ANALIZA FUNKCJONALNA I

1. Wstęp.

W kursie algebry rozważane były tylko odwzorowanie (wielo-)liniowe przestrzeni wektorowych. w szczególnosci, mowa była o formach biliniowych. Dla przestrzeni nad $\mathbb{R}$ głównym przykładem był iloczyn skalarny: dodatnia, dwuliniowa forma symetryczne. Przykładem iloczynu skalarnego dla $V = \mathbb{R}^n$ jest forma

$$f(x, y) = x_1 y_1 + \cdots + x_n y_n, \quad x = (x_1, \ldots, x_n), \ \ y = (y_1, \ldots, y_n).$$

Dla przestrzeni wektorowych nad $\mathbb{C}$ iloczyn skalarny φ nie może być funkcją biliniową, jeżeli ma być zachowana interpretacja $\varphi(v, v)$ jako kwadratu normy. W przestrzeni $\mathbb{C}^n$ iloczyn skalarny zadaje się wzorem

$$\varphi(x, y) = \bar{x}_1 y_1 + \cdots + \bar{x}_n y_n, \quad x = (x_1, \ldots, x_n), \ \ y = (y_1, \ldots, y_n).$$

Istnieje więc potrzeba rozpatrywania odwzorowań posiadających własności jak φ ze względu na pierwszą zmienną.

2. Odwzorowania antyliniowe.

Niech V, W będą przestrzeniami wektorowymi nad $\mathbb{C}$. Odwzorowanie $F: V \to W$ nazywamy *antyliniowym*, jeżeli

$$F(\lambda v + \mu w) = \overline{\lambda} F(v) + \overline{\mu} F(w).$$

STWIERDZENIE 1.

1. *Jeżeli $F: V \to W$, $G: W \to U$ są antyliniowe, to $G \circ F$ jest liniowe.*
2. *Jeżeli jedno z odwzorowań $F: V \to W$, $G: W \to U$ jest liniowe, a drugie jest antyliniowe, to $G \circ F$ jest antyliniowe.*

Zbiór odwzorowań antyliniowych z V do W tworzy przestrzeń wektorową nad $\mathbb{C}$, którą oznaczać będziemy $\mathsf{AL}(V, W)$.

Podobnie jak odwzorowania liniowe, odwzorowanie antyliniowe jest wyznaczone przez swoje wartości na wektorach bazy. Istnieje więc reprezentacja macierzowa odwzorowań antyliniowych:

i-ta kolumna macierzy $[F]_e^f$ jest równa $[Fe_i]^f$.

Ponieważ

$$F(\lambda^1 e_1 + \cdots \lambda^n e_n) = \bar{\lambda}^1 F(e_1) + \cdots \bar{\lambda}^n F(e_n),$$

mamy

$$[F(v)]^f = [F]_e^f \overline{[v]}^e.$$

(Znak sprzężenia nad macierzą oznaczą sprzężenie każdego elementu macierzowego.)

Zajmijmy się szczególnym przypadkiem $W = \mathbb{C}$.

Przestrzeń funkcjonałów liniowych $\mathsf{L}(V, \mathbb{C})$ oznaczać będziemy V^*, a przestrzeń $\mathsf{AL}(V, \mathbb{C})$ funkcjonałów antyliniowych $V^\#$. Będziemy też używać oznaczenia $\langle f, v \rangle$ dla ewaluacji $f(v)$. Niech $F: V \to W$ będzie **odwzorowaniem liniowym**. Odwzorowanie

$$W^* \to V^*: f \mapsto f \circ F$$

jest, jak wiadomo, liniowe. Nazywać je będziemy odwzorowaniem *dualnym* lub *sprzężonym* do F i oznaczać będziemy F^*. Podobnie, odwzorowanie

$$W^\# \to V^\#: f \mapsto f \circ F$$

jest liniowe. Nazywać je będziemy odwzorowaniem *antydualnym* lub *hermitowsko sprzężonym* do F i oznaczać będziemy $F^{\#}$.

Jeżeli $F\colon V \to W$ jest **odwzorowaniem antyliniowym**, to wzór $f \mapsto f \circ F$ zadaje odwzorowanie

$$F^*\colon W^* \to V^{\#}\colon f \mapsto f \circ F,$$

które jest, jak łatwo sprawdzić, liniowe:

$$F^*(\lambda f) = \lambda f \circ F$$

Podobnie, odwzorowanie

$$F^{\#}\colon W^{\#} \to V^*\colon f \mapsto f \circ F$$

jest liniowe.

Mając bazę $e = (e_1, \ldots, e_n)$ w przestrzeni V definiujemy bazę dualną $e^* = (e_1^*, \ldots, e_n^*)$ w V^* kładąc $\langle e_i^*, e_j \rangle = \delta_{ij}$. W taki sam sposób dostajemy też bazę dualną $e^{\#}$ w $V^{\#}$. Z kursu algebry wiemy, że

$$[F^*]^{e^*}_{f^*} = ([F]^f_e)^{\mathsf{T}}, \tag{1}$$

a łatwym rachunkiem sprawdzamy, że

$$[F^{\#}]^{e^{\#}}_{f^{\#}} = (\overline{[F]}^f_e)^{\mathsf{T}}, \tag{2}$$

DEFINICJA 2. Sprzężeniem zespolonym *funkcjonału liniowego (antyliniowego)* f *na* V *nazywamy funkcjonał antyliniowy (liniowy)* $\overline{f}$ *na* V*, zdefiniowany przez*

$$\overline{f}(v) = \overline{f(v)}.$$

Sprzężenie zespolone daje kanoniczny antyizomorfizm (t. j. antyliniowy izomorfizm) V^* i $V^{\#}$. W konsekwencji daje antyizomorfizm przestrzeni $\mathsf{L}(V^*, W^*)$ i $\mathsf{L}(V^{\#}, W^{\#})$, zdefiniowany wzorem

$$G \to \overline{G} \ : \ \overline{G}(f) = \overline{G(\overline{f})},$$

gdzie $G \in \mathsf{L}(V^*, W^*)$. Niech $F \in \mathsf{L}(V, W)$. Oczywistym jest związek

$$F^{\#} = \overline{F}^*.$$

Jak wiadomo z algebry, dla przestrzeni wektorowych wymiaru skończonego mamy kanoniczny izomorfizm liniowy V i $(V^*)^*$, zadany wzorem

$$V \ni v \mapsto \varphi_v \in (V^*)^* \ : \ \varphi_v(f) = \langle f, v \rangle \text{ dla } f \in V^*.$$

Liniowość tego izomorfizmu wynika z faktu, że ewaluacja $\langle f, v \rangle$ jest liniowa ze względu na oba argumenty.

Podobnie, ten sam wzór zadaje bijekcję V i $(V^{\#})^*$. Tym razem jednak $\langle f, v \rangle$ jest liniowe ze względu na f i antyliniowe ze względu na v, więc bijekcja jest izomorfizmem antyliniowym. Mówiąc inaczej, struktura przestrzeni wektorowej nad $\mathbb{C}$, indukowana na V z $(V^{\#})^*$ różni się od wyjściowej. Struktura indukowana różni się mnożeniem przez liczbę: $\lambda \circ v = \bar{\lambda} v$. Zbiór V z taką strukturą przestrzeni wektorowej oznacza się $\overline{V}$.

Pozostaje jeszcze zauważyć, że odwzorowanie

$$V \ni v \mapsto \overline{\varphi}_v \in (V^{\#})^{\#} \ : \ \overline{\varphi}_v(f) = \overline{\langle f, v \rangle} \text{ dla } f \in V^{\#},$$

jest izomorfizmem liniowym.

W przypadku wymiaru nieskończonego wzory powyższe dają injekcje, ale nie surjekcje, przestrzeni V w $(V^*)^*$, $(V^{\#})^*$ i $(V^{\#})^{\#}$.

Podsumowując:

STWIERDZENIE 3.

(1) *Istnieje kanoniczny izomorfizm przestrzeni* $(V^\#)^*$ *i* $(V^*)^\#$.
(2) *Istnieje kanoniczna injekcja antyliniowa przestrzeni* V *w* $(V^*)^\#$.
(3) *Istnieje kanoniczna injekcja antyliniowa przestrzeni* V *w* $(V^\#)^*$.
(4) *Istnieje kanoniczna injekcja liniowa przestrzeni* V *w* $(V^\#)^\#$.

DOWÓD:

(1) Niech $\varphi \in (V^*)^\#$. Wzór

$$V^\# \ni f \mapsto \overline{\varphi}(\overline{f}) \in \mathbb{C}$$

określa, jak łatwo sprawdzić, odwzorowanie liniowe, więc element z $(V^\#)^*$. Oznaczmy go przez $\widetilde{\varphi}$. Przyporządkowanie $\varphi \mapsto \widetilde{\varphi}$ jest izomorfizmem przestrzeni wektorowych. (Co się zmieni, jeżeli w powyższym wzorze $\overline{\varphi}$ zastąpimy przez φ?)
(2) Sprzężenie zespolone daje antyizomorfizm $(V^*)^*$ i $(V^*)^\#$. Sprzężenie zespolone złożone z kanonicznym włożeniem V w $(V^*)^*$ daje żądane włożenie antyliniowe.
(3) Otrzymujemy przez złożenie włożenia z punktu drugiego z izomorfizmem z punktu pierwszego.
(4) Otrzymujemy przez złożenie włożenia z punktu poprzedniego ze sprzężeniem zespolonym. ■

Jako proste ćwiczenie zostawiamy dowód następującego stwierdzenia, będące odpowiednikiem znanego z algebry twierdzenia o odwzorowaniach sprzężonych.

STWIERDZENIE 4. *Dla* F, G *liniowych (antyliniowych) mamy związki*

(1) $(F+G)^\# = F^\# + G^\#, \ \ (F+G)^* = F^* + G^*,$
(2) $(\lambda F)^\# = \overline{\lambda} F^\#, \ \ (\lambda F)^* = \lambda F^*$
(3) $(G \circ F)^\# = F^\# \circ G^\#, \ \ (G \circ F)^* = F^* \circ G^*$
(4) $(F^\#)^\# \supset F, \ \ (F^*)^* \supset F.$

W przypadku skończenie-wymiarowym inkluzje stają się równościami.

Niech teraz $F\colon V \to V^*$ będzie odwzorowaniem liniowym. Odwzorowanie sprzężone $F^*\colon (V^*)^* \supset V \to V^*$ jest scharakteryzowane związkiem

$$\langle F^*(v), w\rangle = \varphi_v \circ F(w) = \langle F(w), v\rangle \tag{3}$$

i podobnie

$$\langle F^\#(v), w\rangle = \overline{\varphi_v \circ F(w)} = \overline{\langle F(w), v\rangle}. \tag{4}$$

Odwzorowanie liniowe $F\colon V \to V^*$ nazywamy *symetrycznym*, jeżeli $F = F^*$ na V, czyli

$$\langle F(v), w\rangle = \varphi_v \circ F(w) = \langle F(w), v\rangle.$$

Odwzorowanie liniowe $F\colon V \to V^\#$ nazywamy *hermitowsko symetrycznym*, jeżeli $F^\# = F$ na V, czyli

$$\langle F(v), w\rangle = \varphi_v \circ F(w) = \overline{\langle F(w), v\rangle}.$$

Podobnie, jeżeli $F\colon V \to V^*$ jest odwzorowaniem antyliniowym, to $\overline{F}\colon V \to V^\#$ jest odwzorowaniem liniowym. Odwzorowanie sprzężone $\overline{F}^\#\colon (V^\#)^\# \supset V \to V^*$ jest więc scharakteryzowane związkiem, zgodnie z (4),

$$\langle \overline{F}^\#(v), w\rangle = \overline{\varphi_v \circ \overline{F}(w)} = \overline{\langle \overline{F}(w), v\rangle} = \langle F(w), v\rangle.$$

Równość $\overline{F}^\# = F$ na V oznacza

$$\langle F(v), w\rangle = \langle F(w), v\rangle.$$

Jeżeli natomiast $F: V \to V^{\#}$ jest odwzorowaniem antyliniowym, to $\overline{F}: V \to V^*$ jest odwzorowaniem liniowym. Odwzorowanie sprzężone $\overline{F}^*: (V^*)^* \supset V \to V^*$ jest scharakteryzowane, zgodnie z (3), związkiem

$$\langle \overline{F}^*(v), w \rangle = \varphi_v \circ \overline{F}(w) = \langle \overline{F}(w), v \rangle = \overline{\langle F(w), v \rangle}.$$

Równość $\overline{F}^* = F$ na V oznacza

$$\langle F(v), w \rangle = \overline{\langle F(w), v \rangle}.$$

3. Przestrzenie z iloczynem skalarnym.

DEFINICJA 5. Iloczynem skalarnym *w przestrzeni wektorowej V nad $\mathbb{C}$ nazywamy odwzorowanie:*

$$h: V \times V \to \mathbb{C}$$

takie, że dla $v, w \in V$ i $\alpha \in \mathbb{C}$ mamy

(1) $h(v, \alpha w) = \alpha h(v, w)$,
(2) $h(v, w + w') = h(v, w) + h(v, w')$,
(3) $h(v, w) = \overline{h(w, v)}$,
(4) $h(v, v) > 0$ *dla* $v \neq 0$ $\;(h(v, v) \in \mathbb{R})$.

Często można spotkać inny zestaw aksjomatów iloczynu skalarnego:

(1') $h(v, \alpha w) = \alpha h(v, w)$, $h(\alpha v, w) = \bar{\alpha} h(v, w)$,
(2') $h(v, w + w') = h(v, w) + h(v, w')$, $h(v + v', w) = h(v, w) + h(v', w)$,
(3') $h(v, v) \in \mathbb{R}$,
(4') $h(v, v) > 0$ dla $v \neq 0$.

Oczywistym jest, że (1-4) pociąga za sobą (1'-4'). W obu też przypadkach h jest odwzorowaniem liniowym ze względu na drugą, a antyliniowym ze względu na pierwszą zmienną. Mówimy, że h jest formą półtoraliniową.

Pokażemy teraz równoważność obu definicji.

Wyprowadzimy najpierw formułę polaryzacyjną dla odwzorowania półtoraliniowego. Mamy

$$h(v + w, v + w) = h(v, v) + h(w, w) + h(v, w) + h(w, v)$$

oraz

$$h(\iota v + w, \iota v + w) = h(v, v) + h(w, w) - \iota h(v, w) + \iota h(w, v),$$

a stąd, mnożąc drugą tożsamość przez ι i dodając do pierwszej, dostajemy (korzystając z półtoraliniowości h)

$$h(v, w) = \frac{1}{2}(h(v+w, v+w) - h(v, v) - h(w, w)) + \iota \frac{1}{2}(h(\iota v+w, \iota v+w) - h(v, v) - h(w, w)). \quad (5)$$

Podobnie, mnożąc drugą tożsamość przez $-\iota$ i dodając do pierwszej, dostajemy

$$h(w, v) = \frac{1}{2}(h(v+w, v+w) - h(v, v) - h(w, w)) - \iota \frac{1}{2}(h(\iota v+w, \iota v+w) - h(v, v) - h(w, w)). \quad (6)$$

Z warunku (3') wynika więc równość (3).

Spostrzeżenie: $\operatorname{Im} h$ jest formą $\mathbb{R}$-dwuliniową, antysymetryczną i niezdegenerowaną. $\operatorname{Re} h$ jest formą $\mathbb{R}$-dwuliniową, symetryczną i niezdegenerowaną, więc zadaje na V strukturę przestrzeni unormowanej. Możemy zatem stosować oznaczenia i konstrukcje właściwe przestrzeniom z normą. Przede wszystkim, jest to przestrzeń metryczna, więc topologiczna.

Ponadto,

(1) $h(v, w)$ oznaczamy $(v \mid w)$
(2) $(h(v, v))^{\frac{1}{2}}$ będziemy oznaczać $\|v\|$ i nazywać *normą* (*długością*) wektora v.

Jak i w przypadku euklidesowym, mamy równość równoległoboku

$$\|v+w\|^2+\|v-w\|^2=2(\|v\|^2+\|w\|^2)$$

DOWÓD:

$$\|v+w\|^2+\|v-w\|^2=(v+w|v+w)-(v-w|v-w)=2(v|v)+2(w|w)$$

■

Przykłady:

(1) $\mathbb{C}^n$ z iloczynem skalarnym

$$(x \mid y)=\overline{x_1}y_1+\cdots+\overline{x_n}y_n.$$

(2) Przestrzeń funkcji ciągłych na odcinku $[a,b]$ z iloczynem skalarnym

$$(f \mid g)=\int_a^b \bar{f}g\ .$$

(3) Przestrzeń $C([a,b])$ funkcji różniczkowalnych w sposób ciągły z iloczynem skalarnym

$$(f \mid g)=\int_a^b \bar{f}g+\int_a^b \bar{f}'g'\ ,$$

gdzie 'prim' oznacza pochodną.

Przestrzenie z drugiego i trzeciego przykładu są wymiaru nieskończonego. Jako przestrzenie metryczne nie są zupełne i można je, metodą standardową, uzupełnić. Otrzymane w ten sposób przestrzenie oznaczamy $\mathsf{H}^0([a,b])$ i $\mathsf{H}^1([a,b])$.

3.1. Podstawowe własności iloczynu skalarnego.

TWIERDZENIE 6 (NIERÓWNOŚĆ SCHWARZA). *Jeśli V jest przestrzenią wektorową z iloczynem skalarnym, to*

$$|\ (v \mid w)\ |\leqslant \|v\|\ \|w\|. \tag{7}$$

Równość zachodzi wtedy i tylko wtedy, gdy v i w są liniowo zależne.

DOWÓD: Jeśli $v=\mathbf{0}$, to twierdzenie jest trywialne.
Jeśli $v\neq\mathbf{0}$, to rozpatrzmy funkcję $\alpha\colon\mathbb{R}\ni t\mapsto\|tv+w\|^2\in\mathbb{R}$, czyli

$$\alpha(t)=t^2(v \mid v)+2t\,\mathrm{Re}(v \mid w)+(v \mid w).$$

Oczywiście $\alpha(t)\geqslant 0$, zatem wyróżnik tego trójmianu jest niedodatni, tzn.,

$$(\mathrm{Re}(v \mid w))^2-(\|v\|\ \|w\|)^2\leqslant 0. \tag{8}$$

Dla pewnego rzeczywistego φ mamy

$$(v \mid w)=e^{\iota\varphi}|(v \mid w)|,$$

zatem

$$|\ (v \mid w)\ |=e^{-\iota\varphi}(v \mid w)=(e^{\iota\varphi}v \mid w)=\mathrm{Re}(e^{\iota\varphi}v \mid w).$$

Stąd i z (8) dostajemy

$$|\ (v \mid w)\ |^2\leqslant\|e^{i\varphi}v\|^2\ \|w\|^2=\|v\|^2\|w\|^2.$$

Jeżeli $w = \lambda v$, to
$$|(v \mid w)| = |\lambda| \|v\|^2 = \|\lambda v\| \, \|v\| = \|v\| \, \|w\|.$$
Niech teraz $|(v \mid w)| = \|v\| \|w\|$ i $|(v \mid w)| = e^{-\iota\varphi}(v \mid w)$. Rozpatrzmy funkcję

$$\begin{aligned}\beta \colon t \mapsto \beta(t) = \|e^{\iota\varphi} tv + w\|^2 = t^2\|v\|^2 + 2t \mid (v \mid w) \mid + \|w\|^2 = \\ = t^2\|v\|^2 + 2t\|v\|\|w\| + \|w\|^2 = (t\|v\| + \|w\|)^2.\end{aligned}$$

β jest równe zero dla $t_0 = -\frac{\|w\|}{\|v\|}$, czyli $w = (e^{\iota\varphi}\frac{\|w\|}{\|v\|})v$. ■

Bezpośrednim wnioskiem z nierówności Schwarza jest nierówność trójkąta:

$$\|v + w\|^2 = \|v\|^2 + 2\operatorname{Re}(v|w) + \|w\|^2 \leqslant \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2.$$

Zatem $\| \, \|$ spełnia aksjomaty normy w przestrzeni wektorowej. Z nierówności Schwarza wynika też, że iloczyn skalarny jest funkcją ciągła na $V \times V$ z normą produktową. Przy okazji zauważmy, że w iloczynie kartezjańskim $V \times W$ przestrzeni z iloczynami skalarnymi, odpowiednio $(\,|\,)_V$ i $(\,|\,)_W$ możemy zdefiniować iloczyn skalarny wzorem

$$((v, w)|(v', w')) = (v|v')_V + (w|w')_W.$$

Ponieważ iloczyn skalarny h jest formą półtoraliniową, odpowiadają mu dwa odwzorowania $F_h \colon V \to V^*$ i ${}_hF \colon V \to V^\#$ zdefiniowane przez

$$\begin{aligned}\langle F_h(v), w\rangle &= (v \mid w) \\ \langle {}_hF(v), w\rangle &= (w \mid v).\end{aligned} \tag{9}$$

F_h jest odwzorowaniem antyliniowym:

$$\langle F_h(\lambda v), w\rangle = (\lambda v \mid w) = \bar{\lambda}\langle F_h(v), w\rangle = \langle \bar{\lambda} F_h(v), w\rangle,$$

natomiast odwzorowanie ${}_hF$ jest liniowe i, jak łatwo zauważyć, hermitowsko symetryczne. Ponadto
$$\langle {}_hF(v), w\rangle = (w \mid v) = \overline{(v \mid w)} = \overline{\langle F_h(v), w\rangle} = \langle \overline{F}_h(v), w\rangle,$$
czyli ${}_hF = \overline{F}_h$. Jeżeli $v \in \ker F_h$, to $F_h(v) = 0$ i

$$0 = \langle F_h(v), v\rangle = h(v, v) = \|v\|^2.$$

Stąd $v = \mathbf{0}$, czyli ${}_hF$ jest injekcją. W przypadku przestrzeni wymiaru skończonego oznacza to, że ${}_hF$ jest izomorfizmem. Nie jest tak dla wymiaru nieskończonego, ale, jak zobaczymy w dalszym ciągu wykładu, po ograniczeniu się do funkcjonałów ciągłych i spełnieniu warunku zupełności przestrzeni, znów dostajemy izomorfizm.

Dla dowolnego podzbioru $A \subset V$ przestrzeni z iloczynem skalarnym definiujemy zbiór

$$A^\perp = \{w \in V \colon \forall v \in A, (w|v) = 0\}.$$

Oczywistym jest, że $A^\perp$ jest domkniętą podprzestrzenią wektorową przestrzeni V. Mają też miejsce następujące oczywiste relacje

(1) Jeżeli $A \subset B$, to $A^\perp \supset B^\perp$.
(2) $(A^\perp)^\perp$ zawiera najmniejszą domkniętą podprzestrzeń zawierającą A.

Zauważmy jeszcze, że $A^\perp = F_h^{-1}(A^\circ)$, gdzie $A^\circ = \{f \in V^* \colon f(v) = 0 \ \forall v \in A\}$. Możemy tu zastąpić F_h przez ${}_hF$ i V^* przez $V^\#$.

4. Przestrzenie Hilberta.

Zupełna przestrzeń z normą jest przestrzenią Banacha. Zupełną przestrzeń z iloczynem skalarnym nazywamy *przestrzenią Hilberta*. Przestrzeń Hilberta jest przestrzenią Banacha, więc można korzystać z rachunku różniczkowego i innych metod omawianych w kursie Analizy. Warto zwrócić uwagę na to, że nie każda norma jest równoważna normie hilbertowskiej, tzn. pochodzącej od iloczynu skalarnego. Oczywiście, jest to możliwe dla wymiaru nieskończonego. W przypadku wymiaru skończonego, wszystkie normy są równoważne. Przypomnę tu, że dwie normy $\|\cdot\|$ i $\|\cdot\|'$ na przestrzeni wektorowej V są równoważne, jeżeli istnieją liczby dodatnie a, b takie, że dla każdego $v \in V$ zachodzi nierówność $a\|v\| \leqslant \|v\|' \leqslant b\|v\|$.

TWIERDZENIE 7 (O RZUCIE PROSTOPADŁYM). *Niech* H *będzie przestrzenią Hilberta z iloczynem skalarnym* h *i niech* $W \subset \mathsf{H}$ *będzie domkniętą podprzestrzenią wektorową. Wówczas*

$$\mathsf{H} = W \oplus W^{\perp}$$

(w sensie algebraicznym i topologicznym).

DOWÓD: Jeżeli $v \in W \cap W^{\perp}$, to $(v \mid v) = 0$, czyli $v = 0$.
Czy $\mathsf{H} = W + W^{\perp}$?
Rozpatrzmy funkcję

$$\mathsf{H} \ni v \mapsto \alpha(v) = \inf_{y \in W} \|v - y\|^2.$$

Z definicji inf istnieje ciąg $(y_n \in W)$ taki, że $\alpha(v) = \lim_{n\to\infty} \|v - y_n\|^2$. Zatem możemy przyjąć, że

$$\alpha(v) \leqslant \|v - y_n\|^2 \leqslant \alpha(v) + \frac{1}{n}.$$

Stąd i z równości równoległoboku

$$\begin{aligned} \|y_n - y_m\|^2 &= \|(v - y_m) - (v - y_n)\|^2 = \\ &= 2(\|v - y_m\|^2 + \|v - y_n\|^2) - \|2v - (y_m + y_n)\|^2 \\ &\leqslant 4\alpha + 2(\frac{1}{n} + \frac{1}{m}) - 4\alpha = 2(\frac{1}{n} + \frac{1}{m}) \to 0 \end{aligned}$$

Zatem ciąg y_n jest ciągiem Cauchy'ego. Ponieważ przestrzeń H jest zupełna, a podprzestrzeń W jest domknięta jego granica $P(v)$ istnieje, należy do W i $\alpha(v) = \|v - P(v)\|$.

Pokażemy, że $v - P(v) \in W^{\perp}$. Niech $y \in W$, wówczas funkcja

$$f(t) \mapsto \|(v - P(v)) + ty\|^2$$

ma minimum w $t = 0$. Z drugiej strony

$$f(t) = \|v - P(v)\|^2 + 2t \ \mathrm{Re}(y \mid v - P(v)) + t^2\|y\|^2,$$

czyli

$$0 = f'\,(0) = 2\,\mathrm{Re}(y \mid v - P(v)).$$

Ponadto

$$\mathrm{Im}(y \mid v - P(v)) = \mathrm{Re}(\imath y \mid v - P(v)) = 0,$$

bo $\imath y \in W$. Oznacza to, że $v - P(v) \in W^{\perp}$. Rozkład $v = P(v) + (v - P(v))$ jest rozkładem w sumie prostej (algebraicznej!) $\mathsf{H} = W \oplus W^{\perp}$. Trzeba jeszcze pokazać, że suma ta jest sumą topologiczną. Jest to natychmiastowy wniosek z poniższego rachunku

$$\|v\|^2 = (v|v) = (P(v)|P(v)) + (v - P(v)|v - P(v)) = \|P(v)\|^2 + \|v - P(v)\|^2.$$

■

Odwzorowanie P zdefiniowane w dowodzie nazywamy rzutem ortogonalnym na podprzestrzeń W.

Łatwym wnioskiem z twierdzenia o rzucie prostopadłym jet twierdzenie o reprezentacji funkcjonału.

TWIERDZENIE 8 (Riesza o reprezentacji funkcjonału liniowego i ciągłego). *Dla każdego funkcjonału liniowego i ciągłego f na przestrzeni Hilberta H istnieje dokładnie jeden wektor $w_f \in H$ taki, że dla każdego wektora $v \in \mathsf{H}$ zachodzi równość*

$$f(v) = (w_f \mid v).$$

Dowód: Niech $W = \ker f$. Jest to domknięta podprzestrzeń liniowa H. Domkniętość wynika z ciągłości f. Na mocy twierdzenia o rzucie prostopadłym $H = W \oplus W^\perp$. Pokażemy, że $W^\perp$ ma wymiar 1. Niech bowiem $v, w \in W^\perp$, to $f(v)w - f(w)v \in W^\perp$ i $f(f(v)w - f(w)v) = 0$, więc $f(v)w - f(w)v \in W$. Stąd $f(v)w - f(w)v = 0$ co oznacza, że v, w są liniowo zależne. Zatem istnieje dokładnie jeden $v_0 \in W^\perp$ taki, że $\|v_0\| = 1$ i $\mathbb{R} \ni f(v_0) > 0$. Połóżmy $w_f = f(v_0)v_0$. Mamy $f(w_f) = (f(v_0))^2 = (w_f \mid w_f)$ i stąd $f(v) = (w_f \,|\, v)$ dla każdego $v \in \mathsf{H}$, bo rozkładając $v = \lambda w_f + w$, gdzie $w \in W$, dostajemy

$$f(v) = f(\lambda w_f + w) = \lambda f(w_f) = \lambda(w_f \mid w_f) = (w_f \mid v).$$

■

Uwaga: OD TEGO MIEJSCA H^* ($\mathsf{H}^\#$) oznaczać będzie przestrzeń funkcjonałów liniowych (antyliniowych) i CIĄGŁYCH na H.

Przestrzenie H^* i $\mathsf{H}^\#$ są przestrzeniami Banacha (przestrzeń odwzorowań liniowych i ciągłych między przestrzeniami Banacha jest przestrzenią Banacha). Twierdzenie Riesza oznacza, że odwzorowania $F_h \colon \mathsf{H} \to \mathsf{H}^*$ i ${}_hF \colon \mathsf{H} \to \mathsf{H}^\#$ (9) są surjektywne. Ponieważ są też injektywne, są izomorfizmami (F_h antyliniowym, a ${}_hF$ liniowym) przestrzeni wektorowych. Pokażemy, że są izometriami, czyli izomorfizmami przestrzeni Banacha.

TWIERDZENIE 9. *Odwzorowania ${}_hF$ i F_h są izometriami (zachowują normę).*

Dowód: Niech $f = F_h(w_f)$. Mamy

$$\|f\| = \sup_{\|v\|=1} |f(v)| = \sup_{\|v\|=1} |(w_f|v)| = \|w_f\|,$$

bo z nierówności Schwarza $|(w_f|v)| \leqslant \|w_f\|$. Analogicznie pokazujemy, że ${}_hF$ jest izometrią.

■

Ponieważ ${}_hF$ i F_h są izometriami, można przy ich pomocy wyposażyć przestrzenie H^* i $\mathsf{H}^\#$ w iloczyn skalarny, zgodny z zastaną normą:

$$(F_h(v) \,|\, F_h(w)) = (w \,|\, v)$$
$$({}_hF(v) \,|\, {}_hF(w)) = (v \,|\, w).$$

Przykład Przestrzeń ℓ^2 ciągów liczb zespolonych, sumowalnych z kwadratem, tzn. $(a_i) \in \ell^2$ jeżeli $\sum |a_i|^2 < \infty$, z iloczynem skalarnym

$$((a_i)|(b_i)) = \sum \bar{a}_i b_i,$$

jest przestrzenią Hilberta. Wykażemy najpierw, że szereg $\sum \bar{a}_i b_i$ jest zbieżny bezwzględnie: z nierówności Schwarza w $\mathbb{R}^N$

$$\sum_1^N |a_i b_i| \leqslant \sqrt{\sum_1^N |a_i|^2}\sqrt{\sum_1^N |b_i|^2} \leqslant \sqrt{\sum_1^\infty |a_i|^2}\sqrt{\sum_1^\infty |b_i|^2} < \infty,$$

czyli szereg definiujący iloczyn skalarny jest zbieżny. Trzeba teraz wykazać zupełność przestrzeni ℓ^2. Niech ciąg $\bar{a}_n$ będzie ciągiem Cauchy'ego w ℓ^2, $\bar{a}_n = (a_{n,i})$. Dla każdego i ciąg $(a_{n,i})$ jest ciągiem Cauchy'ego w $\mathbb{C}$, więc zbieżnym do a_i. Niech $\bar{a} = (a_i)$. Pokażemy, że $\bar{a} \in \ell^2$ i że $\bar{a}_n \to \bar{a}$ w ℓ^2.

5. Bazy w przestrzeni Hilberta.

Niech V będzie przestrzenią wektorową. Bazą w V nazywamy zbiór $B \subset V$ taki, że każdy jego skończony podzbiór jest liniowo niezależny i każdy wektor z V jest skończoną kombinacją liniową wektorów bazy. W przypadku przestrzeni wektorowej z topologią, możemy zmodyfikować pojęcie bazy żądając, by skończone kombinacje liniowe tworzyły zbiór gęsty w V. Nie jest to jednak wystarczające, i w przypadku ogólnej przestrzeni Banacha pojęcie bazy jest dość skomplikowane. W przestrzeni z iloczynem skalarnym możemy mówić o bazie ortogonalnej i bazie ortonormalnej, co znacznie upraszcza problem definicji.

DEFINICJA 10. *Bazą ortonormalną* w przestrzeni Hilberta H nazywamy zbiór $B \subset \mathsf{H}$ taki, że

(1) każdy skończony podzbiór B jest ortonormalny,
(2) zbiór skończonych kombinacji liniowych wektorów z B (powłoka liniowa B) jest gęsty w H.

Warunek drugi można zapisać tak: najmniejszą domkniętą podprzestrzenią zawierającą B jest całe H.

STWIERDZENIE 11. *Niech B i B' będą bazami ortonormalnymi w przestrzeni Hilberta H. Wówczas B i B' są zbiorami równolicznymi.*

DOWÓD: W przypadku wymiaru skończonego jest to fakt znany z algebry. Niech więc zbiory B i B' będą co najmniej $\aleph_0$. Dla skończonego podzbioru $\{e_1, \dots, e_k\} \subset B$ niech W będzie podprzestrzenią rozpiętą przez ten podzbiór. Z twierdzenia o rzucie ortogonalnym, dla każdego $b \in B'$

$$\|b\| = \|Pb\| + \|b - Pb\| \geqslant \|Pb\|,$$

gdzie $P\colon \mathsf{H} \to W$ jest rzutem ortogonalnym. Ponieważ układ wektorów $\{e_1, \dots, e_k\}$ jest ortonormalny,

$$Pb = (e_1|b)e_1 + \cdots + (e_k|b)e_k \quad \text{i} \quad \|Pb\|^2 = |(e_1|b)|^2 + \cdots + |(e_k|b)|^2 \leqslant \|b\| = 1.$$

Wynika stąd, że liczba wektorów z $e \in B$, dla których $|(e|b)| > \dfrac{1}{n}$ jest skończona dla każdego n. Zatem zbiór $e \in B\colon (e|b) \neq 0$ jest przeliczalny. W iloczynie kartezjańskim $B \times B'$ definiujemy podzbiór $A = \{(e, b) \in B \times B'\colon (e|b) \neq 0\}$. Jest on mocy nie większej niż moc zbioru $\mathbb{N} \times B$ i nie mniejszej niż moc zbioru B (dla każdego $b \in B$ istnieje conajmniej jeden wektor $b' \in B'$ taki, że $(b|b') \neq 0$). Zbiór B jest nieskończony, więc moc $\mathbb{N} \times B$ jest równa mocy B. Wnioskujemy stąd, że moc A jest równa mocy B. Zamieniając rolami B i B' dostajemy tezę. ■

Mówimy, że przestrzeń Hilberta jest *ośrodkowa*, jeżeli ma przeliczalną bazę ortonormalną. Równoważnie, jeżeli ma przeliczalny zbiór gęsty.

STWIERDZENIE 12. *Niech B będzie bazą ortonormalną w przestrzeni Hilberta H. Wówczas dla każdego wektora $v \in \mathsf{H}$*

$$v = \sum_{e \in B} (e|v)e. \tag{10}$$

DOWÓD: Z dowodu poprzedniego stwierdzenia wiemy, że tylko przeliczalna liczba wyrazów w sumie (10) jest różna od zera. Możemy więc ograniczyć się do przypadku przestrzeni ośrodkowej z bazą $\{e_1, e_2, \dots\}$. Oznaczmy przez W_n podprzestrzeń rozpiętą przez n pierwszych wektorów bazy. Niech $P_n v = \sum_1^n (e_i|v)e_i$. Oczywiście, $P_n v \in W_n$ oraz $v - P_n v \in W_n^\perp$, czyli $P_n v$ jest rzutem ortogonalnym v na W_n. Z dowodu twierdzenia o rzucie ortogonalnym wiemy, że

$$\|v - P_n v\| = \inf_{w \in W_n} \|v - w\|,$$

z czego wniosek, że ciąg $\|v - P_n v\|$ jest malejący. Ponieważ $\bigcup_n W_n$ jest zbiorem gęstym w H, ciąg $\|v - P_n v\|$ maleje do zera. ■

Niech H będzie ośrodkową przestrzenią Hilberta. Rozkład (10) indukuje odwzorowanie

$$I_B \colon \mathsf{H} \to \ell^2 \colon v \mapsto ((e_i|v)).$$

Odwzorowanie to zachowuje iloczyn skalarny: w oznaczeniach z dowodu Stwierdzenia 12

$$(v|w) = \lim_{n\to\infty} (P_n v|P_n w) = \lim_{n\to\infty} \sum_i^n \overline{(e_i|v)}(e_i|w) = (((e_i|v))|((e_i|w))).$$

Oznacza to, że baza w ośrodkowej przestrzeni Hilberta zadaja jej izomorfizm (będący izometrią) z przestrzenią ℓ^2.

5.1. Ważny przykład. Niech $\mathsf{H} = L^2([0, 2\pi])$, tzn. H jest uzupełnieniem przestrzeni funkcji ciągłych (o wartościach zespolonych) na odcinku $[0, 2\pi]$ z iloczynem skalarnym

$$(f|g) = \int_0^{2\pi} \bar{f}(t)g(t)\mathrm{d}t.$$

Niech $e_k(t) = \frac{1}{\sqrt{2\pi}} e^{\iota kt}$. Sprawdzamy, że funkcje te tworzą układ ortonormalny w $L^2([0, 2\pi])$:

$$(e_l|e_k) = \int_0^{2\pi} e^{-\iota lt} e^{\iota kt}\mathrm{d}t = \begin{cases} 0 & \text{dla } k \neq l \\ 1 & \text{dla } k = l. \end{cases}$$

Funkcje sin i cos rozdzielają punkty w $[0, 2\pi[$, więc z zespolonej wersji twierdzenia Weierstrassa funkcje ciągłe na $[0, 2\pi]$ z warunkiem $f(0) = f(\pi)$ można jednostajnie przybliżać wielomianami od e_1 i e_{-1}. Ale $(e_1)^k = e_k$, więc powłoka liniowa układu (e_k) jest jednostajnie gęsta w $C([0, 2\pi])$ (z warunkiem brzegowym), więc gęsta w $L^2([0, 2\pi])$. Rodzina $B = (e_k)$ tworzy bazę w $L^2([0, 2\pi])$. Odwzorowanie $I_B \colon L^2([0, 2\pi]) \to \ell^2$ nazywa się transformatą Fouriera, a reprezentacja

$$L^2([0, 2\pi]) \ni f(t) = \sum_{-\infty}^{\infty} a_k e^{\iota kt}$$

szeregiem Fouriera funkcji f.

6. Operatory w przestrzeni Hilberta.

Zacznę od przytoczenia (bez dowodu) jednego z fundamentalnych twierdzeń Banacha:

TWIERDZENIE 13 O WYKRESIE DOMKNIĘTYM. *Odwzorowanie $F \colon X \to Y$ między przestrzeniami Banacha jest ciągłe wtedy i tylko wtedy, gdy jego wykres jest domknięty w $X \times Y$.*

To, że wykres odwzorowania ciągłego jest domknięty jest dość oczywiste. Trudny jest dowód wynikania w drugą stronę.

Bezpośrednim wnioskiem jest stwierdzenie, że jeżeli F jest ciągłą bijekcją, to odwzorowanie F^{-1} jest ciągłe.

6.1. Sprzężenie hermitowskie. Niech H będzie przestrzenią Hilberta i niech $F \in \mathsf{End}(\mathsf{H}) = \mathsf{L}(\mathsf{H}, \mathsf{H})$. Dla każdego $w \in V$ odwzorowanie

$$\mathsf{H} \ni v \mapsto (w|Fv) \in \mathbb{C}$$

jest liniowe, zatem na mocy twierdzenia o reprezentacji funkcjonału liniowego istnieje $\widetilde{w} \in V$ takie, że $(w|Fv) = (\widetilde{w}|v)$ dla każdego v. Oznaczmy $\widetilde{w} = F^\dagger w$. Jak łatwo zauważyć, odwzorowanie $F^\dagger$ jest liniowe (addytywność oczywista):

$$(F^\dagger \lambda v \mid w) = (\lambda v \mid F(w)) = \bar{\lambda}(v \mid F(w)) = \bar{\lambda}(F^\dagger v \mid w) = (\lambda F^\dagger v \mid w).$$

Nazywamy je *sprzężeniem hermitowskim* operatora F.

Uwaga. Sprzężenie hermitowskie operatora F możemy zdefiniować używając odwzorowania $F_h \colon \mathsf{H} \to \mathsf{H}^*$:

$$F^\dagger = F_h^{-1} \circ F^* \circ F_h,$$

lub odwzorowania ${}_hF \colon \mathsf{H} \to \mathsf{H}^\#$:

$$F^\dagger = {}_hF^{-1} \circ F^\# \circ {}_hF.$$

STWIERDZENIE 14. *Dla* $F, G \in \mathsf{L}(\mathsf{H})$, $\lambda \in \mathbb{C}$ *mamy*

(1) $(F^\dagger)^\dagger = F$,
(2) $(F+G)^\dagger = F^\dagger + G^\dagger$,
(3) $(\lambda F)^\dagger = \overline{\lambda} F^\dagger$,
(4) $(F \circ G)^\dagger = G^\dagger \circ F^\dagger$.

DOWÓD:

(1) Dla $v, w \in \mathsf{H}$ mamy

$$(w|Fv) = (F^\dagger w|v) = \overline{(v|F^\dagger w)} = \overline{((F^\dagger)^\dagger v|w)} = (w|(F^\dagger)^\dagger v).$$

Stąd $Fv = (F^\dagger)^\dagger v$ i $F = (F^\dagger)^\dagger$.

(2) Oczywiste.

(3) Mamy

$$((\lambda F)^\dagger w|v) = (w|(\lambda F)v) = (w|F(\lambda v)) = (F^\dagger w|\lambda v) = (\overline{\lambda} F^\dagger w|v).$$

Stąd $(\lambda F)^\dagger = \overline{\lambda} F^\dagger$.

(4) $((F \circ G)^\dagger w|v) = (w|F \circ Gv) = (F^\dagger w|Gv) = (G^\dagger \circ F^\dagger w|v)$. ■

Wniosek. Przyporządkowanie $F \mapsto F^\dagger$ jest antyliniowym izomorfizmem w przestrzeni $\mathsf{L}(\mathsf{H})$.

DEFINICJA 15. *Operator* $F \colon \mathsf{H} \to \mathsf{H}$ *nazywamy* hermitowskim, *jeżeli* $F = F^\dagger$, *to znaczy dla wszystkich* $v, w \in \mathsf{H}$ $(w|Fv) = (Fw|v)$.

Łatwo zauważyć, że jeżeli operatory F, G są hermitowskie, to $F + G$ też jest hermitowski, a $F \circ G$ na ogół nie jest hermitowski. Z kolei dla hermitowskiego F, operator λF jest hermitowski wtedy i tylko wtedy, gdy $\lambda \in \mathbb{R}$.

DEFINICJA 16. *Operator* $F \colon \mathsf{H} \to \mathsf{H}$ *nazywamy* izometrią *jeżeli dla wszystkich* $v, w \in \mathsf{H}$ *mamy* $(Fv|Fw) = (v|w)$. *Jeżeli izometria jest surjekcją, to mówimy, że jest odwzorowaniem* unitarnym.

Z równości $\|Fv\| = \|v\|$ wynika ciągłość F, czyli nie trzeba jej zakładać. Oczywiście, operator unitarny jest też injekcją, bo $Fv = 0$ implikuje $\|Fv\| = \|v\| = 0$. Odwzorowanie odwrotne też jest odwzorowaniem unitarnym:

$$(F^{-1}v \,|\, F^{-1}w) = (FF^{-1}v \,|\, FF^{-1}w) = (v \,|\, w).$$

STWIERDZENIE 17.

i) *F jest unitarny wtedy i tylko wtedy, gdy jest surjekcją i dla wszystkich* $v \in \mathsf{H}$ *mamy* $(Fv|Fv) = (v|v)$.
ii) *F jest unitarny wtedy i tylko wtedy, gdy* $F^\dagger F = Id = FF^\dagger$, *czyli* $F^\dagger = F^{-1}$.
iii) *Jeżeli* F, G *są unitarne, to* $F \circ G$ *też jest unitarny.*

DOWÓD:

i) Wynika z formuły polaryzacyjnej (5) i (6).
ii) Mamy $(v|w) = (Fv|Fw) = (v|F^\dagger Fw)$. Stąd równość $F^\dagger F = Id$ jest równoważna izometrii F. Z równości $FF^\dagger = Id$ wynika surjekcja F. Równość ta oznacza też, że F^{-1} jest izometrią.
iii) $(v|w) = (Gv|Gw) = (F(Gv)|F(Gw))$. ■

Łatwo zauważyć, że suma operatorów unitarnych na ogół nie jest operatorem unitarnym. Z kolei dla F unitarnego λF jest unitarny wtedy i tylko wtedy, gdy $|\lambda| = 1$.

Teraz kilka słów o operatorach rzutowych. Operator $P \in \mathsf{L}(\mathsf{H})$ taki, że $P^2 = P$ nazywany jest operatorem rzutowym. Zadaje on rozkład $\mathsf{H} = V \oplus W$ na sumę prostą podprzestrzeni domkniętych: $V = \operatorname{im} P = \ker(Id - P)$ i $W = \ker P$. Operator rzutowy P jest *operatorem rzutu ortogonalnego*, jeżeli $\ker P = (\operatorname{im} P)^\perp$, to znaczy, przestrzeń wzdłuż której się rzutuje jest dopełnieniem ortogonalnym przestrzeni na którą się rzutuje.

STWIERDZENIE 18. *$P\colon V \to V$ jest operatorem rzutu ortogonalnego wtedy i tylko wtedy, gdy $P^2 = P$ i $P = P^\dagger$.*

DOWÓD: Niech $P^2 = P$ i $P^\dagger = P$. Równość $P^2 = P$ oznacza, że P jest operatorem rzutowym. Mamy pokazać, że $(\operatorname{im} P)^\perp = \ker P$. Niech $v \in \ker P$ i niech $w \in \operatorname{im} P$, tzn., $w = Pw'$. Wówczas, ponieważ $P = P^\dagger$,

$$(v|w) = (v|Pw') = (P^\dagger v|w') = (Pv|w) = 0.$$

$\operatorname{im} P$ i $\ker P$ są domknięte i rozpinają H, więc $\ker P = (\operatorname{im} P)^\perp$.

Załóżmy teraz, że P jest operatorem rzutu ortogonalnego na podprzestrzeń V. Niech $v, v_1 \in \mathsf{H}$ i niech $v = w + w', v_1 = w_1 + w_1'$ gdzie $w, w_1 \in V$ a $w', w_1' \in V^\perp$. Mamy wówczas

$$(v|Pv_1) = (v|w_1) = (w|w_1) = (w|w_1 + w_1') = (Pv|v_1).$$

■

Jeżeli P jest operatorem rzutu ortogonalnego na podprzestrzeń V, to $Id - P$ jest operatorem rzutu ortogonalnego na $V^\perp$. Na koniec ciekawostka:

STWIERDZENIE 19. *Jeżeli $F\colon \mathsf{H} \to \mathsf{H}$ jest jednocześnie unitarny i hermitowski, to $\frac{1}{2}(Id - F)$ oraz $\frac{1}{2}(Id + F)$ są operatorami rzutu ortogonalnego.*

DOWÓD: Zauważmy, że $\frac{1}{2}(Id - F)$ jest hermitowski. Ponadto

$$(\frac{1}{2}(Id - F))^2 = \frac{1}{4}(Id - 2F + F^2) = \frac{1}{4}(Id - 2F + F^\dagger F) = \frac{1}{2}(Id - F).$$

Na mocy poprzedniego stwierdzenia dostajemy tezę dla $\frac{1}{2}(Id - F)$. Wystarczy teraz zauważyć, że $\frac{1}{2}(Id + F) = I - \frac{1}{2}(Id - F)$. ■

7. Teoria Dystrybucji.

Jak w skrypcie Analiza III.

8. Odwzorowania Fredholma i zwarte.

Niech X i Y będą przestrzeniami Hilberta. Jak wiemy z Twierdzenia Riesza, możemy utożsamiać funkcjonały liniowe na przestrzeni Hilberta z elementami przestrzeni Hiberta. Możemy, ale nie musimy i w tym rozdziale robić tego nie będziemy. Będziemy natomiast korzystać z kanonicznego izomorfizmu $X = (X^*)^*$ i równości dla odwzorowań $F = (F^*)^*$. Niech $F\colon X \to Y$ będzie odwzorowaniem liniowym i ciągłym. Odzwzorowanie sprzężone

$$F^*\colon Y^* \to X^*\colon f \mapsto f \circ F$$

jest też ciągłe (ograniczone) i jego norma jest równa normie F:

$$\|F\| = \sup_{\|x\|=1} \|Fx\| = \sup_{\|x\|=1} \sup_{\|f\|=1} |\langle f, Fx\rangle| = \sup_{\|f\|=1} \sup_{\|x\|=1} |\langle F^* f, x\rangle| = \sup_{\|f\|=1} \|F^* f\| = \|F^*\|$$

Interesować nas będzie równanie $Fx = b$. Warunkiem rozwiązalności jest $b \in \operatorname{im} F$, więc pytanie o rozwiązalność równania sprowadza się do opisu obrazu odwzorowania F.

Stwierdzenie 1.

$$\overline{\operatorname{im} F} = (\ker F^*)^\circ = \{y \in Y \mid \langle g, y\rangle = 0 \ \forall g \in \ker F^*\}.$$

Dowód: Jeżeli $y \in \operatorname{im} F$, to dla $g \in \ker F^*$ mamy

$$\langle y, g\rangle = \langle Fx, g\rangle = \langle x, F^* g\rangle = 0.$$

Stąd $\operatorname{im} F \subset (\ker F^*)^\circ$, a poniewa $(\ker F^*)^\circ$ jest zbiorem domkniętym, również $\overline{\operatorname{im} F} \subset (\ker F^*)^\circ$.

W drugą stronę: zawieranie $\overline{\operatorname{im} F} \supset (\ker F^*)^\circ$ jest równowane zawieraniu $(\overline{\operatorname{im} F})^\circ \subset \ker F^*$. Jeżeli $g \in (\overline{\operatorname{im} F})^\circ$, to dla każdego $x \in X$

$$0 = \langle Fx, g\rangle = \langle x, F^* g\rangle$$

i stąd $F^* g = 0$. Zatem $(\operatorname{im} F)^\circ \subset \ker F^*$, ale $(\operatorname{im} F)^\circ = (\overline{\operatorname{im} F})^\circ$. ∎

Bezpośrednim wnioskiem jest

Stwierdzenie 2. *Jeżeli $F\colon X \to Y$ jest injekcją w zbiór gęsty, to również $F^*\colon Y^* \to X^*$ jest injekcją w zbiór gęsty.*

Dowód: $\ker F^* = (\operatorname{im} F)^\circ = 0$ i $(\operatorname{im} F^*)^\circ = \ker F^{**} = \ker F = 0$ ∎

Z taką sytuacją spotkaliśmy się już przy omawianiu dystrybucji.

Podstawowym dla dalszych rozważań jest następujące twierdzenie:

Twierdzenie 1. *Dwa warunki są równoważne:*

(1) $\dim \ker F < \infty$ *i* $\operatorname{im} F$ *jest domknięty,*
(2) *z każdego ciągu ograniczonego* (x_n) *takiego, że ciąg* Fx_n *jest zbieżny, można wybrać podciąg zbieżny.*

Dowód:

(2) $\Rightarrow$ (1)

Jeżeli wybierzemy ciąg ograniczony (x_n) taki, że $x_n \in \ker F$, to można z niego wybrać podciąg zbieżny. Stąd domknięta kula jednostkowa w $\ker F$ jest zwarta, czyli wymiar $\ker F$ jest skończony. Oznaczmy $V = (\ker F)^\perp$, ortogonalne dopełnienie $\ker F$. Indukowane odwzorowanie $F\colon V \to Y$ jest injekcją. Pokażemy, że istnieje c takie, że dla $x \in V$ zachodzi nierówność

$$\|x\| \leqslant c\|Fx\|. \tag{11}$$

Istotnie, gdyby takie c nie istniało, to dla każdego j można by znaleźć wektor x_j taki, że $1 = \|x_j\| \geqslant j\|Fx_j\|$. Dla takiego ciągu $Fx_j \to 0$, więc można wybrać podciąg (x_{j_k}), zbieżny do pewnego x. Oczywiście $\|x\| = 1$ i $x \in V$, ale z ciągłości F, $Fx = \lim F(x_{j_k}) = 0$. Sprzeczność, bo F na V jest injekcją.

Nierówność (11) oznacza, że odwzorowanie odwrotne (na $\operatorname{im} F$) jest ciągłe, a ponieważ wykres jest domknięty, to nie można go przez ciągłość rozszerzyć. Stąd jego dziedzina, czyli $\operatorname{im} F$, musi być domknięty.

(1) $\Rightarrow$ (2)

Niech (x_j) będzie ciągiem ograniczonym i takim, że ciąg Fx_j jest zbieżny. Rozłożymy ten ciąg na sumę dwóch: $x_j = y_j + z_j$, gdzie $y_j \in \ker F$ i $z_j \in (\ker F)^\perp$. Ciągi te są też ograniczone, bo $\|x_j\|^2 = \|y_j\|^2 + \|z_j\|^2$. Przestrzeń $\ker F$ jest wymiaru skończonego, więc można wybrać podciąg zbieżny (y_{j_k}). Ponadto $Fz_j = Fx_j$, więc ciąg (Fz_j) jest zbieżny. Ale $F\colon (\ker F)^\perp \to \operatorname{im} F$ jest izomorfizmem ($\operatorname{im} F$ jest domknięty, więc odwrotne odwzorowanie jest ciągłe), więc ciąg (z_j) jest zbieżny. Stąd ciąg $x_{j_k} = y_{j_k} + z_{j_k}$ jest zbieżny. ■

8.1. Odwzorowania zwarte. Odwzorowanie liniowe i ciągłe $K\colon X \to Y$ nazywamy zwartym, jeżeli zbiory ograniczone przeprowadza w prezwarte, tzn. z każdego ciągu ograniczonego (x_n) można wybrać podciąg x_{n_k} taki, że ciąg Kx_{n_k} jest zbieżny. Oczywistym jest, że złożenie odwzorowania ograniczonego (czyli ciągłego) ze zwartym jest odwzorowaniem zwartym (mówimy tylko o odwzorowaniach liniowych).

Stwierdzenie 2. *Odwzorowanie dualne do odwzorowania zwartego jest odwzorowaniem zwartym*

Dowód: Niech $K\colon X \to Y$ będzie odwzorowaniem zwartym, a $C_X\colon X \to X^*$ i $C_Y\colon Y \to Y^*$ izomorfizmami Riesza. Niech ciag (g_n) w Y^* będzie ograniczony. Ciągi (K^*g_n) i $(C_X^{-1}K^*g_n)$ są też ograniczone. Zatem istnieje podciąg g_{n_k} taki, że ciąg $(KC^{-1}K^*g_{n_k})$ jest zbieżny (bo K jest zwarty). Mamy

$$\begin{aligned}
\|K^*(g_n - g_m)\|^2 &= \big(C_X^{-1}K^*(g_n - g_m) \mid C_X^{-1}K^*(g_n - g_m)\big)\\
&= \big\langle K^*(g_n - g_m), C_X^{-1}K^*(g_n - g_m)\big\rangle = \big\langle (g_n - g_m), KC_X^{-1}K^*(g_n - g_m)\big\rangle\\
&= \big(C_Y^{-1}(g_n - g_m) \mid KC_X^{-1}K^*(g_n - g_m)\big)\\
&\leqslant \|C_Y^{-1}(g_n - g_m)\|\|KC_X^{-1}K^*(g_n - g_m)\|\\
&\leqslant M\|KC^{-1}K^*(g_n - g_m)\|
\end{aligned}$$

bo ciąg (g_n) jest ograniczony. Ze zbieżności $(\|KC^{-1}K^*(g_{n_k})\|)$ wynika zbieżność $(K^*g_{n_k})$. ■

Podam teraz bardzo ważny przykład odwzorowań zwartych. Niech Ω będzie obszarem ograniczonym w $\mathbb{R}^n$, z gładkim brzegiem. Przestrzeń $H^s(\Omega)$ definiujemy jako uzupełnienie przestrzeni funkcji gładkich względem normy iloczynu skalarnego

$$(f \mid g)_s = \sum_{|\alpha|\leqslant s} \int_\Omega D^\alpha \overline{f} D^\alpha g.$$

Równoważną normę (dla $s > 0$) dostaniemy biorąc jako iloczym skalarny odwzorowanie

$$(f, g) \mapsto \sum_{|\alpha|=s} \int_\Omega D^\alpha \overline{f} D^\alpha g + \int_\Omega \overline{f} g.$$

Przestrzeń $H^s(\Omega)$ możemy też definiować jako przestrzeń funkcji z $L^2(\Omega)$, których pochodne dystrybucyjne do rzędu s też należą do $L^2(\Omega)$. Mamy oczywisty ciąg gęstych włożeń

$$H^0(\Omega) \supset H^1(\Omega) \supset H^2(\Omega) \supset \cdots H^s(\Omega) \cdots.$$

Włożenia te są odwzorowaniami zwartymi (Lemat Rellicha). Mamy stąd dualny ciąg zwartych i gęstych włożeń

$$(H^0(\Omega))^* \subset (H^1(\Omega))^* \subset (H^2(\Omega))^* \subset \cdots (H^s(\Omega))^* \cdots .$$

Utożsamiamy $H^0(\Omega) = L^2(\Omega)$ z dualną do niej i oznaczamy $(H^s(\Omega))^* = H^{-s}(\Omega)$. Otrzymujemy w ten sposób ciąg zwartych i gęstych włożeń

$$\cdots H^{-s}(\Omega) \supset \cdots \supset H^{-1}(\Omega) \supset H^0(\Omega) \supset H^1(\Omega) \supset H^2(\Omega) \supset \cdots H^s(\Omega) \cdots . \tag{12}$$

8.2. Odwzorowania Fredholma. Niech V będzie podprzestrzenią wektorową przestrzeni Hiberta X. Przestrzeń ilorazowa X/V ma naturalną strukturę przestrzeni wektorowej, ale jeżeli V nie jest podprzestrzenią domkniętą, to nie ma w X/V topologii takiej, że rzut kanoniczny jest odwzorowaniem ciągłym (V jest przeciwobrazem zera, a punkt jest zbiorem domkniętym). Niech więc V będzie podprzestrzenią domkniętą. Mamy oczywiste utożsamienie X/V z $V^\perp$, a stąd naturalną strukturę przestrzeni Hilberta w X/V. Ponieważ $X = V + V^\perp$ i w konsekwencji $X^* = V^\circ + (V^\perp)^\circ$, przestrzeń dualna do X/V może być utożsamiona z V°. W szczególności, dla liniowego i ciągłego odwzorowania $F\colon X \to Y$ mamy

$$\big(Y/\overline{\operatorname{im} F}\big)^* \simeq \big(\overline{\operatorname{im} F}\big)^\circ, \quad Y/\overline{\operatorname{im} F} \simeq \big(\overline{\operatorname{im} F}\big)^\perp.$$

Stąd i ze Stwierdzenia 1 wynika, że $\dim\big(Y/\overline{\operatorname{im} F}\big) = \dim\ker F^*$. Przestrzeń $Y/\overline{\operatorname{im} F}$ nazywana jest kojądrem F.

Definicja 1. Odwzorowanie (liniowe i ciągłe) $F\colon X \to Y$ nazywamy *Fredholma*, jeżeli $\operatorname{im} F$ jest podprzestrzenią domkniętą, a $\dim\ker F$ oraz $\dim\big(Y/\operatorname{im} F\big)$ są skończone.

Stwierdzenie 3. *Jeżeli odwzorowanie $F\colon X \to Y$ jest Fredholma, to dualne odwzorowanie $F^*\colon Y^* \to X^*$ jest też Fredholma.*

Dowód: Mamy $\operatorname{im} F = (\ker F^*)^\circ$ i stąd $\ker F^* = \big(Y/\operatorname{im} F\big)^*$, więc $\dim\ker F^* = \dim\big(Y/\operatorname{im} F\big) < \infty$. Odwzorowanie F indukuje izomorfizm przestrzeni Hilberta $\widetilde{F}\colon X/\ker F \to \operatorname{im} F$, więc też izomorfizm dualny

$$\widetilde{F}^*\colon (\operatorname{im} F)^* \to \big(X/\ker F\big)^*.$$

Ponieważ $\operatorname{im} F = (\ker F^*)^\circ$, to $(\operatorname{im} F)^* = Y^*/\ker F^*$, a z równości $\overline{\operatorname{im} F^*} = (\ker F)^\circ$ wynika

$$\big(X/\ker F\big)^* = (\ker F)^\circ = \overline{\operatorname{im} F^*}.$$

Mamy zatem izomorfizm $\widetilde{F}^*\colon Y^*/\ker F^* \to \overline{\operatorname{im} F^*}$, a z drugiej strony bijekcję $\widetilde{F^*}\colon Y^*/\ker F^* \to \operatorname{im} F^*$. Ponieważ $\widetilde{F}^* = \widetilde{F^*}$ (sprawdzić!), to $\overline{\operatorname{im} F^*} = \operatorname{im} F^*$. Oczywiście $\dim(X^*/\operatorname{im} F^*) = \dim\ker F < \infty$. ∎

Bardzo ważny jest fakt, że zaburzenie odwzorowania Fredholma odwzorowaniem zwartym pozostaje odwzorowaniem Fredholma.

Twierdzenie 3. *Jeżeli $F\colon X \to Y$ jest Fredholma, a $K\colon X \to Y$ zwarty, to $F + K\colon X \to Y$ jest też Fredholma.*

Dowód: Niech (x_n) będzie ograniczonym ciągiem w X takim, że ciąg $\big((F+K)x_n\big)$ jest zbieżny. Ze zwartości K istnieje podciąg (x_{n_k}) taki, że (Kx_{n_k}) jest ciągiem zbieżnym. Zatem (Fx_{n_k}) jest też zbieżny, a ponieważ F jest Fredholma, to (Twierdzenie 1) istnieje podciąg zbieżny $(x_{n_{k_l}})$. Z Twierdzenia 1 $\dim\ker(F+K) < \infty$ i obraz $\operatorname{im}(F+K)$ jest domknięty. Zastępując F, K przez ich dualne i korzystając ze Stwierdzeń 2, 3 $\dim\ker(F+K)^* < \infty$. ∎

8.3. Zagadnienie brzegowe. Dla prostoty, niech $\mathbb{K} = \mathbb{R}$. Niech B będzie funkcją symetryczną, biliniową i ciągła na przestrzeni Hilberta X. Odpowiadające jej odwzorowanie

liniowe $X \to X^*$ jest też ciągłe, a z powodu symetrii B, samosprzężone. Prostym rachunkiem sprawdzamy, że jest to pochodna funkcji $\varphi\colon X \to \mathbb{R}\colon x \mapsto \frac{1}{2}B(x,x)$.
Weźmy teraz, dla obszaru spójnego Ω, $X = H^1(\Omega)$ i $B(f,g) = \int_\Omega \sum_i f_{,i}\, g_{,i}$. Widać, że $B(f,g) = (f \mid g) - \int_\Omega fg$, gdzie iloczyn skalarny jest w $H^1(\Omega)$. Zatem stowarzyszone odwzorowanie liniowe $F\colon X \to X^*$ jest różnicą $F = C_1 - C_0$, gdzie $C_1\colon X \to X^*$ jest izomorfizmem Riesza, a $C_0\colon X \to X^*$ jest izomorfizmem Riesza $H^0(\Omega) \to (H^0(\Omega))^*$, złożonym z kanonicznymi włożeniami $H^1(\Omega) \to H^0(\Omega)$ i $(H^0(\Omega))^* \subset (H^1(\Omega))^*$. Wiemy (12), że włożenia te są gęste i zwarte, zatem odwzorowanie C_0 jest zwarte. Z Twierdzenia 3 wiemy, że odwzorowanie F jest Fredholma. Ponieważ forma B jest symetryczna, F jest odwzorowaniem samosprzężonym. Zatem $\operatorname{im} F = (\ker F)^\circ$. Z dodatniości funkcji kwadratowej $f \mapsto B(f,f)$ i z tego, że F jest pochodną funkcji $\frac{1}{2}B(f,f)$, z $B(f,f) = 0$ wynika $F(f) = 0$. Zatem jądro F jest zbiorem zer funkcji kwadratowej $B(f,f)$. Mamy więc $\ker F = \{f \mid f_{,i} = 0\}$, czyli f z $\ker F$ jest funkcją stałą. Zobaczmy, jak można interpretować $F(f)$ dla gładkich funkcji.Z twierdzenia Gaussa-Greena,

$$\int_\omega \sum_i f_{,i}\, g_{,i} = -\int_\Omega (\triangle f) g + \int_{\partial\Omega} f_n g,$$

gdzie f_n jest składową normalną do brzegu gradientu funkcji f. Para (ρ, D_n) należy do obrazu f jeżeli $-\int_\Omega \rho + \int_{\partial\Omega} D_n = 0$. W elektrostatyce ten warunek oznacza, że całkowity ładunek jest równy zero (tw. Gaussa).

8.4. Alternatywa Fredholma. Bardzo istotne w zastosowaniach jest poniższe twierdzenie.

Twierdzenie 4 Alternatywa Fredholma. *Niech $A\colon X \to X$ będzie zwartym odwzorowaniem przestrzeni Hilberta X w siebie. Możliwe są dwie, wykluczające się sytuacje*

(1) *odwzorowanie $(Id - A)$ jest surjekcją, czyli dla każdego $g \in X$ istnieje rozwiązanie równania $(Id - A)f = g$,*
(2) *jądro odwzorowania $(Id - A)$ jest nietrywialne, czyli równanie jednorodne $(Id - A)f = 0$ ma nietrywialne rozwiązanie, i ponadto wymiary jądra i kojądra $(Id - A)$ są równe.*

Dowód: Twierdzenie to jest oczywiste w przypadku A samosprzężonego, bo wtedy obraz jest anihilatorem jądra (operator $I - A$ jest Fredholma, więc obraz jest domknięty). W przypadku ogólnym wystarczy pokazać, że $\dim\ker(Id - A) = \dim\operatorname{coker}(Id - A)$. Istotnie, równość ta oznacza, że $\ker(Id - A) = \{0\}$ wtedy i tylko wtedy, gdy $\operatorname{im}(Id - A) = X$. Dowód przeprowadzimy w trzech etapach. Oznaczmy $T = Id - A$.

i) $\dim\operatorname{coker} T = 0 \Rightarrow \dim\ker T = 0$
oznaczmy $M_n = \ker T^n$. Oczywiście $\{0\} = M_0 \subset M_1 \subset M_2 \subset \ldots$. Ciąg ten stabilizuje się, bo gdyby $M_n \subsetneq M_{n+1}$ dla każdego n, to istniałby ciąg ortonormalny (f_n) taki, że $f_n \in M_n,\ f_n \notin M_{n+1}$. Dla takiego ciągu

$$\|Af_n - Af_n\|^2 = \|f_n - (Tf_n + f_m - Tf_m)\| = \|f_n\|^2 + \|Tf_n + f_m - Tf_m\|^2 \geqslant \|f_n\|^2 = 1$$

dla $n > m$, bo wtedy $(Tf_n + f_m - Tf_m) \in M_{n-1}$, a $f_n \perp M_{n-1}$. Czyli z ciągu Af_n nie można wybrać podciągu zbieżnego, co jest sprzeczne ze zwartością A. Zatem ciąg M_n stabilizuje się. Przypuśćmy teraz, że $\dim\ker T \neq 0$, tzn. istnieje niezerowy f_1 taki, że $Tf_1 = 0$. Ale $\operatorname{im} T$ jest całą przestrzenią, więc istnieje f_2 takie, że $Tf_2 = f_1$. Podobnie $f_2 = Tf_3$ itd. Ciąg (f_n) ma tą własność, że $f_n \in M_n$ i $f_n \notin M_{n-1}$. Istnienie takiego ciągu oznacza, że ciąg podprzestrzeni (M_n) nie stabilizuje się. Sprzeczność.

ii) $\dim\ker T = 0 \Rightarrow \dim\operatorname{coker} T = 0$
Wystarczy udowodnić (3), czyli,+ że $\dim\ker T = \dim\operatorname{coker} T^*$, $\dim\operatorname{coker} T = \dim\ker T^*$ i zastosować i) do operatora T^*.

iii) $\dim\ker T = \dim\operatorname{coker} T$
Niech $(\varphi_1, \ldots, \varphi_n)$ będzie bazą $\ker T$, a $(\psi_1, \ldots, \psi_m)$ bazą przestrzeni dopełniającej $\operatorname{im} T$. Przypuśćmy, że $m > n$. Modyfikujemy operator A kładąc

$$\widetilde{A} = A - \sum_{i=1}^{n} (\cdot \mid \varphi_i)\psi_i, \quad \widetilde{T} = Id - \lambda\widetilde{A}$$

$\widetilde{A}$ jest operatorem zwartym jak suma zwartego i skończenie-wymiarowego, więc możemy stosować ii) do $\widetilde{T}$. Jeżeli $\widetilde{T}f = 0$, to $Tf = 0$ i $\sum_{i=1}^{n}(f \mid \varphi_i)\psi_i = 0$, a stąd $(f \mid \varphi_i) = 0$. Zatem $f = 0$. Sprzeczność, bo z $m > n$ wynika, że kojądro $\widetilde{T}$ nie jest zerowe.

■

9. Równania całkowe.

9.1. Operator całkowy. Dla prostoty, zajmijmy się równaniami całkowymi dla funkcji na odcinku $I = [a, b]$. *Operatorem całkowym Fredholma* nazywamy operator A postaci

$$Af(x) = \int_a^b K_A(x, y)f(y)\mathrm{d}y.$$

Oszacujmy Af w normie $L^2(I)$:

$$\int_a^b \big[Af(x)\big]^2 \mathrm{d}x = \int_a^b \left(\int_a^b K_A(x,y)f(y)\mathrm{d}y\right)^2 \mathrm{d}x$$
$$\leqslant \int_a^b \left(\int_a^b K_A^2(x,y)\mathrm{d}y \int_a^b f^2(y))\mathrm{d}y\right)\mathrm{d}x = \int_a^b f^2(y)\mathrm{d}y \int_a^b\int_a^b K_A^2(x,y)\mathrm{d}y,$$

czyli

$$\|Af\|_{L^2(I)} \leqslant \|K_A(x,y)\|_{L^2(I\times I)}\|f\|_{L^2(I)}.$$

Oznacza to, że $K_A \in L^2(I \times I)$ definiuje ciągły operator $A\colon L^2(I) \to L^2(I)$ z szacowaniem normy $\|A\| \leqslant \|K_A(x,y)\|_{L^2(I\times I)}$.

Zauważmy, że

(1) $K_{A^*}(x,y) = K_A(y,x)$,
(2) $K_{AB}(x,y) = \int_a^b K_A(x,z)K_B(z,y)\mathrm{d}z$.

Pokażemy teraz, że operator całkowy jest zwarty. Przydatny tu będzie ogólniejszy fakt.

Twierdzenie 5. *$K\colon X \to X$ jest operatorem zwartym wtedy i tylko wtedy, gdy jest granicą, w normie operatorowej, operatorów skończenie-wymiarowych, tj. takich, których obraz jest wymiaru skończonego.*

Dowód: Niech K będzie operatorem zwartym. Obraz kuli jednostkowej jest zbiorem prezwartym, czyli jego domknięcie jest zbiorem zwartym. Dla każdego $\varepsilon > 0$ istnieje skończony zbiór $x_1, \ldots, x_{n_\varepsilon}$ w X taki, że $\bigcup_i K(x_i, \varepsilon)$ zawiera obraz kuli jednostkowej. Oznacza to, że dla każdego $\|x\| < 1$ odległość $K(x)$ od przestrzeni V_ε rozpiętej na $x_1, \ldots, x_{n_\varepsilon}$ jest mniejsza od ε. Stąd wniosek, że $\|K - P_\varepsilon \circ K\| \leqslant \varepsilon$, gdzie P_ε jest rzutem ortogonalnym na V_ε. K jest granicą normową $P_\varepsilon \circ K$.

W drugą stronę. Niech $\|K - K_n\| < \frac{1}{n}$, gdzie K_n jest operatorem skończenie-wymiarowym. Weźmy ciąg ograniczony x_j, $\|x_j\| \leqslant 1$. Istnieje podciąg $x_{j(1,k)}$ taki, że ciąg $K_1(x_{j(1,k)})$ jest zbieżny. Następnie wybieramy podciąg $x_{j(2,k)}$ ciągu $x_{j(1,k)}$, tak, by ciąg $K_2(x_{j(2,k)})$ był zbieżny, itd. Podciąg $x_{j(n,k)}$ jest więc taki, że dla $m \leqslant n$ ciąg $k \mapsto K_m(x_{j(n,k)})$ jest zbieżny. Można przy tym tak wybierać te podciągi, by dla $k, l > n$ zachodziła nierówność $\|K_n(x_{j(n,k)}) - K(x_{j(n,l)})\| < \frac{1}{n}$. Ciąg $K(x_{j(n,n)})$ jest zbieżny, bo dla $m > n$

$$\|K(x_{j(n,n)}) - K(x_{j(m,m)})\| \leqslant \|K(x_{j(n,n)}) - K_n(x_{j(n,n)})\|$$
$$+ \|K_n(x_{j(n,n)}) - K_n(x_{j(m,m)})\| + \|K_n(x_{j(m,m)}) - K(x_{j(m,m)})\| \leqslant \frac{3}{n}.$$

■

Dla pokazania, że operator całkowy A jest zwarty, wystarczy udowodnić, że jest granicą operatorów skończenie-wymiarowych. $L^2(I)$ można wybrać przeliczalną bazę ortonormalną (φ_n). Funkcje $\varphi_{nm}(x, y) = \varphi_n(x)\varphi_m(y)$ tworzą bazę ortonormalną w $L^2(I \times I)$. Mamy więc

$$K_A = \sum_{n,m} \alpha_{nm}\varphi_{nm}, \quad Af = \sum_{n,m} \alpha_{nm}(\varphi_n \mid f)\varphi_m.$$

Oznaczmy przez P_N rzut ortogonalny na podprzestrzeń rozpiętą przez N pierwszych wektorów bazy. Mamy

$$P_N A P_N f = \sum_{n,m=1}^{N} \alpha_{nm}(\varphi_n \mid f)\varphi_m.$$

Stąd $P_N A P_N \to A$ w normie operatorowej, bo jądra zbiegają w $L^2(I \times I)$. Zajmiemy się teraz równaniem całkowym

$$(Id - \lambda A)f = g,$$

gdzie A jest, jak zwykle, operatorem całkowym z jądrem K_A. Istotna jest zwartość tego operatora.

9.2. Wzory Fredholma. Spróbujmy rozwiązać równanie

$$(Id - \lambda A)f = g, \quad Af(x) = \int_a^b K_A(x, y)f(y)\mathrm{d}y. \tag{13}$$

W pierwszym odruchu piszemy $g = (Id - \lambda A)^{-1}f$ i

$$(Id - \lambda A)^{-1} = \frac{1}{Id - \lambda A} = \sum_{k=0}^{\infty} \lambda^k A^k, \tag{14}$$

ale gwarancję zbieżności mamy tylko w przypadku $\|\lambda A\| < 1$, co nam nie wystarcza.

Przedstawmy $(Id - \lambda A)^{-1}$ w postaci $Id + \lambda R$ i spróbujmy wyznaczyć operator R, zwany *rezolwentą* równania. Jak łatwo zauważyć,

$$R = \frac{A}{Id - \lambda A}.$$

Ponieważ A jest operatorem zwartym, to także R jest operatorem zwartym. Może całkowym? Jeżeli tak, to jego jądro $\mathcal{R}$ nazywamy *jądrem rozwiązującym*. Ponieważ mamy tożsamość

$$(Id - \lambda A)R = A,$$

to jądro rozwiązujące spełnia (jeżeli istnieje) równanie

$$\mathcal{R}(x, y) - \lambda \int_a^b K_A(x, z)\mathcal{R}(z, y)\mathrm{d}z = K_A(x, y). \tag{15}$$

Oznaczmy

$$K_A \begin{pmatrix} x_1 & \dots & x_m \\ y_1 & \dots & y_m \end{pmatrix} = \det\big[K(x_i, y_j)\big] \tag{16}$$

i zdefiniujmy

$$d(\lambda) = 1-\frac{\lambda}{1!}\int_a^b K_A\begin{pmatrix} y_1 \\ y_1 \end{pmatrix} \mathrm{d}y_1+\dots+\frac{(-1)^n\lambda^n}{n!}\int\cdots\int K_A\begin{pmatrix} y_1 & \dots & y_n \\ y_1 & \dots & y_n \end{pmatrix}\mathrm{d}y_1\cdots\mathrm{d}y_n+\cdots \tag{17}$$

oraz

$$\begin{aligned} d(x,y;\lambda) = K_A\begin{pmatrix} x \\ y \end{pmatrix} - \frac{\lambda}{1!}\int_a^b K_A\begin{pmatrix} x & y_1 \\ y & y_1 \end{pmatrix}\mathrm{d}y_1 + \cdots \\ + \frac{(-1)^n\lambda^n}{n!}\int\cdots\int K_A\begin{pmatrix} x & y_1 & \dots & y_n \\ y & y_1 & \dots & y_n \end{pmatrix}\mathrm{d}y_1\cdots\mathrm{d}y_n + \cdots . \end{aligned} \tag{18}$$

$d(\lambda)$ nazywane jest *wyznacznikiem Fredholma.*

Twierdzenie 6.

(1) *Szeregi $d(\lambda)$ i $d(x,y;\lambda)$ mają nieskończony promień zbieżności.*

(2) $\mathcal{R}(x,y) = \dfrac{d(x,y;\lambda)}{d(\lambda)}$.

Dowód: Skorzystamy z nierówności Hadamarda

$$\big|\det[a_{ij}]\big| \leqslant \prod_{j=1}^n \sqrt{\sum_{i=1}^n |a_{ij}|^2}.$$

Wynika z niej, że jeżeli $\sup|K_A(x,y)| = M$, to

$$\left|K_A\begin{pmatrix} x_1 & \dots & x_m \\ y_1 & \dots & y_m \end{pmatrix}\right| \leqslant M^m m^{\frac{m}{2}}.$$

Jeżeli więc oznaczymy przez a_n współczynnik przy λ^n w rozwinięciu $d(\lambda)$, to

$$|a_n| \leqslant \frac{M^n n^{\frac{n}{2}}}{n!}(b-a)^n$$

i stąd

$$\sqrt[n]{|a_n|} \leqslant M(b-a)\frac{n^{\frac{1}{2}}}{\sqrt[n]{(n!)}} \xrightarrow[n\to\infty]{} 0.$$

Zatem promień zbieżności szeregu (17) jest nieskończony. Podobnie pokazujemy, że promień zbieżności szeregu (18) jest nieskończony.

Zauważmy teraz, że z rozwinięcia Laplace'a względem pierwszej kolumny mamy tożsamość

$$\begin{aligned} & K_A\begin{pmatrix} x & y_1 & \dots & y_n \\ y & y_1 & \dots & y_n \end{pmatrix} = \\ & K_A(x,y)K_A\begin{pmatrix} y_1 & \dots & y_n \\ y & \dots & y_n \end{pmatrix} - K_A(x,y_1)K_A\begin{pmatrix} y_1 & y_2 & \dots & y_n \\ y & y_2 & \dots & y_n \end{pmatrix} + \cdots \\ & + (-1)^k K_A(x,y_k)K_A\begin{pmatrix} y_1 & \dots & y_k & y_{k+1} & \dots & y_n \\ y & \dots & y_{k-1} & y_{k+1} & \dots & y_n \end{pmatrix} + \cdots \\ & = K_A(x,y)K_A\begin{pmatrix} y_1 & \dots & y_n \\ y & \dots & y_n \end{pmatrix} - K_A(x,y_1)K_A\begin{pmatrix} y_1 & y_2 & \dots & y_n \\ y & y_2 & \dots & y_n \end{pmatrix} - \cdots \\ & - K_A(x,y_k)K_A\begin{pmatrix} y_k & y_1 & \dots & y_{k-1} & y_{k+1} & \dots & y_n \\ y & y_1 & \dots & y_{k-1} & y_{k+1} & \dots & y_n \end{pmatrix} - \cdots \end{aligned}$$

czyli wszystkie składniki, począwszy od drugiego, dają ten sam wkład do całki

$$\int \cdots \int K_A \begin{pmatrix} y_1 & \dots & y_n \\ y_1 & \dots & y_n \end{pmatrix} \mathrm{d}y_1 \cdots \mathrm{d}y_n.$$

Mamy więc

$$\int \cdots \int K_A \begin{pmatrix} x & y_1 & \dots & y_n \\ y & y_1 & \dots & y_n \end{pmatrix} \mathrm{d}y_1 \cdots \mathrm{d}y_n = K_A(x,y) \int \cdots \int K_A \begin{pmatrix} y_1 & \dots & y_n \\ y_1 & \dots & y_n \end{pmatrix} \mathrm{d}y_1 \cdots \mathrm{d}y_n$$
$$- n \int \cdots \int K_A(x,s) K_A \begin{pmatrix} s & y_1 & \dots & y_{n-1} \\ y & y_1 & \dots & y_{n-1} \end{pmatrix} \mathrm{d}s \mathrm{d}y_1 \cdots \mathrm{d}y_{n-1}.$$

Dostajemy stąd tożsamość

$$d(x,y;\lambda) = K_A(x,y) d(\lambda) + \lambda \int K_A(x,s) d(s,y;\lambda) \mathrm{d}s,$$

czyli $\dfrac{d(x,y;\lambda)}{d(\lambda)}$ jest rozwiązaniem równania rezolwenty (15).

■

Przykład 1. Niech $K_A = \sum_{i=1}^m u_i(x) v_i(y)$. Mamy w tym przypadku $K_A \begin{pmatrix} x_1 & \dots & x_n \\ y_1 & \dots & y_n \end{pmatrix} = 0$ dla $n > m$, czyli szeregi (17) i (18) są sumami pierwszych $m+1$ wyrazów.

Na przykład, rozpatrzmy równanie

$$f(x) + 2 \int_a^b (x+y) f(y) \mathrm{d}y = g(x),$$

czyli $K_A(x,y) = (x+y)$ i $\lambda = -2$. Tutaj

$$K_A \begin{pmatrix} x \\ y \end{pmatrix} = K_A(x,y) = x + y, \quad K_A \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} = \begin{vmatrix} x_1 + y_1 & x_1 + y_2 \\ x_2 + y_1 & x_2 + y_2 \end{vmatrix},$$

i stąd

$$d(\lambda) = 1 - \lambda \int_0^1 2y_1 \mathrm{d}y_1 + \frac{\lambda^2}{2!} \iint_0^1 (4y_1y_2 - (y_1+y_2)^2) \mathrm{d}y_1 \mathrm{d}y_2 = 1 - \lambda - \frac{\lambda^2}{12},$$

$$d(x,y;\lambda) = x + y - \lambda \int_0^1 (2y_1(x+y) - (x+y_1)(y_1+y)) \mathrm{d}y_1$$
$$= x + y - \lambda(x+y) + \lambda xy + \frac{1}{2}\lambda(x+y) + \frac{1}{3}\lambda.$$

W szczególności,

$$d(-2) = 1 + 2 - \frac{1}{3} = \frac{8}{3}, \quad d(x,y;-2) = 2(x+y) - 2xy - \frac{2}{3}.$$

Dostajemy stąd rezolwentę

$$\mathcal{R}(x,y) = \frac{3}{4}(x+y) - \frac{3}{4}xy - \frac{1}{4}$$

i rozwiązanie

$$f(x) = g(x) - 2 \int_0^1 \big(\frac{3}{4}(x+y) - \frac{3}{4}xy - \frac{1}{4}\big) g(y) \mathrm{d}y$$