Rozdział 1

Przestrzenie unitarne i przestrzenie Hilberta

Podajemy podstawowe własności iloczynu skalarnego: nierówność CBS (Cauchy'ego – Buniakowskiego – Schwarza), tożsamość równoległoboku oraz wzór polaryzacyjny, pojęcia prostopadłości wektorów ($x \perp y$) i dopełnienia ortogonalnego ($M^{\perp}$) podprzestrzeni.

1.1 Iloczyn skalarny

Definicja 1.1. **Przestrzenią unitarną** nad ciałem $\mathbb{K}(= \mathbb{R}$ lub $\mathbb{C})$ nazywamy parę $(V, \langle \cdot, \cdot \rangle)$, gdzie V jest przestrzenią wektorową nad ciałem $\mathbb{K}$, zaś odwzorowanie $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$, zwane **iloczynem skalarnym**, jest liniowe ze względu na pierwszą zmienną, „**skośnie symetryczne**" (tzn. $\langle y, x \rangle = \overline{\langle x, y \rangle}$) oraz **dodatnio określone** – czyli takie, że $\langle x, x \rangle > 0$ dla $x \in V \setminus \{0\}$.

W przypadku rzeczywistym ($\mathbb{K} = \mathbb{R}$) skośna symetria to po prostu symetria i wówczas iloczyn skalarny to odwzorowanie dwuliniowe symetryczne, dodatnio określone. Gdy $x = 0$ lub $y = 0$, mamy $\langle x, y \rangle = 0$, bo np. dla $x = 0$ można napisać $0 = tx$ dla $t = 0$, więc z liniowości względem pierwszej zmiennej, $\langle tx, y \rangle = t\langle x, y \rangle = 0$. Natomiast z równości $\langle x, y \rangle = 0$ nie wynika zerowanie się któregokolwiek z tych dwu wektorów.

Przykład 1.2. W przypadku $n < \infty$ mozna wykazać izomorfizm dowolnej n-wymiarowej przestrzeni unitarnej z tzw. przestrzenią euklidesową $(\mathbb{R}^n, \langle , \rangle)$, (lub $\mathbb{C}^n$ w przypadku zespolonym) gdzie iloczyn skalarny wektorów $\mathbf{x} = (x_1, \ldots, x_n)$ oraz $\mathbf{y} = (y_1, \ldots, y_n)$ określony jest wzorem

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1\bar{y}_1 + \ldots + x_n\bar{y}_n.$$

(Oczywiście, w przypadku rzeczywistym ($\mathbb{R}^n$) mamy $\bar{y}_k = y_k$.) Dla wektorów z bazy kanonicznej zero-jedynkowej $\mathbf{e}_j$ mamy $\langle \mathbf{e}_j, \mathbf{e}_k \rangle = \delta_{j,k}$ (symbol $\delta_{j,k}$ oznacza 1 dla $j = k$ oraz zero dla $j \neq k$).

Definicja normy: Przypomnijmy, że odwzorowanie przypisujące wektorom $\mathbf{x}$ przestrzeni V ich „długość": $\|\mathbf{x}\|$ nazywamy **normą**, gdy dla $\mathbf{x} \neq 0$ jest $\|\mathbf{x}\| > 0$, zaś dla dowolnych skalarów $\lambda \in \mathbb{K}$ i wektorów $\mathbf{x}, \mathbf{y} \in V$ zachodzą warunki:

$$\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\| \quad \text{(jednorodność) oraz} \quad \|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\| \quad \text{(nierówność trójkäta).}$$

Według definicji, ciąg wektorów $\mathbf{x}_n \in V$ ma granicę równą $\mathbf{x}_0$, co oznaczamy $\mathbf{x}_n \to \mathbf{x}_0$ lub $\mathbf{x}_0 = \lim \mathbf{x}_n$ gdy $\mathbf{x}_0 \in V$ oraz $\lim \|\mathbf{x}_n - \mathbf{x}_0\| = 0$. Ciąg zbieżny, to ciąg posiadający granicę w danej przestrzeni. W przypadku skalarnym (gdy $V = \mathbb{R}$ lub $V = \mathbb{C}$), normą jest moduł liczby i powyższa definicja pokrywa się ze zwykłym pojęciem granicy ciągu liczb. Zamieniając w definicji ciągu Cauchy'ego i w definicji granicy funkcji moduł na normę otrzymamy uogólnienia pojęć ciągu Cauchy'ego, granicy (i ciągłości) funkcji. Przestrzeń unormowaną nazwiemy **przestrzenią zupełną**, gdy każdy ciąg Cauchy'ego jest zbieżny. **Przestrzenią Banacha** nazywamy przestrzeń unormowaną zupełną. Można udowodnić, źe każda przestrzeń unormowana $(V, \| \|)$ zawiera sieę w pewnej przestrzeni zupełnej $(\tilde{V}, \| \|)$ takiej, że każdy punkt $x \in \tilde{V}$ jest granicą pewnego ciągu o wyrazach $x_n \in V$ (mówimy wtedy, że V jest gęstą podprzestrzenią w $\tilde{V}$, zaś przestrzeń $(\tilde{V}, \| \|)$ nazywamy **uzupełnieniem przestrzeni** $(V, \| \|)$).

1.2 Norma a iloczyn skalarny

Podobnie jak w przypadku przestrzeni euklidesowej $\mathbb{K}^n$, w dowolnej przestrzeni unitarnej definiujemy długość wektora $\mathbf{x}$ jako liczbę

$$\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}. \tag{$*$}$$

Liczba pierwiastkowana jest nieujemna, z postulatu dodatniej określoności $\langle \cdot, \cdot \rangle$. Natomiast z liniowości ze względu na pierwszą zmienną i skośnej symetrii, mamy $\langle \lambda \mathbf{x}, \lambda \mathbf{x} \rangle = \lambda \cdot \bar{\lambda} \langle \mathbf{x}, \mathbf{x} \rangle = |\lambda|^2 \|\mathbf{x}\|^2$, skąd wynika jednorodność długości wektora:

$$\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|.$$

Ponadto $\|\mathbf{x}\| > 0$ dla $x \in V \setminus \{0\}$, co również wynika wprost z definicji $\langle \cdot, \cdot \rangle$. Aby wykazać, że $\| \cdot \|$ jest normą potrzebna jest (dla dowodu nierówności trójkąta) następująca **nierówność Cauchy'ego – Buniakowskiego – Schwarza** mająca podstawowe znaczenie dla teorii przestrzeni unitarnych:

Lemat 1.3. (Cauchy, Buniakowski, Schwarz) *W przestrzeni unitarnej dla dowolnych wektorów x, y zachodzi nierówność*

$$|\langle x, y \rangle| \leqslant \|x\| \cdot \|y\|. \tag{CBS}$$

Dowód. Ustalmy dowolnie wektory $x, y \in V$. Możemy tu założyć, że $x \neq 0$. Dla wszystkich wartości parametru $t \in \mathbb{R}$ mamy

$$0 \leqslant \langle tx + y, tx + y \rangle = t^2 \|x\|^2 + 2t \mathrm{Re} \langle x, y \rangle + \|y\|^2.$$

Traktując to wyrażenie jako funkcję zmiennej $t \in \mathbb{R}$ – a mianowicie trójmian kwadratowy – wnioskujemy, że jego wyróżnik nie może być dodatni. Mamy więc

$$0 \geqslant (\mathrm{Re} \langle x, y \rangle)^2 - \|x\|^2 \|y\|^2,$$

co daje nierówność

$$|\mathrm{Re} \langle x, y \rangle| \leqslant \|x\| \|y\|.$$

Wystarczy udowodnić (CBS) w przypadku gdy $\lambda := \langle x, y \rangle \neq 0$. Stosując ostatnią nierówność zamiast do pary x, y – do pary wektorów $\alpha \cdot x, y$, gdzie $\alpha = \bar{\lambda} \cdot |\lambda|^{-1}$, otrzymamy

$$|\mathrm{Re} \langle \alpha \cdot x, y \rangle| = |\langle x, y \rangle| \leqslant \|\alpha \cdot x\| \|y\| = \|x\| \|y\|. \; \bullet$$

Wniosek 1.4. *Wzór* $(*)$ *określa normę. Nazywamy ją* **normą** *danej* **przestrzeni unitarnej** $(V, \langle\cdot,\cdot\rangle)$.

Dowód. Rzeczywiście, ze względu na nieujemność obu jej stron, nierówność trójkąta jest równoważna stwierdzeniu, że
$$\|x+y\|^2 \leqslant \|x\|^2 + 2\|x\|\|y\| + \|y\|^2. \tag{$**$}$$
Ponieważ
$$\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + 2\mathrm{Re}\langle x,y\rangle + \|y\|^2,$$
relację $(**)$ otrzymamy z nierówności (CBS).

Wniosek 1.5. *Iloczyn skalarny jest ciągły jako funkcja dwu zmiennych oraz jednostajnie ciągły na ograniczonych podzbiorach przestrzeni* $V \times V$.

Dowód. Ponieważ $\langle x,y\rangle - \langle x_0,y_0\rangle = \langle x-x_0, y_0\rangle + \langle x_0, y-y_0\rangle + \langle x-x_0, y-y_0\rangle$, ciągłość ta wynika z oszacowania
$$|\langle x_n,y_n\rangle - \langle x_0,y_0\rangle| \leqslant \|x_n-x_0\|\|y_0\| + \|x_0\|\|y_n-y_0\| + \|x_n-x_0\|\|y_n-y_0\|. \bullet$$

Funkcje jednostajnie ciągłe na dowolnej przestrzeni metrycznej posiadają (jednoznacznie określone) przedłużenia na jej uzupełnienie, gdyż przekształcają one ciągi fundamentalne (czyli spełniające warunek Cauchy'ego:
$$\forall \epsilon > 0 \exists_M \forall_{n,k>M} \quad \|x_n - x_k\| < \epsilon$$
) w ciągi fundamentalne. Można stąd wywnioskować, że uzupełnienie $\tilde{V}$ przestrzeni unitarnej V jest również przestrzenią unitarną. Iloczyn skalarny przestrzeni $\tilde{V}$ jest przedłużeniem ciągłym funkcji $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{K}$ na $\tilde{V} \times \tilde{V}$. (Jednostajną ciągłość tego iloczynu uzyskamy przy ustalonym jednym z wektorów i wówczas możemy dokonać przedłużenia ze względu na pierwszą zmienną. Potem rozszerzamy ze względu na drugą zmienną.)

Definicja 1.6. **Przestrzeń Hilberta** to przestrzeń unitarna, która jest zupełna względem normy wyznaczonej przez jej iloczyn skalarny, czyli w której wszystkie ciągi fundamentalne mają granice.

Twierdzenie 1.7. (Tożsamość równoległoboku) *W przestrzeni unitarnej normy dowolnych dwu wektorów* $x, y \in V$ *spełniają następującą równość*
$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2). \tag{PL}$$
Otrzymujemy ją dodając stronami równości $\|x \pm y\|^2 = \|x\|^2 \pm 2\mathrm{Re}\langle x,y\rangle + \|y\|^2$.

Odejmowanie stronami daje jako prawą stronę podwojoną część rzeczywistą z iloczynu skalarnego. W przypadku rzeczywistym z równości
$$\|x+y\|^2 - \|x-y\|^2 = 4\langle x,y\rangle$$
mając daną normę znajdujemy iloczyn skalarny. Również w przypadku zespolonym otrzymujemy w podobny sposób następujący wzór:

Lemat 1.8. (Wzór polaryzacyjny) (nie obowiązuje na kierunku Elektrotechnika

$$4\langle x, y\rangle = \|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2. \tag{P}$$

Można wykazać, że każda norma spełniająca tożsamość równoległoboku pochodzi od pewnego iloczynu skalarnego – a mianowicie iloczynu skalarnego zdefiniowanego przez (P). Sprawdzenie, że spełnione są postulaty z definicji wymaga użycia (PL).

Warto podkreślić, że w dowodach zarówno wzoru (PL), jak i (P) nie korzystamy nigdzie z założenia, że $\forall_{x\neq 0}\ \langle x, x\rangle > 0$ i wzór polaryzacyjny można stosować do form hermitowskich typu

$$\langle x, y\rangle_T := \langle Tx, y\rangle,$$

o ile tylko $T : H \to H$ jest operatorem (czyli odwzorowaniem) liniowym spełniającym warunek symetrii

$$\langle Tx, y\rangle = \langle x, Ty\rangle \qquad (\forall x, y \in H). \tag{SYM}$$

Bardzo pożytecznym wnioskiem jest zachodząca dla tego typu form (przy założeniu *symetrii* każdego z operatorów S, T) implikacja

$$[\langle x, x\rangle_T = \langle x, x\rangle_S\ (\forall x \in H)] \Rightarrow S = T.$$

Na zakończenie powróćmy do rzeczywistej przestrzeni Euklidesowej $\mathbb{R}^n$, gdzie iloczyn skalarny ma bardzo ważną interpretację geometryczną. Zauważmy, że norma euklidesowa $\|\mathbf{x}\| =: a$ przedstawia długość wektora $\mathbf{x}$, czyli odległość punktu $\mathbf{x} = (x_1, \ldots, x_n)$ od początku układu współrzędnych. Odległość punktu $\mathbf{x}$ od punktu $\mathbf{y}$ jest równa normie $\|\mathbf{x} - \mathbf{y}\| =: c$, przy czym $c^2 = \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y}\rangle = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\langle \mathbf{x}, \mathbf{y}\rangle$ z dwuliniowości rzeczywistego iloczynu skalarnego. Niech $b := \|y\|$. Zastosowanie wzoru kosinusów do trójkąta, którego bokami są wektory $\mathbf{x}$, $\mathbf{y}$ oraz $\mathbf{x} - \mathbf{y}$, daje

$$c^2 = a^2 + b^2 - 2ab \cos \angle(\mathbf{x}, \mathbf{y}).$$

Z porównania tych wzorów otrzymujemy następujący wniosek:

Wniosek 1.9. *W przestrzeni euklidesowej $\mathbb{R}^n$ zachodzi relacja*

$$\langle \mathbf{x}, \mathbf{y}\rangle = \|\mathbf{x}\|\|\mathbf{y}\| \cos \angle(\mathbf{x}, \mathbf{y}).$$

W szczególności, iloczyn skalarny dwu wektorów jest zerowy wtedy i tylko wtedy, gdy wektory te są prostopadłe. Natomiast

$$\angle(\mathbf{x}, \mathbf{y}) = \arccos \frac{\langle \mathbf{x}, \mathbf{y}\rangle}{\|x\|\|y\|} \ \text{ gdy } \mathbf{x}, \mathbf{y} \neq 0. \tag{$*\angle$}$$

Definicja 1.10. W przestrzeni unitarnej $(V, \langle\cdot,\cdot\rangle)$ **wektorami prostopadłymi** (albo ortogonalnymi) nazywamy wektory, których iloczyn skalarny jest zerem. Relację tę oznaczamy symbolem $\perp$, tzn. $x \perp y$ oznacza, że $\langle x, y\rangle = 0$. **Dopełnieniem ortogonalnym** zbioru $E \subseteq V$ nazywamy zbiór

$$E^\perp := \{x \in V;\ \langle x, e\rangle = 0\ (\forall e \in E)\}.$$

Zamiast $x \in E^\perp$ piszemy też $x \perp E$.

Uwaga 1.11. Warto w tym miejscu podkreslić, że dla nieeuklidesowych iloczynów skalarnych ta abstrakcyjna prostopadłość nie oznacza prostopadłości geometrycznej i podobnie umowny sens ma kąt między parą wektorów. Widać to wyraźnie na następującym przykładzie:

Przykład 1.12. Dla bazy kanonicznej $\mathbf{e}_1$, $\mathbf{e}_2$ w $\mathbb{R}^2$ (por. Przykł. 1.2) wektory

$$\mathbf{f}_1 := \mathbf{e}_1, \qquad \mathbf{f}_2 := \mathbf{e}_1 + \mathbf{e}_2$$

również tworzą bazę. Zamiast sumy iloczynów współrzędnych kartezjańskich, możemy dodawać iloczyny współrzędnych w tej nowej bazie. Współrzędne te określone są funkcjonałami f_j^* z bazy algebraicznie dualnej do $\{f_1, f_2\}$. Otrzymamy wówczas nowy iloczyn skalarny

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{f_1}^*(\mathbf{x})\mathbf{f_1}^*(\mathbf{y}) + \mathbf{f_2}^*(\mathbf{x})\mathbf{f_2}^*(\mathbf{y}).$$

Przy tym $\langle \mathbf{f_1}, \mathbf{f_2} \rangle = 0$, choć geometryczny kąt pomiędzy tymi wektorami zamiast $\frac{\pi}{2}$, wynosi $\frac{\pi}{4}$.

W sytuacjach gdy chcemy podkreślić, że nie chodzi nam o „prostopadłość geometryczną” (w $\mathbb{R}^n$), będziemy mówić o „prostopadłości algebraicznej” (lub prostopadłości w sensie iloczynu skalarnego). Łatwo przekonać się (ćw.), jakiej prostopadłości dotyczy następujące, znane nam skądinąd:

Twierdzenie 1.13. (Twierdzenie Pitagorasa) *Jeżeli jakieś dwa wektory $y, z \in H$ są prostopadłe, to*

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

Rzeczywiście, wynika to z równości

$$\|y + z\|^2 = \|y\|^2 + \|z\|^2 + 2\mathrm{Re}\langle y, z \rangle.$$

Rozdział 2

Rzut i rozkłady ortogonalne

Rozpoczynamy od podstawowego twierdzenia o rzucie i o rozkładzie ortogonalnym w przestrzeni Hilberta. Pokazujemy, w jaki sposób można sprawdzić, czy dany wektor jest rzutem ortogonalnym ustalonego wektora na podprzestrzeń (odpowiednio na zbiór wypukły domknięty). Podajemy charakteryzację określającą, kiedy dana projekcja jest ortogonalna.

2.1 Punkty najbliższe

Podstawowym pojęciem wyróżniającym przestrzenie unitarne spośród innych przestrzeni unormowanych jest prostopadłość. Kierunek prostopadły jest kierunkiem, w którym można „najkrótszą drogą dotrzeć do danego zbioru” w sensie sprecyzowanym poniżej. Osiągnięcie minimum odległości od zbioru zamiast zwartości wymaga słabszej (a dzięki temu dostępnej w przestrzeniach nieskończenie wymiarowych) własności kuli domkniętej, jaką jest zupełność. Do końca tego wykładu będziemy zakładać taką zupełność: H będzie oznaczać zawsze przestrzeń Hilberta.

Twierdzenie 2.1. *Dowolny niepusty zbiór wypukły i domknięty $E \subset H$ zawiera dokładnie jeden element x_0 minimalnej długości:*

$$(\exists!\, x_0 \in E) \quad \|x_0\| = \inf\{\|z\|;\ z \in E\}.$$

Dowód. Oczywiście, $+\infty > \delta = \inf\{\|z\|;\ z \in E\} \geqslant 0$ oraz dla pewnego ciągu wektorów $z_n \in E$ mamy $\delta = \lim \|z_n\|$. Dla $x, y \in E$ również $\frac{x+y}{2} \in E$, więc zachodzi nierówność

$$\|\frac{x+y}{2}\| \geqslant \delta. \tag{$*$}$$

Tożsamość równoległoboku dla wektorów $\frac{1}{2}z_n$ oraz $\frac{1}{2}z_m$ można zapisać w postaci

$$\frac{1}{4}\|z_n - z_m\|^2 = \frac{1}{2}\|z_n\|^2 + \frac{1}{2}\|z_m\|^2 - \|\frac{z_n + z_m}{2}\|^2.$$

Zastosowanie nierówności $(*)$ dla $x = z_n, y = z_m$ daje

$$0 \leqslant \frac{1}{4}\|z_n - z_m\|^2 \leqslant \frac{1}{2}\|z_n\|^2 + \frac{1}{2}\|z_m\|^2 - \delta^2.$$

Ponieważ prawa strona ostatniej nierówności zmierza przy $n \to \infty$, $m \to \infty$ do $\frac{1}{2}\delta^2 + \frac{1}{2}\delta^2 - \delta^2 = 0$, wynika stąd, że dla n, m dostatecznie dużych wartość $\|z_n - z_m\|$ jest dowolnie mała. Innymi słowy, ciąg $\{z_n\}$ jest ciągiem Cauchy'ego, a zatem, dzięki zupełności H –jest on zbieżny. Jego granica,

$x_0 = \lim z_n$ jest wektorem o normie δ (z ciągłości normy), należącym do E (z domkniętości zbioru E).

Gdy wektor y_0 ma również normę δ, to zamieniając w poprzednim oszacowaniu z_n na x_0, zaś z_m na y_0 otrzymamy $\|x_0 - y_0\| = 0$, co jest możliwe jedynie dla $x_0 = y_0$. •

Wniosek 2.2. (Twierdzenie o rzucie prostopadłym) *Gdy zbiór $F \subset H$ jest wypukły, domknięty i niepusty, to dla każdego wektora $y \in H$ istnieje dokładnie jeden element $z_0{:=}P_F y$ należący do F i minimalizujący odległość od tego punktu, tzn.*

$$\|y - z_0\| = \inf\{\|y - x\|;\, x \in F\} \ (= \mathrm{dist}(y, F)).$$

Dowód. Jeżeli $E = F - y$, to $\mathrm{dist}(y, F) = \min\{\|z\|;\, z \in E\} = \|x_0\|$ dla pewnego $x_0 \in E$. Wystarczy przyjąć $z_0 = y + x_0$, a wówczas $x_0 = z_0 - y$, $\|x_0\| = \mathrm{dist}(E, 0) = \mathrm{dist}(F, y)$. •

Definicja 2.3. Powyższy wektor $z_0{=}P_F y$ nazywamy **rzutem prostopadłym (projekcją ortogonalną)** wektora y na zbiór F. Określenia **rzut prostopadły** używamy też w stosunku do operatora $P_F : H \ni y \to P_F y \in H$

W przypadku gdy F jest domkniętą podprzestrzenią liniową, uzasadnienie nazwy można znaleźć w następnym twierdzeniu.

Przypomnijmy najpierw, że relacja prostopadłości wektora w do zbioru M zapisywana symbolem $w \perp M$, albo $w \in M^\perp$ oznacza zachodzenie równości $\langle w, m\rangle = 0\, (\forall m \in M)$.

Lemat 2.4. *Zbiór $M^\perp$ jest domkniętą podprzestrzenią liniową. Jeżeli $M_1 \subset M_2$, to $M_2^\perp \subseteq M_1^\perp$. Ponadto $\{0\}^\perp = H, H^\perp = \{0\}$.*

Dowód. Pierwsza teza wynika z liniowości i ciągłości funkcjonałów

$$\langle \cdot, y\rangle : H \ni x \to \langle x, y\rangle \in \mathbb{K}.$$

Sprawdźmy dla przykładu domkniętość $M^\perp$. Gdy $\lim_{n\to\infty} \|x_n - x\| = 0$, zaś $x_n \in M^\perp$, to wspomniana ciągłość daje dla dowolnego $y \in M$ relację

$$(0 =)\ \lim \langle x_n, y\rangle = \langle x, y\rangle,$$

więc również $x \in M^\perp$. Ostatnia teza wnika stąd, że gdy $x \perp H$, to $x \perp x$, ale to zachodzi jedynie dla $x = 0$. •

2.2 Kierunek prostopadły

Znaczenie następującego twierdzenia polega przede wszystkim na dostarczeniu efektywnego kryterium sprawdzającego, kiedy dany wektor y_0 jest rzutem prostopadłym (na zbiór F) ustalonego wektora y. Zwłaszcza w przypadku, gdy F jest podprzestrzenią afiniczną[1], zamiast sprawdzania zawartego w definicji (niestety, dość kłopotliwego w praktyce) warunku minimalizowania odległości, wystarczy zbadać prostopadłość wektora $y - y_0$ do związanej z F podprzestrzeni wektorowej $F - y_0$.

Twierdzenie 2.5

[1] podprzestrzeń afiniczna, to obraz podprzestrzeni liniowej przez przesunięcie równoległe.

(i) *Przy założeniach 11.2 wektor* $y_0 := P_F y$ *jest jedynym punktem ze zbioru* F *spełniającym dla każdego* $z \in F$ *nierówność* $\mathrm{Re}\langle y - y_0, y_0 - z\rangle \geqslant 0$.

(ii) *Gdy* $F = \bar{F}$ *jest domkniętą podprzestrzenią afiniczną, to* $y - y_0$ *jest prostopadły do* $F - y_0$.

(iii) *Dla podprzestrzeni liniowej domkniętej* $F \subset H \ni y$ *jeżeli* $y_0 \in F$ *oraz* $y - y_0 \perp F$, *to* $y_0 = P_F y$.

Geometryczny sens nierówności z tezy (i) jest następujący: zbiór wypukły F zawiera się w rzeczywistej półprzestrzeni domkniętej

$$\{z \in H;\, \mathrm{Re}\langle y - y_0, z\rangle \leqslant \alpha\}, \quad \text{dla} \quad \alpha := \mathrm{Re}\langle y - y_0, y_0\rangle,$$

zaś punkt $y_0 \in F$ leży na jej brzegu (czyli na hiperpłaszczyźnie rzeczywistej $\{z \in H;\, \mathrm{Re}\langle y - y_0, z\rangle = \alpha\}$. Takie **hiperpłaszczyzny** nazywamy **podpierającymi** zbiór F w punkcie y_0. Innymi słowy, rzut $P_F y$ wektora y na F jest **punktem podparcia zbioru** F pewną rzeczywistą hiperpłaszczyną afiniczną prostopadłą do $y - y_0$.

Przed przystąpieniem do dowodu warto sformułować najważniejszy wniosek z tezy (ii) powyższego twierdzenia (i z ciągłości projekcji, którą sprawdzimy za chwilę).

Wniosek 2.6. (Twierdzenie o rozkładzie ortogonalnym) *Gdy* M *jest domkniętą podprzestrzenią liniową przestrzeni Hilberta* H, *to dowolny wektor* $x \in H$ *ma jednoznaczny rozkład w postaci sumy*

$$x = x_1 + x_2, \quad x_1 \in M,\ x_2 \perp M,$$

przy czym $x_1 = P_M x$, *zaś* $x_2 = x - P_M x (= P_{(M^\perp)} x)$. *Przestrzeń* H *jest więc sumą prostą topologiczną pary swoich podprzestrzeni:* $H = M + N$, *gdzie* $N \perp M$ *(a mianowicie* $N = M^\perp$*).*

Sumy proste tego typu nazywamy **sumami ortogonalnymi** *i oznaczamy ten fakt pisząc*

$$H = M \oplus N \quad \text{zamiast} \quad M + N.$$

Można nawet wykazać dla dowolnej przestrzeni Banacha X, że *istnienie dla każdej jej podprzestrzeni domkniętej* M *podprzestrzeni domkniętej* N *tworzącej wraz z* M *sumę prostą topologiczną równą* X *implikuje izomorficzną równoważność normy* X *z pewną normą hilbertowską.*

Dowód Twierdzenia 11.5. (i) Ustalmy dowolnie $z \in F$. Funkcja zmiennej $t \in \mathbb{R}$ określona wzorem

$$\psi(t) := \|y - (tz + (1-t)y_0)\|^2 = \|y - y_0\|^2 + 2t\mathrm{Re}\langle y - y_0, y_0 - z\rangle + t^2\|y_0 - z\|^2$$

jest wielomianem, który na odcinku $[0,1]$ osiąga minimum w punkcie $t = 0$, gdyż dla $t \in [0,1]$ mamy $tz + (1-t)y_0 \in F$, zaś $\psi(0) = \|y - y_0\|^2$. Stąd

$$0 \leqslant \psi'(0) = 2\mathrm{Re}\langle y - y_0, y_0 - z\rangle,$$

co daje postulowaną nierówność. Z drugiej stony, nierówność w (i) oznacza, że $\psi'(0) \geqslant 0$, a ponieważ $\psi''(t) = 2\|y_0 - z\|^2 \geqslant 0$, implikuje ona nierówności

$$\psi'(t) \geqslant 0 \quad \text{dla} \quad t \in [0,1].$$

Stąd już wynika warunek minimalizowania przez y_0 odległości, gdyż

$$\|y - y_0\|^2 = \psi(0) \leqslant \psi(1) = \|y - z\|.$$

(ii) W tym przypadku $tz + (1-t)y_0 \in F$ również dla $t < 0$, więc $\psi(\cdot)$ osiąga minimum lokalne (a nawet globalne) w punkcie $t = 0$, stąd $\psi'(0) = 0$. Teza (iii) jest szczególnym przypadkiem „tezy odwrotnej" w punkcie (i), gdyż mamy wówczas $F - y_0 = F$. •

W praktyce, aby wykazać, że jakiś wektor jest rzutem danego wektora, zamiast definicji używamy prawie zawsze warunku (iii) (lub (i), w zależności od tego, czy rzutujemy na podprzestrzeń, czy na zbiór wypukły).

Przykład 2.7. Ustalmy wektor $x_0 \neq 0$ i znajdźmy projekcję ortogonalną y_0 wektora y na podprzestrzeń (domkniętą) $L := \mathbb{K} \cdot x_0$. Ponieważ jako element zbioru L, wektor y_0 musi być postaci cx_0, wystarczy znaleźć wartość skalara c. Z warunku $y - y_0 \perp x_0$ (tu równoważnego relacji $y - y_0 \perp L$) mamy

$$0 = \langle y - cx_0, x_0 \rangle = \langle y, x_0 \rangle - c\langle x_0, x_0 \rangle.$$

Warunek ten jest równoważny równości $c = \frac{\langle y, x_0 \rangle}{\langle x_0, x_0 \rangle}$, czyli

$$P_L y = \frac{\langle y, x_0 \rangle}{\langle x_0, x_0 \rangle} x_0. \tag{$**$}$$

Gdy $\|x_1\| = \frac{x_0}{\|x_0\|}$, również $L = \mathbb{K} \cdot x_1$, mamy więc

$$P_L y = \langle y, x_1 \rangle x_1.$$

Przykład 2.8. Gdy zbiór F jest kulą domkniętą

$$\bar{K}(a, r) = \{x \in H;\ \|x - a\| \leqslant r\},$$

to

$$P_F y = a + \frac{r}{\|y - a\|} \cdot (y - a)$$

dla $y \notin F$. Oczywiście, dla $y \in F$ musi być $P_F y = y$. W przypadku $a = 0$ sprawdzimy warunek (i) korzystając z nierówności Schwarza. (ćw.)

W przypadku ogólnym można wykorzystać „**przesuwalność rzutów**":

$$P_{F+a} y = a + P_F(y - a).$$

Przykład 2.9. Gdy $H = L^2[0, 1]$, to dla ustalonego zbioru mierzalnego $A \subset [0, 1]$ rzut na podprzestrzeń liniową

$$F = \{f \in H;\ f(t) = 0 \ \text{ dla p.w. } t \in [0, 1] \setminus A\}$$

jest operatorem mnożenia przez funkcję charakterystyczną zbioru A, czyli

$$(P_F f)(t) = \chi_A(t) f(t) \quad \text{dla p.w. } t \in [0, 1].$$

2.3 Projekcje jako operatory

Twierdzenie 2.10. *Załóżmy, że $F \subset H$ jest domkniętą podprzestrzenią wektorową. Operator $P_F : H \to H$ rzutu prostopadłego na F ma następujące własności:*

(i) *Odwzorowanie $P_F : H \ni y \to P_F y \in H$ jest liniowe.*

(ii) $F = \{x \in H;\, x = P_F x\} = P_F(H) := \{P_F y;\, y \in H\}$.

(iii) *Operator $P = P_F$ spełnia warunki: $P \circ P = P$ oraz $\|P\| \leqslant 1$, czyli jest operatorem idempotentnym oraz kontrakcją.*

(iv) *Dowolny operator liniowy $P : H \to H$ idempotentny i kontrakcyjny jest rzutem prostopadłym na podprzestrzeń $\{x \in H : x = Px\} = \ker(I - P)$.*

Dowód. (i) wynika z Twierdzenia 11.5 (iii): Aby sprawdzić, że

$$z := \alpha P_F x + \beta P_F y = P_F(\alpha x + \beta y),$$

wystarczy upewnić się, że po pierwsze:

$$z \in F$$

(oczywiste, bo z jest liniową kombinacją wektorów z podprzestrzeni wektorowej F) i po drugie, że wektor $z - (\alpha x + \beta y)$ jest prostopadły do tej podprzestrzeni, czyli

$$z - (\alpha x + \beta y) \perp F.$$

Ale ten wektor jest równy $(\alpha(P_F x - x) + \beta(P_F y - y))$ i jako kombinacja liniowa wektorów prostopadłych do F, jest rzeczywiście prostopadły do F. Wynika to z liniowości iloczynu skalarnego względem pierwszej zmiennej i z definicji prostopadłości.

Teza (ii) wynika z równości $P_F x = x$ dla $x \in F$ i z relacji $P_F H \subseteq F$. Analogicznie wykazujemy zawartą w (iii) tezę o $P \circ P$. Wykażmy kontrakcyjność. Mamy

$$\|P_F x\|^2 = \langle P_F x, P_F x\rangle = \langle P_F x, x\rangle \leqslant \|P_F x\|\|x\|.$$

Ostatnia nierówność jest nierównością (CBS), zaś poprzedzająca ją równość wynika z prostopadłości $x - P_F x$ do F – a więc i do wektora $P_F x \in F$. Gdy (nietrywialny przypadek) $P_F x \neq 0$, dzieląc strony tej nierówności przez $\|P_F x\|$ otrzymamy $\|P_F\| \leqslant 1$. Jeszcze prościej można wykazać, że $\|P_F x\|^2 \leqslant \|x\|^2$, stosując do sumy $x = P_F x + P_{F^\perp} x$ twierdzenie Pitagorasa (Twierdzenie 10.15).
(iv): Z algebry liniowej wiadomo, że P oraz $Q = I - P$ są rzutami na składniki sumy prostej $H = M + Q(H)$, gdzie $M = P(H)$, $Q(H) = \ker P$ oraz $M \cap Q(H) = \{0\}$. Dla $z \perp Q(H)$ niech $x = Pz$, $y = Qz$. $\|P\| \leqslant 1$, więc $\|z\|^2 \geqslant \|Pz\|^2 = \langle z - y, z - y\rangle = \|z\|^2 - 2\mathrm{Re}\langle z, y\rangle + \|y\|^2$. Środkowy składnik tej sumy znika, gdyż $z \perp y$.

Otrzymana nierówność może więc zachodzić jedynie przy $y = 0$, stąd dla $z \in A := Q(H)^\perp$ mamy $Pz = z = P_A z$. Dla $z \in A^\perp = Q(H)$ również $Pz = 0 = P_A z$. Operatory liniowe P oraz P_A równe na zbiorze $A \cup A^\perp$ muszą pokrywać się i na obwiedni liniowej tego zbioru. Ale tą obwiednią jest H. •

Uwaga 2.11. Można stosunkowo łatwo wykazać, że gdy operator liniowy ciągły P jest **idempotentny**, tzn. gdy $P = P^2 \in B(H)$, to warunek $\|P\| \leqslant 1$ w punkcie (iv) Twierdzenia 11.10 można zastąpić przez postulat zachodzenia $(\forall x, y \in H)$ jednego z równoważnych wówczas warunków:

(a) $\langle Px, Py\rangle = \langle Px, y\rangle$,

(b) $\langle x, Py\rangle = \langle Px, y\rangle$.

Definicja 2.12. Projekcje ortogonalne P_A, P_B na domknięte podprzestrzenie $A, B \subset H$ nazwiemy wzajemnie ortogonalnymi, co oznaczymy symbolem $P_A \perp P_B$, gdy $P_A P_B = 0$, lub – co na jedno wychodzi – gdy $B \subset \ker P_A = A^{\perp}$, albo $A \perp B$.

Na zakończenie podajmy dość wygodne (i często stosowane) kryterium aproksymowalności danego wektora ciągiem kombinacji liniowych elementów ustalonego zbioru $M \subset H$ (tzn. warunek na należenie do zbioru $\overline{\mathrm{span}} M$):

Twierdzenie 2.13. *Dla dowolnego podzbioru M przestrzeni Hilberta*

$$\overline{\mathrm{span}} M = M^{\perp\perp},$$

czyli wektor y jest granicą (w topologii normy) ciągu pewnych kombinacji liniowych wektorów ze zbioru M wtedy i tylko wtedy, gdy z relacji: $z \perp M$ wynika prostopadłość: $z \perp y$. W szczególności, M jest podzbiorem „liniowo gęstym” przestrzeni H, czyli $\overline{\mathrm{span}} M = H$ wtedy i tylko wtedy, gdy jedynym wektorem prostopadłym do tego zbioru jest wektor 0.

Dowód. Niech $M_0 = \overline{\mathrm{span}} M$. Jak łatwo sprawdzić, jest to domknięta podprzestrzeń przestrzeni $H_0 := M^{\perp\perp}$. Gdyby inkluzja była ostra, to dokonując rozkładu ortogonalnego przestrzeni Hilberta H_0 postaci $H_0 = M_0 + N_0$ znajdziemy $y \neq 0, y \in H_0$ taki, że $y \perp M_0$. Tym bardziej, $y \perp M$, czyli $y \in M^{\perp}$. Ale $y \in H_0$ oznacza dokładnie tyle, że $y \perp M^{\perp}$, więc otrzymujemy stąd prostopadłość: $y \perp y$ możliwą jedynie dla $y = 0$, co stoi w sprzeczności z wyborem wektora y. •

W niektórych tekstach można znaleźć oznaczenie $A \ominus B$ dla podprzestrzeni $A \cap (B^{\perp})$. W szczególności, oznaczenie $H \ominus B$ zamiast $B^{\perp}$ bywa korzystniejsze, bo zawiera informację o tym, w jakiej przestrzeni rozważamy uzupełnienie ortogonalne podprzestrzeni B.

Rozdział 3

Funkcjonały, bazy ortogonalne

Poznamy jedno z ciekawszych zastosowań rozkładu ortogonalnego: twierdzenie Riesza – Frécheta o postaci funkcjonału. Pozwala ono na definicje: operatora sprzężonego $T^* \in \mathcal{B}(H)$ do operatora $T \in \mathcal{B}(H)$, operatorów samosprzężonych i normalnych. Pojawiają się też (charakterystyczne dla ciągów ortonormalnych) pojęcie słabej zbieżności, nierówność Bessela i tożsamość Parsevala.

3.1 Funkcjonały na przestrzeni H

Dzięki technice rozkładów ortogonalnych można uzyskać znacznie więcej informacji o wektorach i funkcjonałach w przestrzeni Hilberta H, niż w przypadku przestrzeni Banacha X. Przypomnijmy na przykład, że przestrzeń sprzężona X' (przestrzeń funkcjonałów liniowych ciągłych na X) może znacznie różnić się od X. W szczególności, ośrodkowość X (dla $X = \ell^1$) nie musi implikować ośrodkowości przestrzeni sprzężonej: (tu $X' = \ell^\infty$). W przypadku przestrzeni Hilberta okazuje się, że w pewnym sensie, $H' = H$. Mówi o tym następujące **twierdzenie o postaci funkcjonału**:

Twierdzenie 3.1. (F. RIESZ, M. FRÉCHET) *Jeżeli φ jest funkcjonałem liniowym i ciągłym na przestrzeni Hilberta H, to istnieje dokładnie jeden wektor $z_0 \in H$, dla którego*

$$\varphi(x) = \langle x, z_0 \rangle \qquad (\forall x \in H).$$

Dowód. Wystarczy ograniczyć się do przypadku $\varphi \neq 0$. Niech y_0 będzie takim wektorem prostopadłym do $M := \ker\varphi$, że $\varphi(y_0) = 1$. Dla wszystkich $x \in H$ mamy

$$x - \varphi(x) y_0 \in M, \tag{†}$$

dzięki liniowości φ. Ponieważ $y_0 \perp M$, mamy

$$0 = \langle x - \varphi(x) y_0, y_0 \rangle = \langle x, y_0 \rangle - \varphi(x) \langle y_0, y_0 \rangle.$$

Warunki tezy spełnia więc wektor $z_0 = \|y_0\|^{-2} y_0$. Innymi słowy, $\varphi = \langle \cdot, z_0 \rangle$.

Zauważmy, że istnieje dokładnie jeden taki $y_0 \in M^\perp$, że $\varphi(y_0) = 1$. Istotnie, kowymiar ($= \dim(H \ominus M)$) przestrzeni M wynosi 1, co wynika z (†). Po drugie, jeśli istnieje wektor z_0 spełniający warunek z tezy, to jest on prostopadły do $\ker\varphi$ i musi on być postaci $c \cdot y_0$. Musi wtedy być $c = \|y_0\|^{-2}$, co dowodzi jednoznaczności. •

Warto podkreślić, że odwzorowanie: $H \ni z \to \varphi_z := \langle \cdot, z \rangle \in H'$, chociaż izometryczne i bijektywne, jest w przypadku zespolonym antyliniowe:

$$\varphi_{\alpha z + \beta w} = \bar{\alpha}\varphi_z + \bar{\beta}\varphi_w \qquad \alpha, \beta \in \mathbb{K}.$$

Na przestrzeni sprzężonej można więc określić iloczyn skalarny wzorem

$$\langle \varphi_z, \varphi_w \rangle = \overline{\langle z, w \rangle}$$

i norma określona przez ten iloczyn skalarny pokrywa się z normą funkcjonału, tzn. $\|\varphi_z\| = \|z\|$ – co wynika z nierówności (CBS). Izometryczność bijekcji $H \to H''$ wynika wprost z twierdzenia Riesza – Frécheta, bez potrzeby odwoływania się do twierdzenia Hahna – Banacha. Bijekcja ta jako złożenie dwu bijekcji antyliniowych: $H \to H'$ oraz $H' \to H''$ jest odwzorowaniem liniowym. Co ważniejsze, pokrywa się ono z zanurzeniem kanonicznym $j : H \to H''$, gdzie $j(x)(\varphi) := \varphi(x)$, gdy $x \in H$, $\varphi \in H'$ (por. podrozdziały 9.6 oraz 9.10), a więc $j(H) = H''$. Innymi słowy, przestrzenie Hilberta są refleksywne.

Definicja 3.2. Gdy $T \in B(H)$ jest operatorem liniowym ciągłym na przestrzeni H, funkcjonał $\varphi_y \circ T : H \ni x \mapsto \langle Tx, y \rangle$ jest liniowy i ciągły, a więc musi być on postaci φ_z dla pewnego $z := T^*y \in H$. Innymi słowy, T^*y jest jedynym takim wektorem, dla którego zachodzą równości:

$$\langle Tx, y \rangle = \langle x, T^*y \rangle \qquad (\forall x \in H).$$

Operator (liniowy) T^* nazywamy **sprzężonym** do T. **Operator** S nazywamy **samosprzężonym**, gdy $S = S^*$ zaś **normalnym**, gdy $SS^* = S^*S$.

3.2 Ciągi ortogonalne, bazy

Ważnym zastosowaniem techniki rozkładów i rzutów ortogonalnych jest konstrukcja baz ortonormalnych.

Definicja 3.3. Rodzinę wektorów $\{a_j\}_{j \in J} \subset H$ nazwiemy:
ortogonalną, gdy $a_j \perp a_i$ $(\forall j \neq i)$,
ortonormalną, gdy $\langle a_i, a_j \rangle = \delta_{i,j}$ dla wszystkich $i, j \in J$,
bazą ortonormalną (odpowiednio **ortogonalną**), gdy wektory te tworzą zbiór ortonormalny (ortogonalny), który jest liniowo gęsty: $\overline{\mathrm{span}}\{a_j\}_{j \in J} = H$.

W ogólniejszej sytuacji ośrodkowych przestrzeni Banacha X – nie dysponując relacją ortogonalności, sformułować można następującą definicję.

Definicja 3.4. Ciąg wektorów $f_n \in X$ jest **bazą Schaudera** w przestrzeni X, gdy dla każdego wektora $x \in X$ istnieje dokładnie jeden ciąg współczynników skalarnych λ_n $(\overset{\text{ozn}}{=} f_n^*(x))$ taki, że

$$x = \sum_{n=1}^{\infty} \lambda_n f_n \quad (\text{tzn. } \lim_{k \to \infty} \|x - \sum_{n=1}^{k} \lambda_n f_n\| = 0\,).$$

Wkrótce przekonamy się, że baza ortonormalna $\{e_n\}$ jest w szczególności bazą Schaudera. W tym miejscu zauważmy, że gdy $x = \sum_{n=1}^{\infty} \lambda_n e_n$, to z liniowości i ciągłości funkcjonału $\langle \cdot, e_k \rangle$, mamy

$$\lambda_k \,(= \sum_{n=1}^{\infty} \lambda_n \delta_{n,k}) = \sum_{n=1}^{\infty} \lambda_n \langle e_n, e_k \rangle = \langle x, e_k \rangle.$$

Układy ortogonalne są więc liniowo niezależne. Nie każdy układ liniowo niezależny wyznacza jednoznacznie współczynniki λ_n – co widać, na następującym przykładzie: Do układu ortonormalnego $\{\epsilon_n;\, n \in \mathbb{N}\} \subset \ell^2$, gdzie ϵ_n jest ciągiem, którego k-ty wyraz wynosi $\delta_{n,k}$ dołączmy wektor ϵ_∞ będący ciągiem $\{\frac{1}{k}\}$. Z jednej strony, $\epsilon_\infty = \sum_1^\infty \frac{1}{k} \cdot \epsilon_k$, a z drugiej, $\epsilon_\infty = 1 \cdot \epsilon_\infty$. Przyczyna leży w tym, że badając liniową niezależność używamy jedynie kombinacji liniowych skończonej ilości elementów. Co więcej, w nieskończenie wymiarowej przestrzeni Banacha każda baza algebraiczna jest nieprzeliczalna, co można wywnioskować z twierdzenia Baire'a (ćw.).

Choć może się to wydawać rzeczą dziwną, istnieją ośrodkowe przestrzenie Banacha nieposiadające bazy Schaudera. W innych, bardzo naturalnych przestrzeniach Banacha szukanie baz Schaudera, choć zakończone sukcesem, było na tyle trudne, że zajęło matematykom wiele lat.

W przestrzeni Hilberta (np. w 2-wymiarowej przestrzeni euklidesowej $\mathbb{R}^2$) większość baz nie jest złożona z wektorów ortogonalnych. Tym niemniej, istnieje bardzo prosty algorytm konstrukcji baz ortonormalnych w dowolnej przestrzeni Hilberta. Zacznijmy jednak od pewnego ważnego lematu.

Lemat 3.5. (Nierówność Bessela) *Dla ciągu ortonormalnego $\{e_i\}_{i=1}^\infty$ zdefiniujmy P_n jako projekcję ortogonalną P_{H_n} na podprzestrzeń H_n rozpiętą na wektorach e_j; $j \leqslant n$. Wówczas dla $x \in H$, $1 \leqslant n < m < \infty$ mamy*

(i)
$$P_n x = \sum_{j=1}^{n} \langle x, e_j \rangle e_j,$$

(ii)
$$\|P_m x\|^2 = \sum_{1}^{m} |\langle x, e_j \rangle|^2 \quad \text{oraz} \quad \|P_m x - P_n x\|^2 = \sum_{j=n+1}^{m} |\langle x, e_j \rangle|^2,$$

$$\sum_{j=1}^{\infty} |\langle x, e_j \rangle|^2 \leqslant \|x\|^2 \,(< +\infty). \tag{NB}$$

Dowód. Wektor y stojący po prawej stronie (i) należy do H_n, więc stosując charakteryzację rzutu (11.5 (iii)), wystarczy sprawdzić, że $x - y \perp H_n$, czyli że $\langle x - y, e_j \rangle = 0\, (\forall j \leqslant n)$. Ale dla $j \leqslant n$ mamy $\langle e_j, e_k \rangle = \delta_{j,k}$, więc

$$\langle x - y, e_j \rangle = \langle x, e_j \rangle - \sum_{k=1}^{n} \langle x, e_k \rangle \langle e_j, e_k \rangle = 0.$$

Dowód (ii) można przeprowadzić indukcją ze względu na n. Teza jest oczywista, gdy $m = 1$. Ponieważ składniki sumy $P_m x = P_{m-1} x + \langle x, e_m \rangle e_m$ są nawzajem prostopadłe, stosując twierdzenie Pitagorasa ($\|u + w\|^2 = \|u\|^2 + \|w\|^2$ dla $w \perp u$), mamy $\|P_m x\|^2 = \|P_{m-1} x\|^2 + |\langle x, e_m \rangle|^2 \|e_m\|^2$, co daje pierwszy wzór. Drugi wynika z pierwszego, jeśli zauważyć (np. stosując (i)), że $P_m x - P_n x$ jest rzutem ortogonalnym wektora x na podprzestrzeń $\mathrm{span}\{e_{n+1}, \ldots, e_m\}$.
(NB) wynika z (ii) po przejściu granicznym ($\lim\limits_{m \to \infty}$), ponieważ $\|P_m x\| \leqslant \|x\|$. •

Będziemy rozważać jedynie ośrodkowe przestrzenie Hilberta, ale warto nadmienić, że powyższe rozumowanie bez trudu przenosi się na dowolne układy ortonormalne $\{e_j;\ j \in J\}$. Nierówność (NB) oznacza sumowalność rodziny liczb, $|\langle x, e_j\rangle|^2$, skąd łatwo wywnioskować, że dla ustalonego $x \in H$ istnieje przeliczalny podzbiór $J_x \subseteq J$ taki, że $x \perp e_j$ dla $j \in J \setminus J_x$.

Ustalmy obecnie dowolny ciąg liniowo niezależnych wektorów

$$x_n;\ n < N+1,\ x_n \in H,$$

gdzie $N \leqslant \dim(H) \leqslant +\infty,\ \infty + 1 = \infty$. Niech $H_k = \operatorname{span}\{x_j;\ j \leqslant k\}$.

Twierdzenie 3.6. (Gram, Schmidt) *Dowolny układ liniowo niezależny można „wyprostować”: istnieje ciąg ortonormalny $\{e_n\}_{n<N+1}$ o tej własności, że $\operatorname{span}\{x_j;\ j \leqslant k\} = \operatorname{span}\{e_j;\ j \leqslant k\}$ dla wszystkich $k < N+1$. Gdy ponadto zbiór $\{x_n;\ n < N+1\}$ jest liniowo gęsty, to otrzymany w tym procesie ciąg $\{e_n;\ n < N+1\}$ tworzy bazę ortonormalną w przestrzeni H.*

Dowód. Dowód polega na konstrukcji algorytmu (w oparciu o wzór na $P_n := P_{H_n}$). Niech $e_1 = \|x_1\|^{-1}x_1$. Oczywiście, $\mathbb{K} \cdot e_1 = \mathbb{K} \cdot x_1$. Wektor $y_2 = x_2 - P_1 x_2$ jest prostopadły do e_1. Dla $e_2 = \|y_2\|^{-1}y_2$ – para e_1, e_2 tworzy bazę ortonormalną przestrzeni H_2. Kontynuując, gdy mamy już bazę ortonormalną $\{e_1, \ldots, e_{n-1}\}$ w podprzestrzeni H_{n-1}, definiujemy $y_n = x_n - P_{n-1}x_n$. Z liniowej niezależności x_j wynika, że $x_n \notin H_{n-1}$, więc $y_n \neq 0$ i można zdefiniować $e_n = \|y_n\|^{-1}y_n$. Oczywiście, przy ustalonym n wektory $e_1, \ldots, e_n$ tworzą bazę ortonormalną przestrzeni H_n. Na zakończenie zauważmy, że założenie o liniowej gęstości oznacza gęstość podprzestrzeni $\bigcup_n H_n$ w H, skąd otrzymamy liniową gęstość w H układu ortonormalnego $\{e_n\}$. •

Algorytm **ortonormalizacji** (i **ortogonalizacji**) jest więc następujący:

$$y_1 = x_1, \qquad e_1 = \frac{y_1}{\|y_1\|},$$

$$y_{n+1} = x_{n+1} - \sum_{j=1}^{n} \langle x_{n+1}, e_j\rangle e_j, \qquad e_{n+1} = \frac{y_{n+1}}{\|y_{n+1}\|}.$$

Istnieją też bezpośrednie wzory na $e_j, j \leqslant n$ w terminach samych wektorów x_j, $j \leqslant n$. W tym celu używa się tzw. **macierzy Grama** G dla układu n wektorów: $x_1, \ldots, x_n$. Definiujemy ją jako macierz $G \in M(n \times n)$, której wyrazami są $a_{jk} := \langle x_j, x_k\rangle$. Można w szczególności wykazać, że wektory x_j są liniowo niezależne wtedy i tylko wtedy, gdy macierz G jest nieosobliwa, tzn. $\det(G) \neq 0$. Na przykład, w n-wymiarowej przestrzeni euklidesowej $\mathbb{K}^n$ z kanonicznym iloczynem skalarnym niech W będzie macierzą $n \times n$, której j-ty wiersz składa się ze współrzędnych wektora x_j, zaś W^* – jej hermitowskie sprzężenie. Wtedy $G = W \cdot W^*$, a nieosobliwość G jest równoważna nieosobliwości W, która oznacza właśnie liniowa niezależność tych wektorów.

Uwaga 3.7. Równość $P_m w = c_1x_1 + \ldots + c_m x_m$, określająca współczynniki ortonormalizacji, jeśli $x = x_{m+1}$ oznacza np. w przypadku rzeczywistym osiąganie przez funkcję n zmiennych

$$\psi(\alpha_1, \ldots, \alpha_n) = \|w - \sum_{1}^{m} \alpha_j x_j\|^2 = \|w\|^2 + \sum_{k=1}^{n} \left(-2\alpha_k\langle w, x_k\rangle + \sum_{j=1}^{n} \alpha_j\alpha_k\langle x_j, x_k\rangle\right)$$

minimum absolutnego w punkcie $(\alpha_1, \ldots, \alpha_n) = (c_1, \ldots, c_n)$, a ponieważ w takim punkcie $\frac{\partial\psi}{\partial\alpha_k} = 0$, otrzymujemy stąd układ (Cramera) n równań postaci

$$c_1\langle x_1, x_k\rangle + \ldots + c_n\langle x_n, x_k\rangle = \langle w, x_k\rangle \qquad 1 \leqslant k \leqslant n. \tag{$*$}$$

Macierzą tego układu jest macierz Grama G i wzory Cramera dają szukane współczynniki c_j. Metoda ta pozwala na znajdowanie rzutu ortogonalnego wektora w na podprzestrzeń n-wymiarową rozpiętą na wektorach $x_1, \ldots, x_m$. Równania $(*)$ wynikają też z relacji $w - P_m w \perp x_k$.

Przykład 3.8. (Wielomiany ortogonalne). W wielu dziedzinach matematyki i jej zastosowań istotną rolę odgrywają układy ortonormalne uzyskane dzięki ortonormalizacji ciągu jednomianów t^n rozważanych jako elementy przestrzeni Hilberta $L^2((a,b), \rho(t)dt)$, gdzie $\rho(t)dt$ oznacza miarę zadaną na odcinku $(a;b)$ przez pewną funkcję gęstości ρ względem miary Lebesgue'a dt. Z postaci algorytmu ortonormalizacji wynika, że są to wielomiany. Gdy np. $a = -1, b = 1$, zaś $\rho(t) = 1\,(\forall t)$, to otrzymujemy tzw. **wielomiany Legendre'a** : $L_n(x)$. Gęstości $\rho(x) = \exp(-\frac{x^2}{2})$ na $(-\infty, +\infty)$ odpowiadają **wielomiany Hermite'a**, $H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n}(e^{-x^2/2})$.

Powróćmy do ogólnych pytań: Kiedy nierówność Bessela (NB) jest ostra? Kiedy układ ortonormalny nie jest bazą?

Okazuje się, że odpowiedź można zawrzeć w symbolu $\Leftrightarrow$:

Twierdzenie 3.9. (Tożsamość Parsevala) *Dla ciągu ortonormalnego* $\{e_n;\, n < \infty\}$ *następujące warunki są równoważne:*

(i) *Układ ten stanowi bazę przestrzeni* H,

(ii) *w nierówności (NB) mamy równość:*

$$\|x\|^2 = \sum_{j=1}^{\infty} |\langle x, e_j\rangle|^2 \qquad (\forall x \in H), \tag{TP}$$

(iii) $x = \lim_n P_n x \qquad (\forall x \in H)$.

Dowód. Warunek (i) oznacza, że podprzestrzeń $K := \operatorname{span}\{e_j;\, n < \infty\}$ jest równa H. W przeciwnym przypadku istniałby wektor $z \in H \ominus K \setminus \{0\}$. Z relacji $z \perp e_j\,(\forall e_j)$ wynika zerowanie się prawej strony (NB) (oraz (TP)), w odróżnienu od lewej. Natomiast gdy $K = H \ni x$, to ciąg $P_n x$ jest ciągiem Cauchu'ego, co wynika z (NB) stosowanej do $\|P_m x - P_n x\|$. Jeżeli $y = \lim P_n x$, to $\langle x - y, e_j\rangle = 0\ (\forall j)$, skąd wynika równość $x = y$, czyli (iii). Na koniec, z równości (TP) wynika liniowa gęstość zbioru $\{e_j; j < \infty\}$, gdyż jedynym wektorem doń prostopadłym jest 0. •

Wnioskiem z dotychczas poznanych własności baz jest twierdzenie mówiące, że *z dokładnością do unitarnej równoważności jest tylko jedna ośrodkowa przestrzeń Hilberta:* ℓ^2.

Twierdzenie 3.10. (Riesz, Fischer). *Ustalmy bazę ortonormalną* $\{e_n\}$ *przestrzeni* H.

(i) *Jeżeli ciąg skalarów jest sumowalny w kwadracie:* $\alpha := \{\alpha_n\} \in \ell^2$, *to*

$$\Phi(\alpha) := \sum_j \alpha_j e_j$$

jest szeregiem zbieżnym.

(ii) *Tak określone odwzorowanie* $\Phi : \ell^2 \to H$ *jest izomorfizmem unitarnym*[1].

[1] czyli bijekcją liniową, zachowującą iloczyn skalarny

Dowód. Tożsamość Parsevala stosowana do $x = x_m - x_n$, gdzie $x_k = \sum_1^k \alpha_j e_j$ dowodzi, że dla $\alpha \in \ell^2$ ciąg $\{x_n\}_n$ jest ciągiem Cauchy'ego, a jego granicą jest $\Phi(\alpha)$. Izometryczność Φ jest dokładnie tożsamością (TP). Surjektywność można uzyskać zauważając, że dla $h \in H$ ciąg współczynników: $\alpha_j = \langle h, e_j\rangle$ jest „sumowalny z kwadratem”, tzn $\alpha \in \ell^2$ (dzięki (NB) lub (TP)). Ponadto mamy relacje $h - \Phi(\alpha) \perp e_j$ $(\forall j)$, a więc $h = \Phi(\alpha)$. •

Wynik tej obserwacji, czyli relację $\Phi \circ \Phi^{-1} = Id_H$ zapiszmy w postaci

$$x = \sum_1^\infty \langle x, e_j\rangle e_j \tag{SF}$$

Wzór ten nazywamy **rozwinięciem wektora** x **w szereg ortogonalny** lub **(abstrakcyjnym) szeregiem Fouriera** wektora x w bazie $\{e_n\}_{n=1}^\infty$. Izometryczność Φ^{-1} jest zarazem równoważna relacji

$$\langle x, y\rangle = \langle \Phi^{-1}x, \Phi^{-1}y\rangle = \sum_{j=1}^\infty \langle x, e_j\rangle\langle e_j, y\rangle. \tag{TP2}$$

Rzeczywiście, obydwie strony relacji (TP2) są formami liniowymi względem zmiennej x, zaś antyliniowymi względem zmiennej y. Formy te pokrywają się na „przekątnej” zbioru $H \times H$, czyli na zbiorze par (x, y), dla których $x = y$ (na mocy (TP)). Wzór polaryzacyjny (wzór (P) z Lematu 10.10), w którego dowodzie nie korzysta się z relacji $\langle x, x\rangle > 0$ dla $x \neq 0$, dowodzi równości tych dwu form już na zbiorze wszystkich par $(x, y) \in H \times H$.

3.3 Wielomiany trygonometryczne

W dalszym ciągu będziemy zajmowali się jednym konkretnym układem ortonormalnym w $L^2[0, 1]$, zwanym **układem trygonometrycznym**. Najbardziej przejrzystą formę ma jego zapis w postaci zespolonej. Przyjmijmy oznaczenie $\Gamma = \{z \in \mathbb{C}; |z| = 1\}$ dla okręgu jednostkowego (o parametryzacji $z(t) := e^{2\pi it}$, $t \in [0, 1]$). Niech

$$e_n(z) := z^n, \quad e_n(z(t)) = e^{2\pi int} = \cos 2\pi nt + i \sin 2\pi nt, \quad n \in \mathbb{Z}. \tag{cT}$$

Skończone kombinacje liniowe elementów układu trygonometrycznego nazywamy **(zespolonymi) wielomianami trygonometrycznymi**. Ich części rzeczywiste – postaci

$$w(t) = \frac{a_0}{2} + \sum_{m=1}^N a_m \cos(2\pi mt) + b_m \sin(2\pi mt) \tag{WT}$$

znane są jako „wielomiany trygonometryczne”. Złożenie wielomianu skończonej ilości zmiennych z elementami rzeczywistego układu trygonometrycznego:

$$1, \quad \frac{1}{\sqrt{2}}\cos(2\pi kt), \quad \frac{1}{\sqrt{2}}\sin(2\pi kt), \quad k = \pm1, \pm2, \ldots \tag{rT}$$

jest bowiem nadal postaci (WT), co najwygodniej sprawdzić w postaci zespolonej, korzystając z oczywistej zależności: $e_n e_m = e_{n+m}$, $\bar{e}_n = e_{-n}$. Funkcje (WT) rozpatrywać można dla $t \in \mathbb{R}$ – są to okresowe funkcje ciągłe, jednostajnie ograniczone, zaś zespolone wielomiany trygonometryczne $\sum_{k=-N}^N c_k e_k(z)$ są ciągłe na okręgu Γ. Z twierdzenia Stone'a – Weierstrassa, tworzą one gęstą podalgebrę w $C(\Gamma, \mathbb{C})$ (w normie $\|f\|_\infty := \sup\{|f(z)|; z \in \Gamma\}$). Stąd rzeczywiste wielomiany trygonometryczne stanowią zbiór gęsty w $C([0, 1], \mathbb{R})$.

Ze względu na inkluzję $C[0,1] \subset L^p[0,1]$ i na nierówności[2] $\|\cdot\|_p \leqslant \|\cdot\|_\infty$ rzeczywiste wielomiany trygonometryczne tworzą gęsty podzbiór przestrzeni $L^p[0,1]$ – oczywiście z wyjątkiem $p = \infty$. Wynika to z następującego faktu.

Lemat 3.11. *Zbiór $C[0,1]$ jest gęsty w $L^p[0,1]$ gdy $1 \leqslant p < \infty$.*

Dowód. Dla $f \in L^p$ niech $f_n(t)$ będzie równa $f(t)$ dla tych t, dla których $|f(t)| \leqslant n$ oraz zero – dla pozostałych t. Jak łatwo sprawdzić, $\|f_n - f\|_p \to 0$ dla $n \to \infty$. Zbiór (klas równoważności) funkcji mierzalnych ograniczonych jest więc gęsty w $L^p[0,1]$. Z kolei, funkcje mierzalne ograniczone można jednostajnie (więc i w $\|\cdot\|_p$) aproksymować funkcjami prostymi. Wystarczy więc umieć aproksymować w L^p funkcje charakterystyczne zbiorów mierzalnych – funkcjami ciągłymi. Ze względu na regularność miary (Lebesgue'a) – zbiory mierzalne można (w sensie miary) „przybliżać” ich nadzbiorami otwartymi. Na koniec, zbiory otwarte $G \subset [0,1]$ są przeliczalnymi sumami odcinków. Wynika stąd liniowa gęstość w badanej przestrzeni funkcji charakterystycznych przedziałów (np. o końcach wymiernych $\Rightarrow$ ośrodkowość L^p). Te ostatnie można bardzo łatwo aproksymować funkcjami ciągłymi (np. kawałkami liniowymi). •

Jedną z podstaw teorii szeregów Fouriera jest następująca obserwacja:

Lemat 3.12. *Dla $p = 2$ układ (rT) jest bazą ortonormalną.*

Dowód elementarny (w oparciu o wzory trygonometryczne) można znaleźć w podręcznikach analizy (np. u Fichtenholza). My przeprowadzimy go w oparciu o wzór całkowy Cauchy'ego.

W przypadku zespolonym niech μ będzie znormalizowaną długością łuku

$$\int f(z)d\mu = \int_0^1 f(z(t))dt \ (= \frac{1}{2\pi}\int_0^{2\pi} f(e^{is})ds).$$

Wzór całkowy Cauchy'ego w punkcie $w = 0$ dla funkcji g analitycznych w otoczeniu domkniętego koła jednostkowego

$$g(w) = \frac{1}{2\pi i}\oint_\Gamma \frac{g(z)}{z-w}dz$$

przy parametryzacji okręgu Γ wzorem: $z(t) := e^{2\pi i t}$, $t \in [0,1]$ przyjmie dla $w = 0$ postać

$$g(0) = \int g d\mu.$$

Podstawiając $g = e_n$ i uwzględniając fakt, że $\int e_{-n}d\mu$ jest sprzężeniem liczby zespolonej $\int e_n d\mu$ otrzymujemy $\int e_n d\mu = e_n(0) = \delta_{n,0}$. Zastosowanie równości $e_n\overline{e_m} = e_{n-m}$ daje ortonormalność układu (cT). Ponieważ mamy relacje prostopadłości $(e_n \pm e_{-n}) \perp (e_m \pm e_{-m})$ dla $n \neq m$ (ćw.), wynika stąd ortonormalność rzeczywistego układu trygonometrycznego (rT).

3.4 Komentarze

Fourier w 1807 twierdził, że *każda funkcja jest sumą szeregu*

$$f(t) = \sum_{k=-\infty}^{\infty} \hat{f}(k)e^{2\pi ikt} \quad dla \quad \hat{f}(k) := \int f(t)e^{-2\pi ikt}dt. \tag{sF}$$

[2] Przestrzeń $L^p[0,1]$ można zdefiniować jako uzupełnienie przestrzeni $C[0,1]$ względem normy: $\|f\|_p := (\int_0^1 |f(t)|^p)^{\frac{1}{p}}$. Oprócz (klas) funkcji całkowalnych w sensie Riemanna trzeba jeszcze dołączyć pewne inne funkcje-całkowalne z p-tą potęgą w sensie Lebesgue'a, przy czym funkcje równe prawie wszędzie należy utożsamiać.

W tym czasie żadne z obecnie używanych pojęć: funkcji (Dirichlet, Cantor), całki (Riemann, 1856), zbieżności topologicznej nie było precyzyjnie sformułowane. W przypadku funkcji ciągłych mogło chodzić o znaną wówczas zbieżność jednostajną. Już około 1876 roku Du Bois-Reymond skonstruował jednak funkcję ciągłą okresową, dla której sumy częściowe szeregu (sF) nie są ograniczone w punkcie $t = 0$. W roku 1923 Kołmogorow znalazł funkcję sumowalną $f \in L^1[0,1]$ o rozbieżnym prawie wszędzie szeregu (sF). Zbieżność (sF) prawie wszędzie dla dowolnych $f \in L^2[0,1]$ została udowodniona dopiero w roku 1966 (L. Carleson). Teoria szeregów Fouriera stała się motorem postępu dla matematyki z przełomu XIX i XX wieku, mając zarazem fundamentalne znaczenie dla technicznych zagadnień analizy sygnałów, drgań i wielu innych konkretnych modeli matematycznych.

Rozdział 4

Szeregi Fouriera

4.1 Wielomiany trygonometryczne

W dalszym ciągu będziemy zajmowali się jednym konkretnym układem ortonormalnym w $L^2[0,1]$, zwanym **układem trygonometrycznym**. Najbardziej przejrzystą formę ma jego zapis w postaci zespolonej. Przyjmijmy oznaczenie $\Gamma = \{z \in \mathbb{C};\ |z| = 1\}$ dla okręgu jednostkowego (o parametryzacji $z(t) := e^{2\pi it}$, $t \in [0,1]$). Niech

$$e_n(z) := z^n, \quad e_n(z(t)) = e^{2\pi int} = \cos 2\pi nt + i \sin 2\pi nt, \quad n \in \mathbb{Z}. \tag{cT}$$

Skończone kombinacje liniowe elementów układu trygonometrycznego nazywamy **(zespolonymi) wielomianami trygonometrycznymi**. Ich części rzeczywiste – postaci

$$w(t) = \frac{a_0}{2} + \sum_{m=1}^{N} a_m \cos(2\pi mt) + b_m \sin(2\pi mt) \tag{WT}$$

znane są jako „wielomiany trygonometryczne". Złożenie wielomianu skończonej ilości zmiennych z elementami rzeczywistego układu trygonometrycznego:

$$1, \quad \frac{1}{\sqrt{2}} \cos(2\pi kt), \quad \frac{1}{\sqrt{2}} \sin(2\pi kt), \quad k = \pm 1, \pm 2, \ldots \tag{rT}$$

jest bowiem nadal postaci (WT), co najwygodniej sprawdzić w postaci zespolonej, korzystając z oczywistej zależności: $e_n e_m = e_{n+m}$, $\bar{e}_n = e_{-n}$. Funkcje (WT) rozpatrywać można dla $t \in \mathbb{R}$ – są to okresowe funkcje ciągłe, jednostajnie ograniczone, zaś zespolone wielomiany trygonometryczne $\sum_{k=-N}^{N} c_k e_k(z)$ są ciągłe na okręgu Γ. Z twierdzenia Stone'a – Weierstrassa, tworzą one gęstą podalgebrę w $C(\Gamma, \mathbb{C})$ (w normie $\|f\|_\infty := \sup\{|f(z)|;\ z \in \Gamma\}$). Stąd rzeczywiste wielomiany trygonometryczne stanowią zbiór gęsty w $C([0,1], \mathbb{R})$.

Ze względu na inkluzję $C[0,1] \subset L^p[0,1]$ i na nierówności[1] $\|\cdot\|_p \leqslant \|\cdot\|_\infty$ rzeczywiste wielomiany trygonometryczne tworzą gęsty podzbiór przestrzeni $L^p[0,1]$ – oczywiście z wyjątkiem $p = \infty$. Wynika to z następującego faktu.

Lemat 4.1. *Zbiór $C[0,1]$ jest gęsty w $L^p[0,1]$ gdy $1 \leqslant p < \infty$.*

[1] Przestrzeń $L^p[0,1]$ można zdefiniować jako uzupełnienie przestrzeni $C[0,1]$ względem normy: $\|f\|_p := (\int_0^1 |f(t)|^p)^{\frac{1}{p}}$. Oprócz (klas) funkcji całkowalnych w sensie Riemanna trzeba jeszcze dołączyć pewne inne funkcje- całkowalne z p-tą potęgą w sensie Lebesgue'a, przy czym funkcje równe prawie wszędzie należy utożsamiać.

Dowód. Dla $f \in L^p$ niech $f_n(t)$ będzie równa $f(t)$ dla tych t, dla których $|f(t)| \leqslant n$ oraz zero – dla pozostałych t. Jak łatwo sprawdzić, $\|f_n - f\|_p \to 0$ dla $n \to \infty$. Zbiór (klas równoważności) funkcji mierzalnych ograniczonych jest więc gęsty w $L^p[0,1]$. Z kolei, funkcje mierzalne ograniczone można jednostajnie (więc i w $\|\cdot\|_p$) aproksymować funkcjami prostymi. Wystarczy więc umieć aproksymować w L^p funkcje charakterystyczne zbiorów mierzalnych – funkcjami ciągłymi. Ze względu na regularność miary (Lebesgue'a) – zbiory mierzalne można (w sensie miary) „przybliżać" ich nadzbiorami otwartymi. Na koniec, zbiory otwarte $G \subset [0,1]$ są przeliczalnymi sumami odcinków. Wynika stąd liniowa gęstość w badanej przestrzeni funkcji charakterystycznych przedziałów (np. o końcach wymiernych $\Rightarrow$ ośrodkowość L^p). Te ostatnie można bardzo łatwo aproksymować funkcjami ciągłymi (np. kawałkami liniowymi). $\bullet$

Jedną z podstaw teorii szeregów Fouriera jest następująca obserwacja:

Lemat 4.2. *Dla $p = 2$ układ (rT) jest bazą ortonormalną.*

Dowód elementarny (w oparciu o wzory trygonometryczne) można znaleźć w podręcznikach analizy (np. u Fichtenholza). My przeprowadzimy go w oparciu o wzór całkowy Cauchy'ego.

W przypadku zespolonym niech μ będzie znormalizowaną długością łuku

$$\int f(z)d\mu = \int_0^1 f(z(t))dt \ (= \frac{1}{2\pi}\int_0^{2\pi} f(e^{is})ds).$$

Wzór całkowy Cauchy'ego w punkcie $w = 0$ dla funkcji g analitycznych w otoczeniu domkniętego koła jednostkowego

$$g(w) = \frac{1}{2\pi i}\oint_\Gamma \frac{g(z)}{z-w}dz$$

przy parametryzacji okręgu Γ wzorem: $z(t) := e^{2\pi i t}$, $t \in [0,1]$ przyjmie dla $w = 0$ postać

$$g(0) = \int g d\mu.$$

Podstawiając $g = e_n$ i uwzględniając fakt, że $\int e_{-n}d\mu$ jest sprzężeniem liczby zespolonej $\int e_n d\mu$ otrzymujemy $\int e_n d\mu = e_n(0) = \delta_{n,0}$. Zastosowanie równości $e_n\overline{e_m} = e_{n-m}$ daje ortonormalność układu (cT). Ponieważ mamy relacje prostopadłości $(e_n \pm e_{-n}) \perp (e_m \pm e_{-m})$ dla $n \neq m$ (ćw.), wynika stąd ortonormalność rzeczywistego układu trygonometrycznego (rT).

4.2 Komentarze

Fourier w 1807 twierdził, że *każda funkcja jest sumą szeregu*

$$f(t) = \sum_{k=-\infty}^{\infty} \hat{f}(k)e^{2\pi ikt} \quad dla \quad \hat{f}(k) := \int f(t)e^{-2\pi ikt}dt. \tag{sF}$$

W tym czasie żadne z obecnie używanych pojęć: funkcji (Dirichlet, Cantor), całki (Riemann, 1856), zbieżności topologicznej nie było precyzyjnie sformułowane. W przypadku funkcji ciągłych mogło chodzić o znaną wówczas zbieżność jednostajną. Już około 1876 roku Du Bois-Reymond skonstruował jednak funkcję ciągłą okresową, dla której sumy częściowe szeregu (sF) nie są ograniczone w punkcie $t = 0$. W roku 1923 Kołmogorow znalazł funkcję sumowalną $f \in L^1[0,1]$ o rozbieżnym prawie wszędzie szeregu (sF). Zbieżność (sF) prawie wszędzie dla dowolnych $f \in L^2[0,1]$ została udowodniona dopiero w roku 1966 (L. Carleson). Teoria szeregów Fouriera stała się motorem postępu dla matematyki z przełomu XIX i XX wieku, mając zarazem fundamentalne znaczenie dla technicznych zagadnień analizy sygnałów, drgań i wielu innych konkretnych modeli matematycznych.

Badanie szeregów Fouriera funkcji z $L^p[0,1]$ zaczynamy od najłatwiejszego przypadku $p = 2$. Następnie wykazujemy twierdzenie znane jako Lemat Riemanna – Lebesgue'a i kryteria stwierdzające regularność funkcji na podstawe jej współczynników Fouriera. Wyrażamy sumy częściowe szeregu Fouriera w postaci całki splotowej Dirichleta i podajemy twierdzenia Diniego oraz Dirichleta o zbieżności tego szeregu w danym punkcie, z zastosowaniem do obliczania pewnych szeregów liczbowych.

4.3 Przypadek L^2

W wykładzie tym interesować nas będzie układ trygonometryczny, którego podstawową własność (ortonormalności) uzyskaliśmy w ostatniej części poprzedniego wykładu. Obecnie możmy już przejść do istoty zagadnienia. Zbadamy mianowicie zależność pomiędzy funkcją całkowalną $f \in L^1[0, 2\pi]$, a jej współczynnikami Fouriera

$$\hat{f}(n) := \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-int} dt$$

oraz spróbujemy określić, w jakim sensie formalny szereg Fouriera

$$\sum_{-\infty}^{+\infty} \hat{f}(n) e^{int} \tag{sF}$$

wyznacza funkcję f. Najprostsze obserwacje dotyczą funkcjonałów wyznaczania n-tego współczynnika, $\chi_n : L^1(0, 2\pi) \ni f \mapsto \hat{f}(n)$.

Lemat 4.3. *Funkcjonały χ_n są liniowe oraz ciągłe na przestrzeniach $L^p[0, 2\pi]$, $1 \leqslant p < \infty$. Zachodzi oszacowanie*

$$|\hat{f}(n)| \leqslant \|f\|_p \text{ dla } 1 \leqslant p \leqslant \infty. \tag{†}$$

Operatory sum częściowych,

$$S_N : L^p[0, 2\pi] \ni f \to S_N[f] := \sum_{k=-N}^{N} \hat{f}(k) e^{ikt} \in L^p[0, 2\pi] \tag{‡}$$

są więc liniowe i ciągłe.

Dowód. Oszacowanie (†) można wywnioskować (stosując nierówność Höldera) stąd, że norma funkcji e^{-int} w każdej z przestrzeni $L^q[0, 2\pi]$ (względem unormowanej miary Lebesgue'a $\frac{dt}{2\pi}$ wynosi 1 ($\forall 1 \leqslant q \leqslant \infty$) . •

W przypadku przestrzeni Hilberta ($p = 2$) sytuacja jest stosunkowo prosta: dla baz ortonormalnych $\{f_j\}_1^\infty \subset H$ twierdzenie Riesza – Fischera i tożsamość Parsevala dowodzą zbieżności w topologii normy **abstrakcyjnego szeregu Fouriera** dowolnego wektora $x \in H$, czyli szeregu

$$x = \sum_1^\infty \alpha_j f_j, \qquad \alpha_j = \langle x, f_j \rangle, \quad \sum_j |\alpha_j|^2 < \infty.$$

Współczynniki rozwinięcia, α_j stanowią ciąg $\alpha \in \ell^2$. Na odwrót, ciąg spełniający warunek $\sum_j |\alpha_j|^2 < \infty$ jest ciągiem współczynników pewnego wektora $x \in H$. Biorąc jako bazę $\{f_n\}_1^\infty$ układ (cT) $= \{e_j;\, j \in \mathbb{Z}\}$ w przypadku zespolonym, lub (rT) – w przypadku rzeczywistym, otrzymujemy podstawową charakteryzację zbieżności (sF) w przestrzeni L^2.

Twierdzenie 4.4. *Dla $f \in L^1[0, 2\pi]$ zachodzi równoważność*

$$f \in L^2[0, 2\pi] \Leftrightarrow \{\hat{f}(n)\}_{n=-\infty}^{+\infty} \in \ell^2.$$

Szereg ten jest zbieżny w normie przestrzeni $L^2[0, 2\pi]$. Operatory sum częściowych, S_n stanowią ciąg projekcji ortogonalnych zbieżnych do identyczności w silnej topologii operatorowej, czyli

$$\|f - S_n[f]\|_2 \longrightarrow 0 \qquad \text{przy} \quad n \to \infty \quad (\forall f \in L^2[0, 2\pi]).$$

Natomiast dla $p = 1$, a nawet dla funkcji ciągłych, nie ma tak prostej odpowiedniości między współczynnikami Fouriera $f \in L^p[0, 2\pi]$, a własnościami szeregu (sF). Na przykład, dla współczynników Fouriera funkcji z $L^1[0, 2\pi]$ niecałkowalnej w kwadracie, poprzez wymnożenie ich przez odpowiednio dobrany ciąg liczb równych ± 1 można uzyskać szereg Fouriera, który nie odpowiada żadnej z funkcji całkowalnych. Nie istnieje więc żadne kryterium (typu WKW) określające w terminach modułów wyrazów ciągu, kiedy jest to ciąg współczynników funkcji całkowalnej na przedziale $[0, 2\pi]$.

Istnieją natomiast pewne warunki konieczne dla bycia ciągiem współczynników funkcji o zadanej klasie regularności. Zacznijmy od najważniejszego kryterium, dotyczącego funkcji całkowalnych.

Twierdzenie 4.5. (Lemat Riemanna – Lebesgue'a) *Ciąg współczynników Fouriera funkcji $f \in L^1[0, 2\pi]$ jest zbieżny do zera, tzn.*

$$\lim_{|n| \to \infty} \hat{f}(n) = 0 \quad \text{dla} \quad \hat{f}(n) := \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-int} dt.$$

Warto zauważyć, że również $\int_a^b f(t) e^{-int} dt \to 0$ dla $a < b$, $a, b \in \mathbb{R}$, co wynika z pomnożenia f przez funkcję charakterystyczną odpowiedniego przedziału, gdy $0 \leqslant a < b \leqslant 2\pi$. W ogólnym przypadku można skorzystać z okresowości (zamieniając ewentualnie całkę na skończoną sumę całek tego typu).

Dowód. Dla funkcji z przestrzeni $L^2[0, 2\pi]$ teza wynika z należenia ciągu $\{\hat{f}(n)\}_{n=-\infty}^{+\infty}$ do ℓ^2 (z nierówności Bessela). Ale przestrzeń ta stanowi gęsty podzbiór w $L^1[0, 2\pi]$ Ustalając $\epsilon > 0$ znajdziemy $g \in L^2[0, 2\pi]$ bliskie f tak, że $\|f - g\|_1 \leqslant \frac{\epsilon}{2}$. Teraz dla funkcji g, na mocy pierwszej części dowodu, znajdziemy $M > 0$ tak, by $|\hat{g}(n)| < \frac{\epsilon}{2}$ dla wszystkich $n \geqslant M$. Dla tych indeksów n mamy $|\hat{f}(n)| \leqslant |\hat{g}(n)| + |\widehat{f - g}(n)| < \epsilon$. (Korzystamy tu z nierówności (†).) •

4.4 Szereg Fouriera a regularność funkcji

W pewnym sensie odwrotnym zagadnieniem jest pytanie, na ile założenia o wzroście współczynników Fouriera pozwalają wnioskować o regularności funkcji i o typie zbieżności jej szeregu Fouriera. Tu oszacowania (przez 1) normy supremowej elementów układu trygonometrycznego dają bezpośredni, lecz bardzo przydatny wniosek.

Twierdzenie 4.6

(i) *Gdy ciąg modułów współczynników Fouriera funkcji f jest sumowalny ($\{\hat{f}(n)\}_{n<\infty} \in \ell^1$), to f jest ciągła, zaś zbieżność jej szeregu Fouriera do f jest jednostajna.*

(ii) *Jeżeli dla ustalonej liczby naturalnej k mamy $\{n^k \cdot \alpha_n\} \in \ell^1$, to α_n stanowią ciąg współczynników Fouriera pewnej funkcji okresowej klasy C^k, której szereg Fouriera jest zbieżny jednostajnie wraz z pochodnymi rzędu $\leqslant k$.*

Dowód. Wystarczy zauważyć, że moduł k-tej pochodnej z funkcji e^{int} szacuje się przez $n(n-1)\cdot\ldots\cdot(n-k+1) \leqslant n^k$ i zastosować obowiązującą w przestrzeniach zupełnych zasadę: „szeregi o sumowalnych normach wyrazów są zbieżne”. •

Jedną z zalet szeregów Fouriera jest bardzo prosty wzór na szereg funkcji pochodnej. Dla funkcji u klasy C^1 na okręgu Γ (o parametryzacji $z(t) = e^{2\pi it}$, pozwalającej też traktować u jako funkcję okresową na $\mathbb{R}$), stosując całkowanie przez części, otrzymamy bowiem wyrażenie współczynników pochodnej: $\widehat{u'}(n)$ w postaci

$$\int_0^1 u'(z(t))z(t)^{-n}dt = u(z(t))(z(t))^{-n}|_0^1 + 2\pi in \int_0^1 u(z(t))(z(t))^{-n} = 2\pi in\widehat{u}(n).$$

Zerowanie się pierwszego ze składników sumy otrzymanej w wyniku całkowania przez części wynika z okresowości funkcji $u(z(t))(z(t))^{-n}$. Stosując k-krotnie otrzymaną równość otrzymamy

Twierdzenie 4.7. (O SZEREGU FOURIERA POCHODNEJ) *Jeżeli funkcja okresowa u jest klasy C^k, to współczynniki Fouriera u oraz jej k-tej pochodnej związane są wzorem*

$$\widehat{u^{(k)}}(n) = (2\pi in)^k\widehat{u}(n).$$

Ponadto ciąg $\{n^k\widehat{u}(n)\}$ jest zbieżny do zera (więc ograniczony).

W szczególności, ciągi współczynników Fouriera dla funkcji okresowych klasy C^k są typu $o(n^{-k})$ przy $n \to \infty$ („o-małe”), co dla $k = 2$ gwarantuje sumowalność. Mamy więc następujący wniosek z ostatnich dwu twierdzeń:

Wniosek 4.8. *Szereg Fouriera funkcji okresowej klasy C^2 jest do niej zbieżny jednostajnie.*[2]

Wniosek 4.9. *Ciąg skalarów α_n jest ciągiem współczynników Fouriera pewnej funkcji okresowej klasy C^∞ wtedy i tylko wtedy, gdy $(\forall k \in \mathbb{N})$ ciąg $\{n^k\alpha_k\}$ jest ograniczony. Zbieżność jednostajna ma wtedy miejsce dla szeregów Fouriera zarówno samej funkcji, jak i jej pochodnych dowolnego rzędu.*

Jak już wspomniano, sama ciągłość funkcji okresowej nie wystarcza dla zbieżności jednostajnej jej szeregu Fouriera. Istnieją pewne słabsze od dwukrotnej różniczkowalności warunki wystarczające dla takiej zbieżności. Jak wykazał Bernstein, jednym z nich jest np. istnienie stałych $\frac{1}{2} < \alpha \leqslant 1$, $M > 0$ takich, że dla wszystkich $s, t \in \mathbb{R}$ mamy

$$|u(z(s)) - u(z(t))| \leqslant M|s-t|^\alpha.$$

Dowód pomijamy.

4.5 Zbieżność w danym punkcie

Natomiast zbieżność punktowa (sF) do danej funkcji wymaga znacznie słabszych założeń. Co ważniejsze, ma ona „charakter lokalny” i zachodzi, między innymi, w punktach ciągłości $f \in L^1[0,1]$.

[2]Wystarczy, by $f \in C^1[0,2\pi]$, $f(0) = f(2\pi)$, bo mamy oszacowanie jednostajne szeregu $\sum|\widehat{f}_n e_n(z)|$ przez iloczyn ℓ_2-norm z ciągów: $(n\widehat{f_n})$ oraz $(\frac{1}{n})$.

Metoda dowodu zasługuje na bliższe poznanie. Zaczniemy od konstrukcji pewnego operatora całkowego, zwanego **całką Dirichleta**, który wyraża sumę częściową S_N odpowiadającą indeksom o module $\leqslant N$.

Stosując wzory Eulera do $z^{-m}(t) + z^m(t) = 2\cos(2\pi mt)$ i wzór

$$2\sin\alpha\,\cos\beta = sin(\alpha-\beta) + \sin(\alpha+\beta)$$

(wymnażając $\sum_{-N}^{N} z^m(t)$ przez $\sin\pi t$, a po dokonaniu uproszczeń – ponownie dzieląc) można otrzymać wzór

$$D_N(t) := \sum_{m=-N}^{N} e^{2\pi imt} = \frac{\sin\pi(2N+1)t}{\sin\pi t}. \tag{$*$}$$

Gdy $t \in \mathbb{Z}$, prawą stronę należy rozumieć w sensie asymptotycznym. Funkcje D_N zwane **jądrami całki Dirichleta** spełniają (jednostajne względem wskaźników $N \in \mathbb{N}$) oszacowania

$$|D_N(t)| \leqslant \frac{1}{\sin\pi t},$$

które są również jednostajne względem $t \in [\epsilon, 1-\epsilon]$ przy ustalonej (małej) wartości $\epsilon > 0$. Są to wielomiany trygonometryczne, których współczynniki Fouriera $\widehat{D_N}(m)$ są równe 1 dla $|m| \leqslant N$ oraz równe 0 dla $|m| > N$. Ich znaczenie wyjaśnia następujący wzór całkowy:

Lemat 4.10. *Dla $f \in L^1[0,1], s \in [0,1]$ mamy*

$$S_N[f](s) = \int_0^1 D_N(s-t)f(t)dt.$$

Dowód. Rzeczywiście,

$$S_N[f](s) = \sum_{m=-N}^{N} \widehat{f}(m)e^{2\pi ims} = \sum_{m=-N}^{N} \int_0^1 f(t)e^{-2\pi imt}e^{2\pi ims}dt,$$

co po zamianie kolejności całkowania z sumowaniem daje tezę. •

Dla funkcji stałej równej 1 mamy $S_N[1](s) = 1$, co w zestawieniu z otrzymanym wzorem i z własnościami funkcji *sinus* daje

$$\int_{\frac{1}{2}}^{1} D_N(t)\,dt = \frac{1}{2} = \int_0^{\frac{1}{2}} D_N(t)\,dt.$$

Jeżeli teraz w punkcie s_0 funkcja f ma granice lewostronne i prawostronne ($f(s_0+0)$ oraz $f(s_0-0)$), to niech $z_0 := \frac{1}{2}(f(s_0+0) + f(s_0-0))$. Podstawiając za $\frac{1}{2}$ ostatnie całki, otrzymamy ze względu na okresowość odpowiednich funkcji wzór

$$S_N[f](s_0) - z_0 = \int_0^{\frac{1}{2}} (f(s_0+t) - f(s_0+0) + f(s_0-t) - f(s_0-0))D_N(t)dt \tag{$**$}$$

Jest to bardzo typowa dla tego typu zagadnień sytuacja: na części obszaru całkowania (na przedziale $[\epsilon, \frac{1}{2}]$) funkcje D_N są w pewnym sensie „małe”, a na pozostałej części mała jest funkcja

$$\psi(t) := f(s_0+t) - f(s_0+0) + f(s_0-t) - f(s_0-0).$$

Trudności biorą się stąd, że L^1-normy jąder D_N nie są wspólnie ograniczone względem N. Jest parę sposobów podejścia do tego problemu, a każdy z nich prowadzi do innych warunków wystarczających.

Po pierwsze, mianownik D_N można „przyłączyć do ψ". Niech mianowicie

$$\varphi(t) := \frac{\psi(t)}{\sin \pi t}.$$

Powiemy, że funkcja ta jest całkowalna w pobliżu zera, gdy

$$(\exists h > 0) \int_0^h |\varphi(t)| dt < \infty$$

lub, co na jedno wychodzi, gdy

$$\int_0^h \frac{|\psi(t)|}{t} dt < \infty.$$

Warunek taki zachodzi np. dla funkcji spełniających nierówność typu

$$|f(s_0 + t) - f(s_0)| \leqslant K|t|^\alpha, \quad K > 0, \alpha \leqslant 1$$

w dostateczne małym otoczeniu badanego punktu s_0. W szczególności (dla $\alpha = 1$) takie nierówności spełnia funkcja f, mająca w punkcie s_0 skończone pochodne: prawo - i lewostronną. W tych przypadkach zbieżność $S_N[f](s_0)$ do $\frac{1}{2}\psi_0 = \frac{1}{2}(f(s_0 - 0) + f(s_0 + 0)) = z_0$ wynika z następującego kryterium.

Twierdzenie 4.11. (Dini) *Szereg Fouriera funkcji całkowalnej f jest zbieżny w punkcie s_0 do z_0, gdy funkcja $\frac{|\psi(t)|}{t}$ jest całkowalna w pobliżu zera.*

Dowód. Gdy prawą stronę wzoru (∗∗) zapiszemy jako sumę całek:

$$\left\{\int_0^\delta + \int_\delta^{\frac{1}{2}}\right\} \psi(t) D_N(t) dt = \left\{\int_0^\delta + \int_\delta^{\frac{1}{2}}\right\} \varphi(t) \sin[\pi(2N+1)t] dt = I_1 + I_2,$$

to drugi składnik, $I_2 = \int_\delta^{\frac{1}{2}} \varphi(t) \sin[\pi(2N+1)t] dt$. Zmierza on do zera przy $N \to \infty$ dzięki uwadze sformułowanej po wypowiedzi Lematu Riemanna – Lebesgue'a. Funkcja φ jest bowiem całkowalna jako iloczyn funkcji całkowalnej ψ (co wynika z relacji $f \in L^1[0,1]$ oraz funkcji $\frac{1}{\sin \pi t}$ – która jest ograniczona na przedziale $[\delta, \frac{1}{2}]$.

Pierwszy składnik będzie, dzięki założeniu, dowolnie mały, o ile wartość $\delta > 0$ wybierzemy dostatecznie blisko zera. (W zasadzie dopiero ustaliwszy to δ przystępujemy do rozważania I_2.) •

Oprócz warunku Diniego najistotniejszy jest warunek Dirichleta: Jeżeli f jest w otoczeniu punktu s_0 wyrażalna jako różnica dwu funkcji rosnących, to powiemy, że f jest **funkcją o wahaniu ograniczonym** w otoczeniu s_0.

Twierdzenie 4.12. (Dirichlet) *Szereg Fouriera funkcji całkowalnej, której wahanie jest ograniczone w otoczeniu punktu s_0 jest zbieżny do wartości $z_0 = \frac{1}{2}(f(s_0 - 0) + f(s_0 + 0))$.*

Dowód. Oczywiście, funkcje monotoniczne posiadają granice jednostronne, co uzasadnia definicję wartości z_0. Utrzymując bez zmian rozumowanie dotyczące całki I_2 z poprzedniego twierdzenia, wystarczy wykazać odpowiednią zbieżność I_1 do zera. W tym celu wystarczy rozważyć przypadek, w którym f jest rosnąca w otoczeniu s_0. Wówczas można zastosować do

$$I_1^+ := \int_0^\delta (f(s_0+t) - f(s_0+0))D_N(t)dt,$$

jako całki z iloczynu funkcji rosnącej i funkcji D_N jedno z twierdzeń o całkowej wartości średniej znajdując pewien „punkt pośredni” $\beta \in (0,\delta)$, dla którego

$$I_1^+ = (f(s_0+\delta) - f(s_0+0)) \cdot \int_\beta^\delta D_N(t)\,dt.$$

(Wykorzystujemy mianowicie równość $\int_a^b g(t)h(t)\,dt$ wartości $g(b)\int_\eta^b h(t)\,dt$ dla odpowiednio dobranego $\eta \in (a,b)$, gdy g jest funkcją rosnącą, nieujemną.) Czynnik przed ostatnią całką zmierza do zera przy $\delta \to 0$, zaś sama całka jest ograniczona względem N. Aby się o tym przekonać, można sprowadzić zagadnienie do równoważnej ograniczoności ciągu

$$\int_\beta^\delta \frac{\sin Nt}{t}dt = \int_{N\beta}^{N\delta} \frac{\sin u}{u}du,$$

która wynika ze zbieżności całki $\int_0^\infty \frac{\sin u}{u}du$. Wspomniana równoważność wynika stąd, że różnica funkcji podcałkowych jest funkcją ograniczoną (przez stałą niezależną od N). Podobnie postępujemy ze składnikiem

$$I_1^- = \int_0^\delta (f(s_0-t) - f(s_0-0))D_N(t)\,dt.$$

Ponieważ $I_1 = I_1^+ + I_1^-$, otrzymujemy stąd tezę. •

4.6 Komentarze

Znaczenie warunku Dirichleta polega na tym, że jest on spełniony przez większość używanych w praktyce funkcji (na przykład dla funkcji „kawałkami monotonicznej” (monotonicznym na każdym ze skończonej liczby przedziałów wypełniających cały odcinek $[0,2\pi]$). Gdy funkcja ma przeliczalną ilość ekstremów lokalnych, warunek ograniczoności wahania oznacza, iż suma łącznej długości jej przyrostów i „ubytków” pomiędzy punktami kolejnych ekstremów lokalnych (na przedziałach monotoniczności jest skończona.

Równie ważne jest twierdzenie Riemanna o lokalizacji mówiące, że na zbieżność i wartości sumy szeregu Fouriera funkcji okresowej sumowalnej ma wpływ jedynie jej zachowanie w otoczeniu danego punktu. Otrzymujemy je zauważając, że dla różnicy dwu funkcji pokrywających się w otoczeniu danego punktu występuje tylko drugi fragment całki Dirichleta, który zmierza do zera na mocy dowodu twierdzenia Dirichleta.

Poprzez prostą zamianę zmiennych wszystkie dotychczasowe wyniki przenoszą się na przypadek funkcji o dowolnym okresie. Najczęściej rozważa się funkcje 2π-okresowe. Dla funkcji f o okresie $2T$ możemy wyrazić jej szereg Fouriera w postaci rzeczywistej wzorem

$$\frac{a_0}{2} + \sum_{n=1}^\infty \left(a_n \cos\frac{\pi n x}{T} + b_n \sin\frac{\pi n x}{T}\right),$$

gdzie

$$a_n = \frac{1}{T}\int_{-T}^T f(s)\cos\frac{\pi n s}{T}ds, \quad b_n = \frac{1}{T}\int_{-T}^T f(s)\sin\frac{\pi n s}{T}ds.$$

Twierdzenia o punktowej zbieżności szeregów Fouriera mają ciekawe zastosowania do obliczania sum pewnych szeregów liczbowych. Dla przykładu, definiując funkcję o okresie 1 tak, by dla $x \in [-\frac{1}{2}, \frac{1}{2}]$ przyjmowała wartości

$$f(x) = |x|,$$

mamy $\widehat{f}(0) = \frac{1}{4}$, $\widehat{f}(2j) = 0\,(\forall j \in \mathbb{Z} \setminus \{0\}$ oraz $\widehat{f}(2j+1) = -[(2j+1)\pi]^{-2}$, co daje rozwinięcie w szereg w postaci zespolonej

$$|t| = \frac{1}{4} - \sum_{-\infty}^{+\infty} \frac{1}{\pi^2(2j+1)^2} e^{2\pi i(2j+1)t},$$

więc dla $t = 0$ otrzymujemy rozwinięcie

$$\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \ldots + \frac{1}{(2j+1)^2} + \ldots.$$

Podobnie, rozwijając funkcję o okresie 1 równą x^2 na przedziale $[-\frac{1}{2}, \frac{1}{2}]$ otrzymamy[3] wzór

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

Rozwinięcie funkcji o okresie 1 danej wzorem $f(x) = x$ na tym samym przedziale prowadzi z kolei po odpowiedniej zamianie zmiennych do wzoru

$$\frac{\pi - x}{2} = \sum_{n=1}^{\infty} \frac{\sin \pi n x}{n} \qquad x \in (0, 2\pi).$$

Sprawdzenie tych wzorów można polecić jako interesujące zadanie.

[3] Wzór ten wynika też z tożsamości Parsevala dla funkcji $\frac{\pi - x}{2}$.