

4 Szeregi funkcyjne c.d.

Definicja. W przestrzeni $C^1[a, b]$ funkcji $f : [a, b] \rightarrow \mathbb{R}$ mających ciągle pochodne definiujemy normę wzorem

$$\|f\|_{C^1[a, b]} := |f(a)| + \sup\{|f'(t)| : t \in [a, b]\} \quad \text{czyli} \quad \|f\|_{C^1[a, b]} = |f(a)| + \|f'\|_{[a, b]}. \quad (1)$$

Ciąg (f_n) jest zbieżny do f w tej normie, czyli $\|f_n - f\|_{C^1[a, b]} \rightarrow 0$ przy $n \rightarrow \infty$, gdy ciąg liczbowy $(f_n(a))$ zmierza do $f(a)$ oraz ciąg pochodnych zmierza jednostajnie do pochodnej z f , czyli $f'_n \Rightarrow f'$ w $[a, b]$. Sytuację wyjaśnia następujące twierdzenie, które podam bez dowodu (osoby zainteresowane szczegółami mogą go znaleźć w 6. wykładzie dla kierunku *Matematyka* na tej stronie [www](#)).

Twierdzenie. (O ZBIEŻNOŚCI WRAZ Z POCHODNYMI) Jeśli ciąg pochodnych f'_n funkcji $f_n \in C^1[a, b]$ jest zbieżny jednostajnie w $[a, b]$ do pewnej funkcji g , zaś ciąg $f_n(c)$ jest zbieżny przynajmniej w jednym punkcie $c \in [a, b]$, to ciąg f_n jest zbieżny jednostajnie do pewnej funkcji $f \in C^1[a, b]$ oraz $f' = g$.

Teraz możemy uzupełnić naszą wiedzę o sumie $S(x) := \sum_{n=0}^{\infty} a_n(x - x_0)^n$ szeregu potęgowego wewnątrz jego przedziału zbieżności, czyli na odcinku $D := (x_0 - R, x_0 + R)$, gdzie R jest promieniem zbieżności. Tu k -te sumy częściowe $S_k(x) = \sum_{n=0}^k a_n(x - x_0)^n$ mają pochodne $S'_k(x) = \sum_{n=1}^k n a_n(x - x_0)^{n-1}$ zbieżne jednostajnie w mniejszych przedziałach $D_r := [x_0 - r, x_0 + r]$ ($\forall r < R$). Każdy punkt z przedziału D należy do któregoś takiego D_r wraz z pewnym swoim otoczeniem. Z powyższego twierdzenia wnioskujemy (ponieważ szereg jest zbieżny do a_0 w punkcie x_0 i ma pochodne $S'_k(x)$ jednostajnie zbieżne), że suma szeregu, czyli funkcja $S(x)$ jest różniczkowalna i ma pochodną równą $S'(x) = \sum_{n=1}^{\infty} n a_n(x - x_0)^{n-1}$. Innymi słowy, mamy

Wniosek. Można albo: najpierw liczyć sumę szeregu potęgowego (czyli funkcję $S(x)$), a potem jej pochodną, albo policzyć sumę szeregu pochodnych i na jedno wyjdzie.

Pozwala to np. na wyliczenie $S(x)$ dla pewnych szeregów gdzie np. $a_n = \frac{1}{n}$, $n \geq 1$ lub $a_n = \frac{1}{(n+2)n}$. Chodzi o to, by "pozbywając się n z mianownika" doprowadzić do liczenia sumy zwykłego szeregu geometrycznego (w razie potrzeby - o wszystkich wyrazach pomnożonych np. przez x lub przez x^2). Ponieważ $(x^n)' = n x^{n-1}$, dla $S(x) = a_0 + \sum_{n=1}^{\infty} \frac{1}{n} x^n$ na mocy ostatniego wniosku,

$$S'(x) = \sum_{n=1}^{\infty} n \cdot \frac{1}{n} \cdot x^{n-1}$$

-a to jest suma szeregu geometrycznego, równa $\frac{1}{1-x}$ dla $|x| < 1$. Ponadto $S(0) = a_0$. Stąd

$$S(x) = S(0) + \int_0^x \frac{1}{1-t} dt = a_0 - \ln(1-x)$$

w pierwszym przypadku. W drugim

$$S'(x) = \sum_{n=1}^{\infty} n \cdot \frac{1}{(n+2)n} x^{n-1} = x^{-3} \sum_{n=1}^{\infty} \frac{1}{n+2} x^{n-1+3}.$$

Różniczkując jeszcze raz obliczamy pochodną z ostatniego szeregu:

$$\left(\sum_{n=1}^{\infty} \frac{1}{n+2} x^{n+2}\right)' = \sum_{n=1}^{\infty} x^{n+1} = x^2 \sum_{j=0}^{\infty} x^j = \frac{x^2}{1-x}.$$

Ta ostatnia suma dla $x = 0$ przyjmuje wartość 0, więc jej wartość otrzymamy równą $\int_0^x \frac{t^2}{1-t} dt$. Powtarzamy ten proces, cofając się przez całkowanie do sumy wyjściowej.

Jak łatwo się domysleć, n -a właściwie, $n+1$ możemy usunąć z licznika (np. gdy $b_n = \frac{n+1}{n+2}$) poprzez całkowanie szeregu wyraz po wyrazie, potem różniczkując, by z kolei "wyczyścić mianownik". Na przykład, $\int_0^x \sum_{n=0}^{\infty} (n+1)t^n dt = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$. Stąd $\sum_{n=0}^{\infty} (n+1)x^n = \frac{d}{dx}(\frac{1}{1-x}) = \frac{1}{(1-x)^2}$ dla $|x| < 1$. Podobnie, choć dłużej, wyliczamy $\sum b_n x^n$. Tu prościej jednak będzie przekształcić b_n do postaci $1 - \frac{1}{n+2}$.

Również trzeba stosować powyższy wniosek lub jego odpowiednik dla całek przy rozwijaniu funkcji w szereg Taylora, np. dla rozwinięcia $\ln(1+x)$ wspomnianego w poprzednim wykładzie.

Funkcje wyrażalne poprzez ich szereg Taylora, czyli funkcje analityczne -są z konieczności bardzo regularne, np. są klasy C^∞ wewnątrz obszaru zbieżności. Funkcje wyrażające niektóre wartości fizyczne (skoki napięcia, natężenie fali dźwiękowej - czy ogólnie: *sygnały*) nie są tak regularne, mogą nawet być nieciągłe, przedstawiać impulsy skokowe itp. Aby przedstawić je jako sumy szeregów musimy zamiast wielomianów -użyć bardziej "elastycznych narzędzi" -są nimi np. funkcje trygonometryczne postaci $\sin nx$, $\cos nx$ dla $n \in \mathbb{N}$. Innymi słowy, zaczynamy nowy rozdział.

4.1 Szeregi Fouriera

Szeregi Fouriera stanowią podstawowe narzędzie w analizie sygnałów, teorii komunikacji, termodynamice, czy elektronice. Ich znaczenia dla rozwoju techniki (a także matematyki) nie sposób przecenić. Jedną z różnic będzie sposób, w jaki będzie badana zbieżność szeregu. Do opisu zbieżności posłużymy się pojęciem normy średniokwadratowej, która okaże się właściwym narzędziem do badania zbieżności szeregów Fouriera. Najpierw jednak określimy, jakie funkcje będziemy rozwijali w szeregi Fouriera.

4.1.1 Funkcje okresowe

Rozważać będziemy funkcje zmiennej rzeczywistej, określone na odcinku $[a, b]$. Najczęściej będzie to przedział $[0, 1]$, $[0, 2\pi]$ lub $[-\pi, \pi]$. Przy tak ustalonym przedziale $[a, b]$ o długości $L := b - a$ będziemy mówili, że *funkcja* $f : \mathbb{R} \rightarrow \mathbb{C}$ (lub $f : [a, b] \rightarrow \mathbb{C}$ *jest okresowa*, gdy $f(x+L) = f(x)$, co w przypadku f określonej na $[a, b]$ oznacza jedynie równość $f(a) = f(b)$). Każdą funkcję określoną na przedziale $[a, b]$ (jak również każdą funkcję okresową $f : [a, b] \rightarrow \mathbb{C}$ można jednoznacznie przedłużyć do funkcji okresowej na całej osi rzeczywistej. Jeśli $x \in \mathbb{R}$, to istnieje $k \in \mathbb{Z}$ taka, że $x - kL \in [a, b]$ i wystarczy jako $f(x)$ przyjąć wartość $f(x - kL)$.

W przypadku odcinka $[0, 2\pi]$ (lub $[-\pi, \pi]$) najważniejsze przykłady funkcji okresowych, to $f_n(t) = \sin(nt)$, $g_n(t) = \cos(nt)$, $e_k(t) = e^{ikt}$, gdzie $n \in \mathbb{N} \cup \{0\}$, $k \in \mathbb{Z}$. Funkcje e_k tworzą tzw. *zespólny układ trygonometryczny*, zaś f_n, g_n -rzeczywisty układ trygonometryczny. (Zauważmy, że $\operatorname{Re}(e_n) = g_n$, $\operatorname{Im}(e_n) = f_n$, g_0 jest funkcją stałą równą 1, funkcji f_0 równej stale zero nie zaliczamy do układu.) Fourier zaproponował, by każdą funkcję okresową $f : [0, 2\pi] \rightarrow \mathbb{C}$ przedstawić przy użyciu rzeczywistego szeregu (szereg trygonometryczny) postaci

$$\frac{1}{2}A_0 + \sum_{n=1}^{\infty} A_n \cos(nt) + B_n \sin(nt),$$

lub w postaci zespolonego szeregu Fouriera

$$c_n e^{int}.$$

Jeśli traktujemy te szeregi jako szeregi formalne, to (również formalne) działania algebraiczne, czy operacje różniczkowania, lub całkowania będą dostarczać

prosty wzór, równania różniczkowe przekształcimy w równania algebraiczne dla poszczególnych współczynników. Cała trudność polega na zagadnieniu zbieżności takich szeregów. Jednostajną zbieżność otrzymamy jedynie dla funkcji okresowych klasy C^1 , jednak najważniejsze są szeregi funkcji mniej regularnych, tam nawet zbieżność punktowa może nie mieć miejsca. Zdefiniujemy normę najbardziej odpowiednią dla naszych celów.

4.1.2 Norma L^2 (średniokwadratowa)

Jeśli $f : [a, b] \rightarrow \mathbb{R}$ jest całkowalna w sensie Riemanna, niech

$$\|f\|_2 := \left(\int_a^b |f(t)|^2 dt \right)^{\frac{1}{2}}.$$

Taki sam wzór możemy zastosować dla funkcji zespolonej $f : [a, b] \rightarrow \mathbb{C}$, o ile jej kwadrat modułu jest całkowalny (wystarczy, by części: rzeczywista i urojona z f były całkowalne). Taką normę nazywamy **L^2 -normą, lub normą średniokwadratową**. Nazwę można uzasadnić zwłaszcza w przypadku, gdy długość przedziału $b - a$ wynosi 1. Wówczas cała jest faktycznie średnią wartością z funkcji podcałkowej po przedziale $[a, b]$, licząc tę normę bierzemy wtedy średnią całkową z kwadratu modułu funkcji f . (Zbieżność względem analogicznej normy: $\|f\|_1 := \int_a^b |f(t)| dt$, czyli warunek $\|f_n - f\|_1 \rightarrow 0$ nazywa się czasami „przeciętną zbieżnością” lub zbieżnością w średniej.) To uśrednianie jest typowe dla naszego postrzegania. Na przykład, patrząc na jakiś obraz – nie widzimy poszczególnych punktów, bo te mają „rozmiar zerowy”, lecz średnie nasycenie i jasność konkretnych barw w danych (możliwe, że bardzo małych) obszarach danego obrazu. Podobnie analizujemy dźwięki w danych przedziałach czasowych. W „punkcie czasowym” nie moglibyśmy np. odczuć wysokości tonu (częstotliwości dźwięku). Odpowiedź na to, dlaczego używamy wykładnika 2, a nie 1 lub 4 przyjdzie nieco później – ma ona pewien związek z twierdzeniem Pitagorasa i z regułą równoległoboku, w naszych rozważaniach będziemy, w pewnym sensie, stosować geometrię euklidesową.

Czy jednak $\|\cdot\|_2$ jest rzeczywiście normą? Za chwilę zobaczymy, że zachodzi nierówność trójkąta: $\|f + g\|_2 \leq \|f\|_2 + \|g\|_2$. Warunek jednorodności: $\|\alpha f\|_2 = |\alpha| \|f\|_2$ jest oczywisty, gdyż stałą (dokładniej, $|\alpha|^2$) można wyłączyć przed znak całki. W przypadku funkcji ciągłych to faktycznie jest norma, bo postulat tożsamości: $\|f\|_2 = 0 \Rightarrow f = 0$ można w przypadku f ciągłej dość łatwo wykazać (metodą nie wprost). Ale rozważamy też funkcje nieciągłe i wtedy niestety, są dwa poważne mankamenty:

- postulat tożsamości ($\|f\|_2 = 0 \Rightarrow f = 0$) na ogół nie zachodzi;
- przestrzeń $R[a, b]$ z $\|\cdot\|_2$ normą nie jest zupełna.

Są jednak „standardowe procedury” pozwalające ominąć te niedogodności: Utożsamiamy każde dwie funkcje f_1, f_2 , dla których $\|f_1 - f_2\|_2 = 0$. Jest to proces tworzenia przestrzeni ilorazowej względem tak określonej relacji. Jest to relacja równoważności: przechodniość wynika z nierówności trójkąta, zaś symetria – z jednorodności normy. Na szczęście, ta relacja jest zgodna ze strukturą wektorową: (np. zgodność z dodawaniem oznacza, że gdy $\|f_1 - f_2\|_2 = 0$ oraz $\|g_1 - g_2\|_2 = 0$, to sumy: $h_1 := f_1 + g_1$ oraz $h_2 := f_2 + g_2$ są równoważne, czyli $\|h_1 - h_2\|_2 = 0$). Ponadto wówczas iloczyn przez dany skalar α funkcji f_1, f_2 są równoważne i normy $\|f_1\|_2$ oraz $\|f_2\|_2$ są równe, co pozwala przypisać taką wartość całej klasie równoważności (lub, jak kto woli, klasie abstrakcji), czyli zbiorowi $[f]$ wszystkich funkcji równoważnych z daną $f \in R[a, b]$. Tak więc $\|[f]\|_2 := \|f\|_2$. W ten sposób odzyskujemy postulat jednoznaczności – w przestrzeni wektorowej złożonej z takich klas, gdzie definiujemy działania $[f] + [g] := [f + g]$, $\alpha[f] := [\alpha f]$. Łatwo wykazać, że tak zdefiniowane działania spełniają aksjomaty przestrzeni wektorowej. Wektorem zerowym jest tu klasa równoważności funkcji stałej 0, czyli zbiór $[0] = \{f \in R[a, b] : \|f\|_2 = 0\}$. Na przykład, funkcja równa 1 w punkcie a i w punkcie b , zaś równa zero wewnątrz

naszego odcinka jest okresowa, całkowalna, należy do $[0]$ (proszę sprawdzić). W praktyce klasę równoważności funkcji f zapisuje się takim samym symbolem f , co jest wygodne, lecz należy uważać, gdyż nie można w takim przypadku podać wartości f w danym punkcie. Jedyne, co można wtedy określić, to „wartości przeciętne na zbiorze (miary dodatniej)”. Jeśli w klasie $[f]$ jest funkcja ciągła f_0 , to jest ona wyznaczona jednoznacznie i wtedy jej wartości będą „najbardziej reprezentatywne” dla klasy $[f]$ (o ile taka f_0 jest okresowa).

Drugi problem jest znacznie poważniejszy: okazuje się, że można albo dokonać abstrakcyjnej konstrukcji rozszerzenia do przestrzeni zupełnej (tzw. uzupełnienie przestrzeni), albo rozszerzyć pojęcie całkowalności -wprowadzając tzw. całkowalność w sensie Lebesgue’a.

Uzupełnienie można uzyskać biorąc zbiór klas równoważności ciągów Cauchy’ego względem naszej normy. Ciągi $(f_n), (g_n)$ uznaje się tu za równoważne, jeżeli $\lim \|f_n - g_n\|_2 = 0$. Ciągi takie dodajemy „wyraz po wyrazie”, otrzymując dodawanie wektorowe dla odpowiadających im klas. Normę klasy $[(f_n)]$ definiujemy jako $\lim \|f_n\|_2$. Można wykazać, że te klasy równoważności tworzą przestrzeń Banacha. Szczegóły pomijam. Otrzymana przestrzeń Banacha, oznaczana symbolem $L^2[a, b]$, jest przestrzenią zawierającą klasy równoważności funkcji całkowalnych w sensie Riemanna (względem relacji równoważności $\|f_1 - f_2\|_2 = 0$) i do takich funkcji będziemy się w dalszym ciągu ograniczali.

(Cały dotychczasowy fragment materiału z tego podrozdziału, z wyjątkiem definicji $\|f\|_2$ jest nadobowiazkowy, ma ułatwić korzystanie z bardziej zaawansowanej literatury.)

Kluczowe znaczenie odgrywa **iloczyn skalarny**. Dla $f, g : [a, b] \rightarrow \mathbb{C}$ całkowalnych w sensie Riemanna definiujemy go wzorem

$$\langle f, g \rangle := \int_a^b f(t) \overline{g(t)} dt. \quad (2)$$

Przestrzeń wektorową z iloczynem skalarnym nazywamy przestrzenią Hilberta, jeśli jest ona zupełna względem normy $\|f\| := \sqrt{\langle f, f \rangle}$. Przestrzeń $R[a, b]$ nie jest zupełna, ale jej rozszerzenie: $L^2[a, b]$ już jest przestrzenią Hilberta. Jak już wspomniałem, będziemy jednak starali się ograniczać do funkcji całkowalnych w sensie Riemanna.

Oznaczeń $f \cdot g, f \circ g$ nie możemy stosować dla iloczynu skalarnego funkcji f, g , bo pierwsze oznacza funkcję $([a, b] \ni t \rightarrow f(t)g(t))$, drugie jest symbolem złożenia funkcji. W wielu podręcznikach zamiast nawiasu graniastego, używa się zwykłych nawiasów, co moim zdaniem może mylić się z parą uporządkowaną (f, g) . W przypadku rzeczywistym kreska oznaczająca sprzężenie zespolone nie musi występować, operacja sprzężenia nie zmienia liczby rzeczywistej.

Własności iloczynu skalarnego:

- **liniowość względem pierwszej zmiennej**:

$$\langle \alpha\phi + \beta\psi, g \rangle = \alpha\langle \phi, g \rangle + \beta\langle \psi, g \rangle;$$

- **„skośna symetria”**: $\langle g, f \rangle = \overline{\langle f, g \rangle}$ (w przypadku rzeczywistym jest to zwykła symetria);
- **dodatniość**: $\langle f, f \rangle \geq 0$, przy czym równość mamy wtedy i tylko wtedy, gdy f jest równoważna ze stałą funkcją zero (tak jest wtedy i tylko wtedy, gdy zerowa jest miara Lebesgue’a zbioru tych $t \in [a, b]$, dla których $f(t) \neq 0$.)

Normę średniokwadratową wyliczamy ze wzoru

$$\|f\|_2 = (\langle f, f \rangle)^{\frac{1}{2}}, \quad \text{czyli} \quad \|f\|_2^2 = \langle f, f \rangle.$$

Nierówność Schwarz’a ma postać

$$|\langle f, g \rangle| \leq \|f\|_2 \|g\|_2 \quad (3)$$

Jej dowód dla uproszczenia przedstawię w przypadku rzeczywistym. W dalszym ciągu dla uproszczenia zapisu zamiast $\|\cdot\|_2$ będę używał symbolu $\|\cdot\|$. W oparciu o własności iloczynu skalarnego zauważmy, że

$$\|f + g\|^2 = \|f\|^2 + 2\operatorname{Re}\langle f, g \rangle + \|g\|^2. \quad (4)$$

Wstawiając dla $t \in \mathbb{R}$ zamiast g wektor tg otrzymamy stale nieujemny trójmian kwadratowy zmiennej t postaci $t^2\|g\|^2 + t \cdot 2\langle f, g \rangle + \|f\|^2$. (Pozwoliłem sobie na opuszczenie "Re", bo to przypadek rzeczywisty.) Wyróżnik "delta" jest więc niedodatni, co daje nierówność $4\langle f, g \rangle^2 - 4\|f\|^2\|g\|^2 \leq 0$, co po trywialnym przekształceniu i po spierwiastkowaniu stronami daje dowodzoną nierówność Schwarza. Wynika z niej nierówność trójkąta dla normy $\|\cdot\|$. Sprawdzamy ją podnosząc strony do kwadratu: ma być $\|f + g\|^2 \leq (\|f\| + \|g\|)^2$, co po zastosowaniu powyższego wzoru dla lewej strony i wzoru na kwadrat sumy liczb -dla prawej, sprowadza zagadnienie do nierówności Schwarza.

Definicja. Mówimy, że wektory f, g są ortogonalne, czyli prostopadłe względem danego iloczynu skalarnego $\langle \cdot, \cdot \rangle$, gdy $\langle f, g \rangle = 0$. Fakt ten oznaczamy zapisując $f \perp g$. Ze wzoru (4) wynika dla $f \perp g$ słynne [twierdzenie Pitagorasa](#):

$$\langle f, g \rangle = 0 \Rightarrow \|f + g\|^2 = \|f\|^2 + \|g\|^2.$$

Zauważmy, że gdy $f \perp g$, to również $g \perp f$, natomiast relacja $f \perp f$ może zachodzić jedynie dla $\|f\|_2 = 0$ (czyli gdy $f = 0$ prawie wszędzie).

Definicja. Układ wektorów $(f_j)_{j \in J}$ nazwiemy *układem ortogonalnym*, gdy $j \neq k \Rightarrow f_j \perp f_k$. Układ ortogonalny wektorów o normie 1 nazwiemy *układem ortonormalnym*. Układ $(e_j)_{j \in J}$ w przestrzeni H z iloczynem skalarnym nazwiemy *bazą ortonormalną*, gdy jest to taki układ ortonormalny, że jedynym wektorem prostopadłym do wszystkich e_j jest wektor zerowy:

$$(\forall_{j \in J} f \perp e_j) \Rightarrow f = 0. \quad (5)$$

Warunek ten, znany jako *warunek zupełności układu* jest równoważny temu, by każdy element $f \in H$ z dowolną dokładnością dał się przybliżyć kombinacjami liniowymi wektorów e_j .