,b] \subset \mathbb{R}: \quad f \in L_1[a,b].$$

This means that f is integrable on any compact set. For example, any continuous function belongs to $L_{1,\mathrm{loc}}$.

Define a distribution F_f by

$$\langle F_f, \varphi \rangle := \int_{\mathbb{R}} f(x)\varphi(x)\,dx.$$

Then, $\forall \varphi \in \mathscr{D}(-a,a)$:

$$\left|\langle F_f, \varphi \rangle\right| \leqslant \int_{-a}^{a} |f(x)| \cdot |\varphi(x)|\,dx \leqslant \int_{-a}^{a} |f(x)|\,dx \cdot \max_{[-a,a]} |\varphi(x)|;$$

thus,

$$C(a) = \int_{-a}^{a} |f(x)|\, dx, \qquad \max_{[-a,a]} |\varphi(x)| = \|\varphi\|_{C[-a,a]}.$$

Therefore, $F_f \in \mathscr{D}'$. In other words, any locally integrable function defines a distribution. In further, we will make no distinction between a function and a functional it defines: for $f \in L_{1,\text{loc}}$, we will use the same letter for the functional, i.e., $F_f \equiv f$; the action of this functional is given by

$$\langle f, \varphi \rangle := \int_{\mathbb{R}} f(x)\varphi(x)\, dx,$$

where $f \in \mathscr{D}'$ on the left-hand side.

Regular and Singular Distributions

Definition 29.7. *If for $F \in \mathscr{D}'$ there exists $f \in L_{1,\text{loc}}$ such that*

$$\forall \varphi \in \mathscr{D}: \quad \langle F, \varphi \rangle = \int_{\mathbb{R}} f(x)\varphi(x)\, dx,$$

then F is called a ***regular*** *distribution.*

Otherwise, F is called a ***singular*** *distribution.*

Theorem 29.2. *$\delta(x)$ is a singular distribution.*

Remark 29.2. *Although $\delta(x)$ is a singular distinction, the integral notation is used sometimes; it is nothing more than a notation.*

Proof by contradiction. Suppose there exists $f \in L_{1,\text{loc}}$ such that

$$\langle \delta(x), \varphi(x) \rangle = \int f(x)\varphi(x)\, dx.$$

Define the function (a smooth approximation of an indicator) $\varphi_{a,b,\varepsilon}$, see Fig. 29.2.

For this function, we have

$$\varphi_{a,b,\varepsilon}\Big|_{[a,b]} \equiv 1, \qquad \varphi_{a,b,\varepsilon}(x)\Big|_{\mathbb{R}\setminus[a-\varepsilon,b+\varepsilon]} \equiv 0;$$

on $(a-\varepsilon, a)$ and $(b, b+\varepsilon)$ the function is continued in a smooth way; one can see that

$$\varphi_{a,b,\varepsilon}(x) \to \chi_{[a,b]} \quad \text{as} \quad \varepsilon \to 0;$$

the convergence is pointwise.

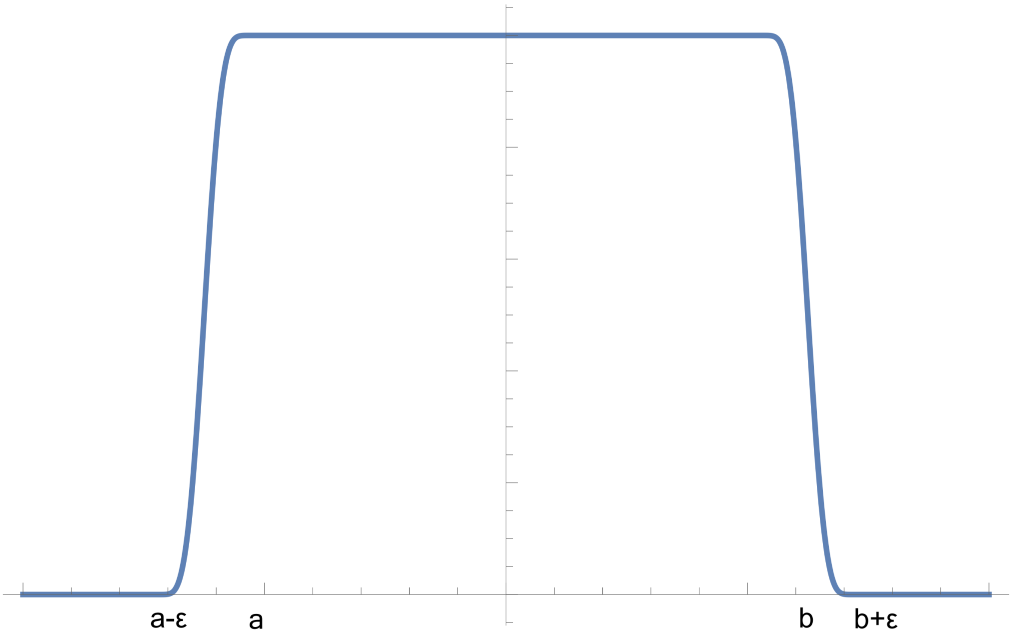

Рис. 29.2. Graph of $\varphi_{a,b,\varepsilon}(x)$.

Such a function $\varphi_{a,b,\varepsilon} \in \mathscr{D}$ exists due to the following reasoning. Consider the derivative $\varphi'_{a,b,\varepsilon}$, see Fig. 29.3.

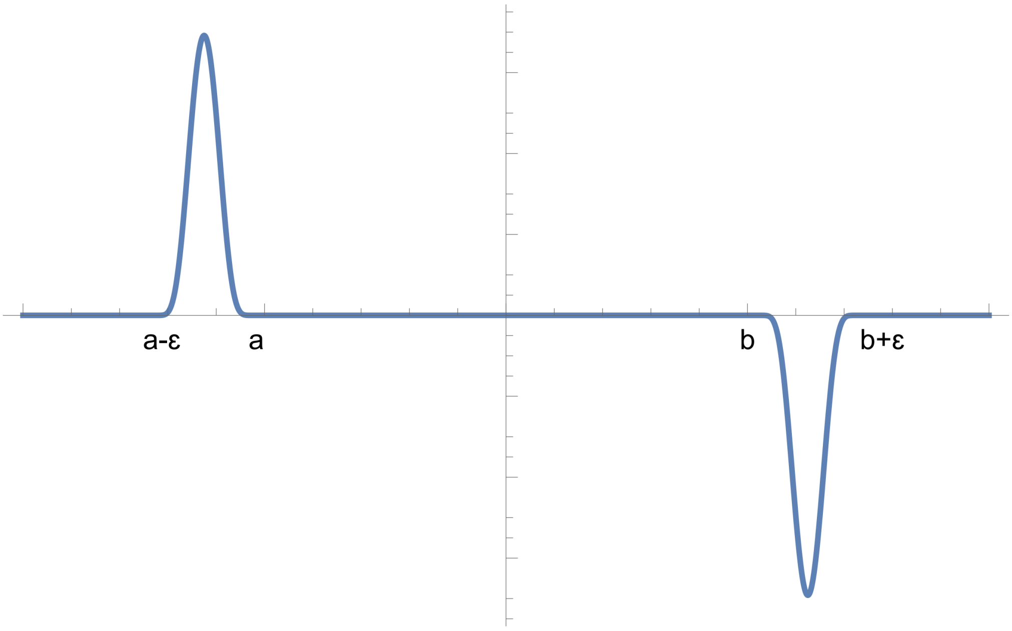

Рис. 29.3. Graph of $\varphi'_{a,b,\varepsilon}(x)$.

It is a linear combination of two hat functions (mentioned previously) that are infinitely differentiable and compactly supported: $\varphi_{a-\varepsilon,a} - \varphi_{b,b+\varepsilon}$. Thus, one can define

$$\varphi_{a,b,\varepsilon}(x) = C\int_{-\infty}^{x} \big(\varphi_{a-\varepsilon,a}(t) - \varphi_{b,b+\varepsilon}(t)\big)\,dt.$$

Next, consider for $a > 0$ and a sufficiently small ε, $0 < a-\varepsilon < a < b < b+\varepsilon$,

$$\langle \delta(x), \varphi_{a,b,\varepsilon}\rangle = 0.$$

By assumption,

$$\langle \delta(x), \varphi_{a,b,\varepsilon}\rangle = \int_{\mathbb{R}} f(x)\varphi_{a,b,\varepsilon}(x)\,dx;$$

since $|\varphi(x)| \leqslant 1$, we have the following bound for the integrand:

$$|f(x)\varphi_{a,b,\varepsilon}(x)| \leqslant |f(x)|,$$

so $f \in L_1[a-\varepsilon_1, b+\varepsilon_1]$ for some fixed ε_1. By Lebesgue's dominated convergence theorem,

$$\int_{\mathbb{R}} f(x)\varphi_{a,b,\varepsilon}(x)\,dx \to \int_a^b f(x)\,dx \quad \text{as} \quad \varepsilon \to 0,$$

therefore,

$$\int_a^x f(t)\,dt = 0 \quad \forall x > a > 0.$$

Differentiating this equality, we obtain

$$f(x) = 0 \quad \forall x > 0.$$

Similarly, one can prove that for $a-\varepsilon < a < b < b+\varepsilon < 0$, we have

$$f(x) = 0 \quad \forall x < 0.$$

Thus, $f(x) = 0$ almost everywhere, so

$$\forall \varphi \in \mathscr{D}: \quad \int_{\mathbb{R}} f(x)\varphi(x)\,dx = 0;$$

however,

$$\langle \delta(x), \varphi_{-1,1,1}\rangle = 1 \neq 0,$$

see the graph of $\varphi_{-1,1,1}$ in Fig. 29.4.

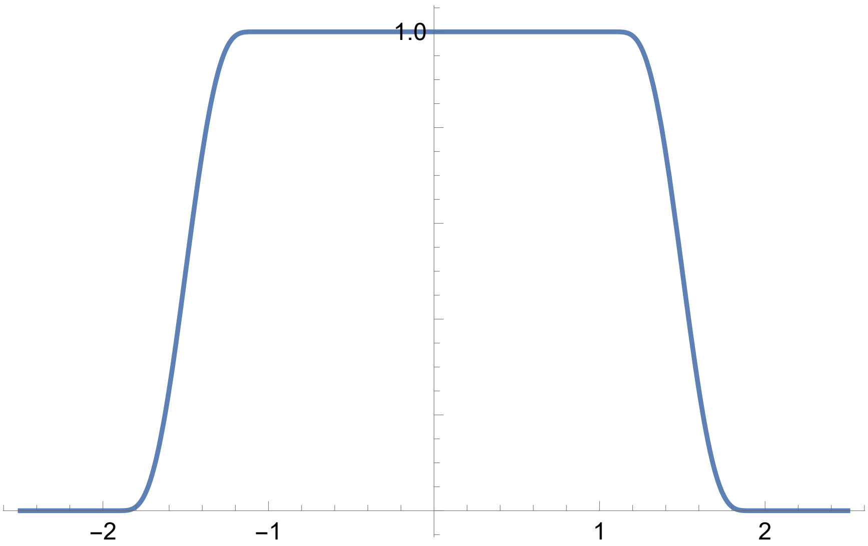

Рис. 29.4. Graph of $\varphi_{-1,1,1}(x)$.

Thus, we arrive at the contradiction. □

Schwartz Space and Tempered Distributions

Now we will consider another space of test functions:

$$\mathcal{S} := \{\varphi \in C^{\infty}(\mathbb{R}) : \ \forall p, q \in \mathbb{N} \cup \{0\} \ \|\varphi\|_{p,q} := \sup_{\mathbb{R}} |x^p \varphi^{(q)}(x)| < \infty\};$$

this is the space of rapidly decaying functions (any derivative decreases faster than $1/x^p$ for any p).

One can see that $\mathscr{D} \subset \mathcal{S}$, since any $\varphi \in \mathscr{D}$ vanishes outside a compact set. For the function

$$\varphi(x) = e^{-x^2} \in \mathcal{S},$$

we have $\varphi \notin \mathscr{D}$.

In the definition above, $\|\varphi\|_{p,q}$ is a family of seminorms, so $\mathcal{S}$ is a locally convex space. One can consider the space of linear continuous functionals on $\mathcal{S}$; the space $\mathcal{S}'$ is called a space of *tempered distributions.*

Theorem 29.3. *The embedding $\mathscr{D} \hookrightarrow \mathcal{S}$ is dense and continuous.*

Proof.

1) It is obvious that $\mathscr{D} \subset \mathcal{S}$.

2) *Continuity.* Suppose that $\varphi \in \mathscr{D}(-a,a)$. To prove the continuity of embedding, we must prove that the operator

$$J : \mathscr{D} \to \mathcal{S}, \quad J\varphi = \varphi$$

is continuous. Consider

$$\|J\varphi\|_{p,q} \equiv \sup_{\mathbb{R}} |x^p \varphi^{(q)}(x)| \leqslant \max_{[-a,a]} |x^p| \cdot \|\varphi\|_{C^q[-a,a]},$$

so $C(a) := \max\limits_{[-a,a]} |x^p|$, which gives the bound with respect to the standard norm that is an admissible seminorm on $\mathscr{D}$.

3) *Density.* Suppose $\varphi \in \mathcal{S}$. Let us consider the functions

$$\varphi_{-1,1,1}(x), \ldots, \varphi_{-n,n,n} = \varphi_{-1,1,1}\left(\frac{x}{n}\right).$$

Consider

$$\varphi_n := \varphi \cdot \varphi_{-n,n,n} \in \mathscr{D}.$$

The *density* means that one can appoximate any function $\varphi \in \mathcal{S}$ with functions $\varphi_n \in \mathscr{D}$ in terms of seminorms in $\mathcal{S}$. Consider $\|\varphi - \varphi_n\|_{p,q}$:

$$\|\varphi - \varphi_n\|_{p,q} = \sup_{\mathbb{R}} \left| x^p \left(\varphi - \varphi \cdot \varphi_{-n,n,n}\right)^{(q)}(x) \right| = \sup_{\mathbb{R}} \left| x^p \left(\varphi(x)\left(1 - \varphi_{-n,n,n}\right)\right)^{(q)}(x) \right|.$$

Let us study the function $\left(\varphi(x)\left(1 - \varphi_{-n,n,n}\right)\right)^{(q)}(x)$:

$$\left(\varphi(x)\left(1 - \varphi_{-n,n,n}(x)\right)\right)^{(q)}(x) = \varphi^{(q)}(x)\left(1 - \varphi_{-n,n,n}(x)\right) - \sum_{j=1}^{q} C_q^j \varphi^{(q-j)}(x) \varphi_{-n,n,n}^{(j)}(x). \tag{29.1}$$

Note that $\varphi_{-n,n,n} \equiv 1$ on $[-n,n]$, see Fig. 29.5.

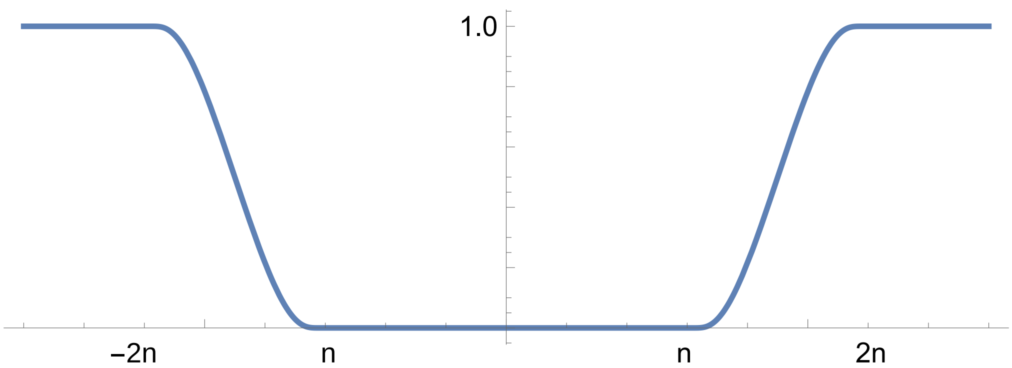

Рис. 29.5. Graph of $1 - \varphi_{-n,n,n}(x)$.

Since $\varphi^{(q)}$ decays rapidly at $|x| \to \infty$, we have

$$\sup_{\mathbb{R}} \left|\varphi^{(q)}\left(1 - \varphi_{-n,n,n}(x)\right)\right| \leqslant \sup_{|x| \geqslant n} |x^p \varphi^{(q)}| \to 0 \quad \text{as} \quad n \to \infty.$$

Consider the terms under the sum in (29.1). The function $\varphi^{(q-j)}$ decays rapidly, and

$$\varphi^{(j)}_{-n,n,n} = \left(\varphi_{-1,1,1}\left(\frac{x}{n}\right)\right)^{(j)} = \frac{1}{n^j}\varphi^{(j)}_{-1,1,1}\left(\frac{x}{n}\right) \to 0 \quad \text{as} \quad n \to \infty$$

due to the chain rule, so every summand has a factor of $1/n^j$ for nonzero j, and the power p is independent of n; therefore, every summand of (29.1) vanishes as $n \to \infty$. □

This theorem obviously implies the following statement.

Corollary 29.1. $\mathscr{D}' \supset \mathcal{S}'$.

Distributions with Compact Support

Consider the following space of test functions:

Definition 29.8. $\mathcal{E}$ *is* $C^\infty(\mathbb{R})$ *with seminorms*

$$p_n(\varphi) = \|\varphi\|_{C^n[-n,n]}.$$

A more general family of seminorms on $\mathcal{E}$ can be defined as follows. Let $\{K_m\}_{m=1}^\infty$ be compact sets such that

$$\cup_m K_m = \mathbb{R}.$$

Then

$$p_{n,K_m} = \|\varphi\|_{C^n(K_m)}.$$

The family of seminorms in the definition above is more convenient due to the fact that it has a single parameter n, while these norms are parametrized by n, m. However, these families are equivalent.

One can see that

$$\mathcal{E} \supset \mathcal{S} \supset \mathscr{D},$$

since in $\mathcal{E}$, all conditions for the decay rate of functions are removed; one can prove that the embedding $\mathcal{S} \hookrightarrow \mathcal{E}$ is dense and continuous.

Definition 29.9. $\mathcal{E}'$ *is the space of linear continuous functions on* $\mathcal{E}$*; this space is called the space of* ***distributions with compact support****.*

For the distributions, the following inclusions hold:

$$\mathcal{E}' \subset \mathcal{S}' \subset \mathcal{D}'.$$

Since distributions are not functions, we must define the support of a distribution.

Definition 29.10. *Let $f \in \mathcal{D}'$ We say that $f \equiv 0$ at a point x if exists a neighborhood $U(x)$ such that*

$$\forall \varphi \in \mathcal{D}, \quad \operatorname{supp} \varphi \subset U(x), \quad \langle f, \varphi \rangle = 0.$$

For instance, $\delta(x) = 0$ for all $x > 0$.

Definition 29.11. *Given $f \in \mathcal{D}'$,*

$$\operatorname{supp} f = \mathbb{R} \backslash \{x : \ f \equiv 0 \ \textit{at } x\}.$$

One can show that $\mathcal{E}'$ is indeed a space of compactly supported distributions, i.e., it consists of $f \in \mathcal{D}'$ with compact $\operatorname{supp} f$.

Lecture 30. Convolution

Example: Green's Function

Consider the equation

$$-f'' + f = \delta(x). \tag{30.1}$$

Recall the notation

$$D = i\frac{d}{dx};$$

in terms of this operator, equation (30.1) can be rewritten as

$$(D^2 + I)f = \delta(x).$$

Taking the Fourier transform, we obtain

$$(M^2 + I)\widehat{f} = \frac{1}{\sqrt{2\pi}}, \quad \text{that is,} \quad (y^2 + 1)\widehat{f}(y) = \frac{1}{\sqrt{2\pi}}.$$

Therefore,

$$\widehat{f}(y) = \frac{1}{\sqrt{2\pi}}\frac{1}{y^2 + 1},$$

so

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \frac{e^{ixy}}{y^2 + 1}\,dy.$$

In Lecture 28, we calculated the direct Fourier transform of $1/(x^2+1)$; the only difference with the inverse one is the sign. Let us repeat the reasoning.

We aim to calculate the integral in terms of residues on the complex plane. For $x > 0$, consider the contour Γ_R as depicted in Fig. 30.1:

$$\frac{1}{2\pi}\int_{\Gamma_R} \frac{e^{ixy}}{y^2 + 1}\,dy.$$

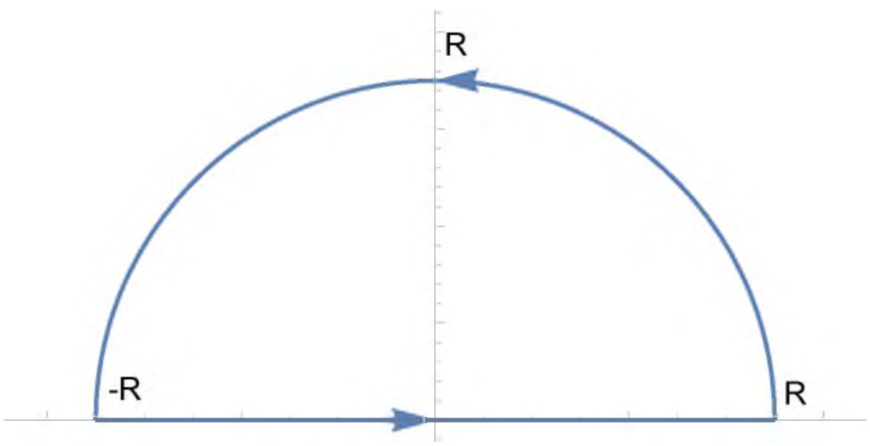

Рис. 30.1. The contour Γ_R.

For $x < 0$, consider the contour in the lower half-plane as depicted in Fig. 30.2.

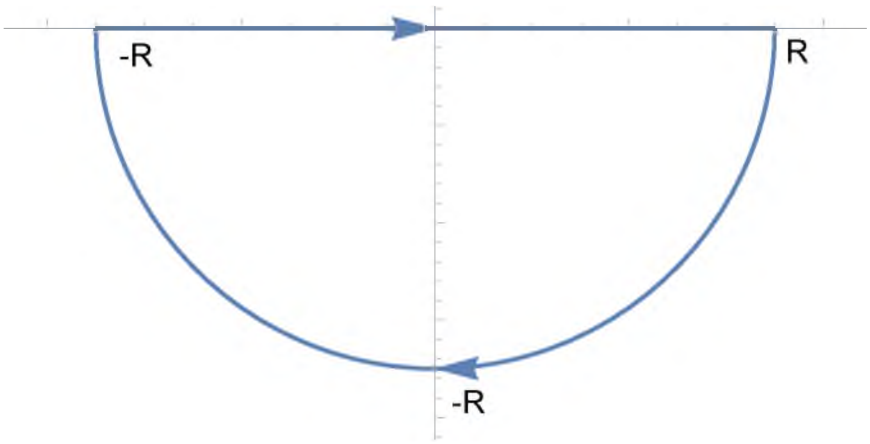

Рис. 30.2. The contour Γ_R in the lower half-plane.

One can see that the value of the integrand on respective semicircles decays exponentially as $R \to \infty$. Therefore, the integral over Γ_R coincides with the integral over the real axis as $R \to \infty$:

$$\frac{1}{2\pi}\int_{\Gamma_\infty}\frac{e^{ixy}}{y^2+1}\,dy = \frac{1}{2\pi}\int_{\mathbb{R}}\frac{e^{ixy}}{y^2+1}\,dy.$$

The same holds for the contour Γ_R in the lower half-plane.

Next, the integrand has two residues, at the points $y = \pm i$. Therefore,

$$\frac{1}{2\pi}\int_{\Gamma_\infty}\frac{e^{ixy}}{y^2+1}\,dy = \frac{1}{2\pi}2\pi i\,\mathrm{Res}_i = i\begin{cases}\dfrac{e^{-x}}{2i}, & x>0,\\ -\dfrac{e^{-x}}{-2i}, & x<0,\end{cases}$$

where an additional minus in the second row is due to the inverse orientation of Γ in the lower half-plane. Thus,

$$f(x) = frac12\pi\int_{\Gamma_\infty}\frac{e^{ixy}}{y^2+1}\,dx = \frac{e^{-|x|}}{2}.$$

This is Green's function for equation (30.1). If we consider the equation

$$-f'' + f = g,$$

then its solution is the convolution with Green's function:

$$f(x) = \frac{1}{2}\int_{-\infty}^{\infty} e^{-|x-t|}g(t)\,dt.$$

Convolution

Definition 30.1. *The **convolution** of f and g at a point x is defined by*

$$(f*g)(x)=\int_{-\infty}^{\infty}f(t)g(x-t)\,dt.$$

One can see that it is well-defined, e.g., in L_2.

Theorem 30.1. *Given f, $g\in L_2(\mathbb{R})$, there exists $f*g\in C_0(\mathbb{R})$.*

Proof of this theorem is a simple exercise that will be considered later.

Let us prove the following statement:

Theorem 30.2. *Given f, $g\in L_1(\mathbb{R})$, there exists $f*g\in L_1(\mathbb{R})$.*

This makes the space $L_1(\mathbb{R})$ an algebra with $*$ as a multiplication.

Proof. One can see that $f(u)g(v)\in L_1(\mathbb{R}^2)$. Thus,

$$\int_{\mathbb{R}}f(u)g(v)\,du\in L_1(\mathbb{R}).$$

Let us substitute $u=t$, $v=x-t$. (In the original integral, we integrate over the lines $v=v_0$; with this substitution, we will integrate over the lines $v=x-u$.) Thus,

$$\int_{\mathbb{R}}f(t)g(x-t)\,dt\in L_1(\mathbb{R}). \qquad \square$$

Properties of Convolution in $L_1(\mathbb{R})$

1) $f*g=g*f$.

   This property is evident.

2) $f*(\alpha g+\beta h)=\alpha f*g+\beta f*h$ for f, g, $h\in L_1$.

   This property follows from the linearity of the integral.

   One can also consider $*$ as a linear operator. Let us fix f. Define the operator of convolution with the function f:

$$S_f g=f*g.$$

   One can see that $\|S_f\|\leqslant\|f\|_{L_1}$.

3) $(f*g)*h=f*(g*h)$.

   This property follows from Fubini's theorem.

4) Let $f \in L_1 \cap C^1$, and $f' \in L_1$. Then

$$\exists (f*g)' = f' * g.$$

5) $\operatorname{supp} f*g = \operatorname{supp} f + \operatorname{supp} g$ meaning that

$$\operatorname{supp} f*g \ni x = t_1 + t_2,$$

where $t_1 \in \operatorname{supp} f$ and $t_2 \in \operatorname{supp} g$.

Since

$$(f*g)(x) = \int_{\mathbb{R}} f(t)g(x-t)\,dt,$$

one can see that

$$x = t + (x-t),$$

and if one of t, $x-t$ is not in the support of f and g respectively, then the corresponding factor vanishes, so is the resulting function.

6) $\widehat{f*g} = \sqrt{2\pi}\widehat{f}\cdot\widehat{g}$.

Consider

$$\widehat{f*g} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\Big(\int_{-\infty}^{\infty} f(t)g(x-t)\,dt\Big)e^{-ixy}\,dx.$$

Using Fubini's theorem, one can switch the integration order:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\Big(\int_{-\infty}^{\infty} f(t)g(x-t)\,dt\Big)e^{-ixy}\,dx = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(t)\Big(\int_{\mathbb{R}} g(x-t)e^{-ixy}\,dx\Big)\,dt$$

and change the variables in the integral with respect to x: $x-t=s$. Then,

$$\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(t)\Big(\int_{\mathbb{R}} g(x-t)e^{-ixy}\,dx\Big)\,dt = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(t)\Big(\int_{\mathbb{R}} g(s)e^{-i(t+s)y}\,ds\Big)\,dt =$$
$$= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(t)e^{-ity}\Big(\int_{\mathbb{R}} g(s)e^{-isy}\,ds\Big)\,dt = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(t)e^{-ity}\,dt\cdot\int_{\mathbb{R}} g(s)e^{-isy}\,ds,$$

which is equal to $\sqrt{2\pi}\widehat{f}\cdot\widehat{g}$.

Convolution of a Distribution with a Test Function

Let $f \in \mathscr{D}'$, $\varphi \in \mathscr{D}$, and $f \in L_1(\mathbb{R})$. Then

$$f*\varphi = \int_{\mathbb{R}} f(t)\varphi(x-t)\,dt = \langle f(t), \varphi(x-t)\rangle.$$

Next, define it $\forall f \in \mathscr{D}'$, $\varphi \in \mathscr{D}$:

Definition 30.2. *Let $f \in \mathscr{D}'$ and $\varphi \in \mathscr{D}$. Define*

$$(f * g)(x) = \langle f(t), \varphi(x-t) \rangle.$$

Since the action of $f \in \mathscr{D}'$ on a function $\varphi \in \mathscr{D}$ gives a constant, and since we have the dependence on the shift with respect to x, we obtain a function in x. Moreover, this function is smooth.

Consider, for example, the convolution with the delta function. By definition,

$$(\delta * \varphi)(x) \langle \delta(t), \varphi(x-t) \rangle;$$

the functional $\delta(t)$ evaluates the function at $t = 0$, thus, we must set t to 0 in this formula:

$$\langle \delta(t), \varphi(x-t) \rangle = \varphi(x).$$

Thus, the convolution with the delta function is the identity operator:

$$S_\delta = I.$$

Next, consider the convolution with the derivative of delta function:

$$(\delta' * \varphi)(x) = \langle \delta'(t), \varphi(x-t) \rangle = -\langle \delta(t), -\varphi(x-t) \rangle = \varphi'(x),$$

so

$$S_{\delta'} = \frac{d}{dx}.$$

This can be employed as follows. Let L be a differential operator:

$$Lf = \sum_{k=0}^{n} a_k f^{(k)} = g.$$

Suppose that we are interested in obtaining a particular solution. Let us write the equation

$$Lf = \delta(x)$$

and solve it in terms of distributions. A solution of this equation is called a **fundamental solution** of a differential operator L, and denoted by $\mathcal{E}(x)$. Then a particular solution can be expressed via

$$f_{par} = \mathcal{E}(x) * g.$$

Why is it a solution? Consider

$$Lf_{par} = L\mathcal{E}(x) * g$$

taking the derivative to the first argument, we get

$$L\mathcal{E}(x) * g = \delta(x) * g = Ig = g.$$

Convolution of Distributions

Let $f,\ g \in \mathscr{D}'$. Can we define $f * g$? In general, the answer is no. However, for some situations it is possible.

Let $f,\ g \in \mathscr{D}$ (note that $\mathscr{D}$ is dense in $\mathscr{D}'$ in weak-$*$ topology). Since $f * g$ is a regular distribution, the action is the integral:

$$\langle f * g, \varphi \rangle = \int \Big(\int f(t) g(x-t)\, dt \Big) \varphi(x)\, dx;$$

due to Fubini's theorem,

$$\int_{\mathbb{R}} \Big(\int_{\mathbb{R}} f(t) g(x-t)\, dt \Big) \varphi(x)\, dx = \int_{\mathbb{R}} f(t) \Big(\int_{\mathbb{R}} g(x-t) \varphi(x)\, dx \Big)\, dt.$$

Substituting $x = t + s$, $dx = ds$, we get

$$\int_{\mathbb{R}} f(t) \Big(\int_{\mathbb{R}} g(x-t) \varphi(x)\, dx \Big)\, dt = \int_{\mathbb{R}} f(t) \Big(\int_{\mathbb{R}} g(s) \varphi(t+s)\, ds \Big)\, dt = \big\langle f(t), \langle g(s), \varphi(t+s) \rangle \big\rangle.$$

One might be tempted to use this as a definition, but it might turn out to be incorrect. In the integral with respect to s, we have the action of $g(s)$ on $\varphi(t+s)$, which gives a function $\psi(t)$. Of course, this function is smooth; however, its support may be noncompact, so it is possible that $\psi \notin \mathscr{D}$.

Due to the properties of the convolution, the support of the resulting function is a sum of supports. Thus, one can guarantee that ψ is compactly supported by imposing the condition that g is a compactly supported distribution. This condition is quite strict.

Another approach is to impose the condition that the supports of g and φ lie in semiaxes so that the intersection of supports is compact.

Define

$$\mathscr{D}_+ \equiv \mathscr{D}(0, +\infty) = \{ \varphi \in \mathscr{D} :\ \operatorname{supp} \varphi \subset [0, +\infty) \}$$

and

$$\mathscr{D}'_+ = \{ f \in \mathscr{D}' :\ \operatorname{supp} f \subset [0, +\infty) \}.$$

Then there exists $f * g \in \mathscr{D}'_+$.

Now, we can define the action of the convolution of two distributions.

Definition 30.3. *Let $\varphi \in \mathscr{D}_+$ and $g \in \mathscr{D}'_+$. Define*

$$\langle f * g, \varphi \rangle := \big\langle f(t), \langle g(s), \varphi(t+s) \rangle \big\rangle.$$

Let us calculate the convolution

$$1 * \delta' * \theta.$$

Let us arragne the brackets as follows: $(1*\delta')*\theta$. Since $S_{\delta'} = \frac{d}{dx}$, we get $1*\delta' = 0$, therefore,

$$(1*\delta')*\theta = 0*\theta = 0.$$

Next, for

$$1*(\delta'*\theta),$$

we have $1*(\delta'*\theta) = 1*\delta = 1$. The problem here is that $1 \notin \mathscr{D}'_+$.

Theorem 30.3. *On $\mathscr{D}'_+$, the convolution is associative.*

Applications of Convolutions of Distributions

Where can we apply the convolution of distributions? Recall the following operator:

$$(Af)(x) = \int_0^x f(t)\,dt.$$

Its nth power is

$$(A^n f)(x) = \int_0^x \frac{(x-t)^{n-1}}{(n-1)!} f(t)\,dt.$$

It looks similar to taking the convolution, however, here we have the integration over $(0,x)$ instead of the entire real axis. Note that for $f \in \mathscr{D}_+$, we have

$$\int_{-\infty}^0 \cdots = 0.$$

Thus, for f of this class, we have the integration over $(0,+\infty)$. Next, to replace the upper limit with x, we can use some standard class of functions in distributions:

$$x_+^\alpha = \begin{cases} x^\alpha, & x > 0, \\ 0, & x < 0, \end{cases}$$

and similarly for x_-^α; one can see that $x_+^\alpha \in \mathscr{D}'_+$. Consider

$$(x-t)_+^\alpha = \begin{cases} (x-t)^\alpha, & x > t, \\ 0, & x < t, \end{cases}$$

so we can integrate over $t < x$. Thus, $(A^n f)(x)$ can be rewritten in terms of convolution:

$$\int_0^x \frac{(x-t)^{n-1}}{(n-1)!} f(t)\,dt = \frac{(x-t)_+^{n-1}}{(n-1)!} * f.$$

Here, we have an integer parameter n, while one can consider $\alpha \in \mathbb{R}$. This allows one to define the fractional derivative.

Let us take $\alpha \in (0,1)$ and consider

$$\psi_\alpha = \frac{x_+^{\alpha-1}}{\Gamma(\alpha)},$$

where Γ denotes the gamma function; recall that $\Gamma(n) = (n-1)!$. Then, the integration of order α can be defined by

$$I_\alpha f := \psi_\alpha * f.$$

For negative α, this defines the fractional derivative.

Example 30.1. *Let us find $I_{1/2}\theta$. For $I_1\theta$, we have*

$$I_1\theta = x_+.$$

By definition,

$$I_{1/2}\theta = \int_0^x \frac{(x-t)^{1/2}}{\Gamma(1/2)}\,dt = -\frac{2}{\sqrt{\pi}}(x-t)^{1/2}\Big|_0^x = \frac{2}{\sqrt{\pi}}x_+^{1/2}.$$

Note that the operator I_α has a group property:

$$I_\alpha I_\beta = I_{\alpha+\beta}$$

in $\mathscr{D}'_+$. One can verify that $I_{1/2}I_{1/2} = I_1$ in a straightforward way.

Exercises

Let us consider the following exercises.

1) Find the convolution $\chi_{[1,2]} * \chi_{[3,4]}$. These functions belong to L_1, so the convolution can be calculated as follows:

$$\int_{\mathbb{R}} \chi_{[1,2]}(t) \int_1^2 \chi_{[3,4]}(x-t)\,dt.$$

Since $\chi_{[3,4]}(x-t) = 1$ for $3 \leqslant x-t \leqslant 4$, i.e., $x-4 \leqslant t \leqslant x-3$, that is, we can rewrite it in the form

$$\chi_{[3,4]}(x-t) = \chi_{[x-4,x-3]}(t).$$

For small x (see Fig. 30.3), the supports do not intersect so the resulting function is 0.

Рис. 30.3. The supports of $\chi_{[1,2]}$ and $\chi_{[x-4,x-3]}$, $x-3<1$.

Next, moving the support of $\chi_{[x-4,x-3]}(t)$ to the right, we obtain a nonzero value of the result for $x-4 \leqslant 1 \leqslant x-3 \leqslant 2$, or simply $4 \leqslant x \leqslant 5$:

$$\int_1^{x-3} dt = x-4.$$

For $1 \leqslant x-4 \leqslant 2$, which simplifies to $5 \leqslant x \leqslant 6$, we get

$$\int_{x-4}^{2} dt = 6-x.$$

Thus, the resulting function is

$$(\chi_{[1,2]} * \chi_{[3,4]})(x) = \begin{cases} x-4, & 4 < x < 5, \\ 6-x, & 5 \leqslant x < 6, \\ 0, & \text{otherwise}, \end{cases}$$

see Fig. 30.4.

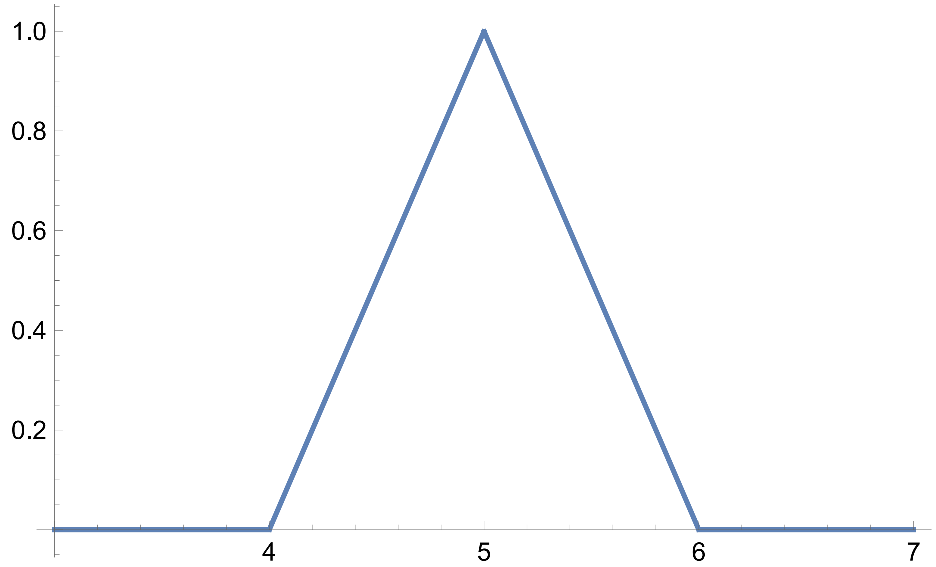

Рис. 30.4. Graph of $(\chi_{[1,2]} * \chi_{[3,4]})(x)$.

Here, one can verify the property of supports: the arguments of $*$ are supported on $[1,2]$ and $[3,4]$, while the result is supported on $[4,6]$.

2) Find $\widehat{\theta}$.

Let us find $\widehat{1/(x+i0)}$ at first. By definition,

$$\left\langle \frac{1}{x+i0}, \varphi \right\rangle = \lim_{\varepsilon\to 0} \int_{\mathbb{R}} \frac{\varphi(x)}{x+i\varepsilon}\, dx.$$

For $\varepsilon \neq 0$, $1/(x+i\varepsilon) \in L_2(\mathbb{R})$. Recall that in L_2, the Fourier transform is unitary. Thus, we can calculate the Fourier transform, and take the limit then. That is, we must consider

$$\widehat{F}\left(\frac{1}{x+i\varepsilon}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{-ixy}}{x+i\varepsilon}\, dx.$$

This integral can be calculated in terms of residues. Let us consider the contour Γ_R as depicted in Fig. 30.5 for $y<0$ and Fig. 30.6 for $y>0$.

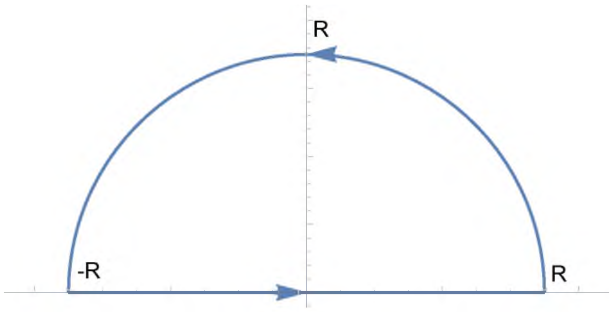

Рис. 30.5. The contour Γ_R for $y<0$.

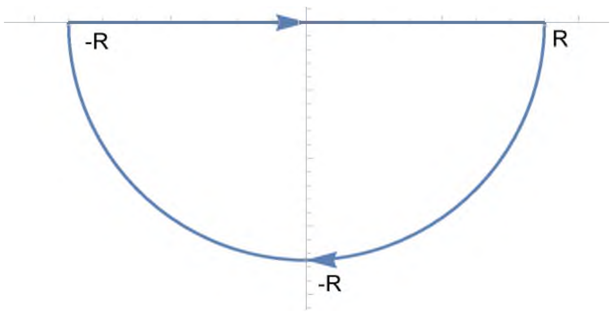

Рис. 30.6. The contour Γ_R for $y>0$.

Since we have the residue at $-i\varepsilon$, there are no residues inside Γ_R for $y<0$; this residue is inside the contour for $y>0$ (note that for Γ_R in the lower half-plane, the orientation is inverse). Therefore,

$$\int_{-\infty}^{\infty} \frac{e^{-ixy}}{x+i\varepsilon}\, dx = 2\pi i \begin{cases} 0, & y<0, \\ -\dfrac{e^{-\varepsilon y}}{1}, & y>0 \end{cases} \xrightarrow{\varepsilon\to 0} 2\pi i \begin{cases} 0, & y<0, \\ -1, & y>0 \end{cases} = -2\pi i \theta(x).$$

Thus,

$$\widehat{\frac{1}{x+i0}} = -i\sqrt{2\pi}\theta(y).$$

In any space where $\widehat{F}\check{F} = I$ (that is, in $\mathcal{S}$, L_2, and $\mathcal{S}'$), we have $\widehat{F}^2 f = f(-x)$. Therefore,

$$-i\sqrt{2\pi}\widehat{\theta(y)} = \widehat{\widehat{\frac{1}{x+i0}}} = \frac{1}{-x+i0},$$

so

$$\widehat{\theta(y)} = \frac{i}{\sqrt{2\pi}}\frac{1}{x-i0}.$$

3) Let $f,\ g \in L_2(\mathbb{R})$. Prove that $f * g \in C_0(\mathbb{R})$.

Consider

$$(f*g)(x) = \int_{\mathbb{R}} f(t)g(x-t)\,dt = \big(f(t), g(x-t)\big).$$

Next, let us use the unitarity of the Fourier transform in L_2:

$$\big(f(t), g(x-t)\big) = \big(\widehat{F}f, \widehat{F}g(x-t)\big).$$

Let us calculate the second factor:

$$\widehat{F}g(x-t) = \int_{\mathbb{R}} g(x-t)e^{-ity}\,dt.$$

Substituting $x-t = s$, $t = x-s$, $dt = -ds$, we get

$$\int_{\mathbb{R}} g(x-t)e^{-ity}\,dt = \int_{\infty}^{-\infty} g(s)e^{i(x-s)y}\,(-ds).$$

Next, we can interchange the integration limits, which gives a minus, and cancel the minus with another one in front of ds:

$$\int_{\infty}^{-\infty} g(s)e^{i(x-s)y}\,(-ds) = e^{ixy}\int_{-\infty}^{\infty} g(s)e^{-isy}\,ds,$$

so

$$\widehat{g(x-t)} = e^{ixy}\widehat{g}.$$

Further, we can rewrite the inner product as

$$\big(\widehat{F}f, \widehat{F}g(x-t)\big) = \int_{\mathbb{R}} \widehat{f}(y)\widehat{g}(y)e^{ixy}\,dy.$$

The product of two L_2-functions $\widehat{f}(y)\widehat{g}(y)$ is a function from $L_1(\mathbb{R})$; therefore, we have the inverse Fourier transform of the function from L_1, so, recalling that

$$\check{F} : L_1 \to C_0,$$

we complete the proof.

Convolution of an L_2-Function with a Distribution from $\mathcal{S}'$

Let $f \in \mathcal{S}'$ such that $\widehat{f} \in L_\infty(\mathbb{R})$, and $g \in L_2(\mathbb{R})$. Let us define the convolution as an operator:

$$f * g = S_f g;$$

the condition $\widehat{f} \in L_\infty(\mathbb{R})$ is important for this operator to be bounded.

Since

$$\widehat{F} : L_2 \to L_2, \qquad \widehat{F} : S' \to S',$$

we have the following diagram:

$$\begin{array}{ccc} L_2 & \xrightarrow{S_f} & L_2 \\ {\scriptstyle \widehat{F}}\big\downarrow & & \big\downarrow{\scriptstyle \widehat{F}} \\ L_2 & \xrightarrow{B} & L_2, \end{array}$$

and this diagram commutes if $\widehat{F} S_f = B\widehat{F}$. For a certain class of functions, the Fourier transform of the convolution is a product of Fourier images; thus, since

$$\widehat{F} S_f = B\widehat{F} g,$$

we have

$$\sqrt{2\pi}\widehat{f}\widehat{g} = B\widehat{g},$$

so S_f is similar to a multiplication operator:

$$Bh = \sqrt{2\pi}\widehat{f}h.$$

Moreover, this immediately implies that the spectrum of S_f consists of essential values of the function $\sqrt{2\pi}\widehat{f}$.

Example 30.2. *Consider the Hilbert transform:*

$$(Af)(x) = \frac{1}{\pi}\,\mathrm{v.p.}\int_{-\infty}^{\infty} \frac{f(t)}{x-t}\,dt.$$

Let us find the spectrum of this operator.

This is a convolution of the following form:

$$\frac{1}{\pi}\,\mathrm{v.p.}\int_{-\infty}^{\infty} \frac{f(t)}{x-t}\,dt = \frac{1}{\pi}\mathcal{P}\frac{1}{x} * f.$$

Thus, we need to find $\widehat{\mathcal{P}(1/x)}$.

For $\mathcal{P}$*, there is the Sokhotski-Plemelj formula:*

$$\frac{1}{x+i0} = \mathcal{P}\frac{1}{x} - i\pi\delta(x).$$

Therefore,

$$\widehat{\frac{1}{x+i0}} = \widehat{\mathcal{P}\frac{1}{x}} - i\pi\widehat{\delta(x)},$$

which gives

$$-i\sqrt{2\pi}\theta = \widehat{\mathcal{P}\frac{1}{x}} - i\sqrt{\frac{\pi}{2}},$$

so

$$\widehat{\mathcal{P}\frac{1}{x}} = i\sqrt{\frac{\pi}{2}} - i\sqrt{2\pi}\theta = i\sqrt{2\pi}\Big(\frac{1}{2}-\theta\Big) = -i\sqrt{\frac{\pi}{2}}\,\mathrm{sgn}\,y.$$

Further, the Hilbert transform A is similar to an operator B:

$$Bh = \frac{1}{\pi}\sqrt{2\pi}\Big(-i\sqrt{\frac{\pi}{2}}\,\mathrm{sgn}\,y\Big)h \equiv -i\,\mathrm{sgn}\,yh.$$

Therefore, $\sigma(B) \equiv \sigma_p(B) = \pm i$; these are eigenvalues of infinite multiplicity. The same holds for A.

Pseudo-Differential Operators

Consider the operator

$$Lf = \sum_{|\alpha|=0}^{m} a_\alpha(x) D^\alpha f, \tag{30.2}$$

where $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multi-index, $|\alpha| = \alpha_1 + \cdots + \alpha_m$, and

$$D^\alpha = D_{x_1}^{\alpha_1} D_{x_2}^{\alpha_2} \ldots D_{x_n}^{\alpha_n}, \quad D_{x_j} = -i\frac{\partial}{\partial x_j}.$$

For D_{x_j}, we have

$$\widehat{F} D_{x_j} = M\widehat{F} \equiv \xi\widehat{F},$$

that is,

$$\widehat{F} D_{x_j} f = \xi\widehat{f}.$$

For (30.2), suppose that $f \in \mathcal{S}$. In this space, we have the inverse Fourier transform:

$$\check{F}\widehat{F} = I.$$

Let us write

$$Lf = L\widehat{F}\check{F}f = \Big(\sum_\alpha a_\alpha(x) D^\alpha \widehat{F}\Big)\check{f}.$$

Next, we can use the commutation formula for D^α and $\widehat{F}$:

$$\Big(\sum_\alpha a_\alpha(x) D^\alpha \widehat{F}\Big)\check{f} = \Big(\sum_\alpha a_\alpha(x)\xi^\alpha\Big)\check{f} = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\sum_\alpha e^{i(x,\xi)} a_\alpha(x)\xi^\alpha f\, d\xi.$$

The function

$$p(x,\xi) = \sum_{\alpha} a_{\alpha}(x)\xi^{\alpha}$$

is called a **symbol** of operator (30.2). Note that we began with considering a differential operator, and arrived at the integral representation with some *symbol* p. For a usual differential operator, the symbol is a polynomial in ξ.

However, one can consider an arbitrary $p(x,\xi)$ as a symbol with the following condition:

$$|p(x,\xi)| < C(1+|\xi|)^m.$$

Then, the resulting operator is called a **pseudo-differential** operator.

Consider the following example. For the Laplacian operator,

$$\Delta f = \sum_{k=1}^{n} \frac{\partial^2 f}{\partial x_k^2},$$

the symbol is $p(x,\xi) = \xi_1^2 + \xi_2^2 + \cdots + \xi_n^2 = |\xi|^2$.

Considering nonpolynomial $p(x,\xi)$, one can obtain, for instance, fractional derivatives and integrals.

МЕХАНИКО-
МАТЕМАТИЧЕСКИЙ
ФАКУЛЬТЕТ
МГУ ИМЕНИ
М.В. ЛОМОНОСОВА

teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ