Videregående vektorrumsteori

Bergfinnur Durhuus

Januar 2003

Indhold

Indledning

Vektorrumsformalismen viser sig at være nyttig til formulering og analyse af mange forskelligartede fysiske problemstillinger. Vi har allerede set forskellige begreber fra vektorrumsteori dukke op i forbindelse med løsning af differentialligninger, både sædvanlige og partielle, samt i Fourierrækketeorien og mere generelt i Sturm-Liouville teorien. Eksempelvis fremstillede vi løsninger til varmeledningsligningen som linearkombinationer af løsninger til det tilsvarende homogene problem, og i Sturm-Liouville teorien bemærkede vi vigtigheden af ortogonaliteten af egenfunktionerne.

Som et yderligere eksempel på en fysisk disciplin, hvori vektorrumsteori indgår på afgørende vis, kan nævnes kvantemekanik, som vi lejlighedsvis kommer til at berøre i disse noter.

Fælles for de vektorrum, der indgår i ovennævnte eksempler, er, at de er af uendelig dimension. Vores mål i det følgende er at udvide dele af den fra første år kendte teori om endeligdimensionale indre produkt rum til en vis klasse af uendeligdimensionale vektorrum, kaldet *Hilbert rum*, som er stor nok til at omfatte ovennævnte eksempler og mere til. Dette muliggør samtidig en forenklet formulering og en præcisering af en del af resultaterne vedrørende differentialligninger, vi har mødt i Kreyszigs bog.

I Kapitel 1 genopfriskes mestendels velkendte begreber fra elementær vektorrumsteori, dog udvidet til komplekse vektorrum. I Kapitel 2 indføres Hilbert rum og grundlæggende resultater vedrørende disse udledes og illustreres. Kapitel 3 omhandler lineære afbildninger mellem Hilbert rum og i særdeleshed differentialoperatorer. Et appendiks er tilføjet, som uddyber nogle aspekter af Kapitel 2.

I disse noter henviser betegnelsen LA til lærebogen Robert Messer: Linear Algebra, 1994-udgaven, og AEM til Erwin Kreyszig: Advanced Engeneering Mathematics, 8. udgave. Messers bog benyttes i Matematik 1, men de resultater, der henvises til, vil også kunne findes i Matematik A lærebogen af Edwards og Penney.

1 Elementær vektorrumsteori

1.1 Nogle eksempler på vektorrum

Et vektorrum er som bekendt en mængde V udstyret med en addition, som definerer summen $v + w$ af to vektorer v, w i V, og en skalarmultiplikation, som definerer produktet λv af en skalar (dvs. et tal) λ og en vektor v i V. Disse opfylder visse standard regneregler, som angivet i LA Definition 1.1, og som vi udlader at gengive her. Afhængig af om skalarerne er reelle tal eller komplekse tal taler vi om et *reelt vektorrum*, henholdsvis *komplekst vektorrum*, eller om et *vektorrum over* $\mathbb{R}$, henholdsvis $\mathbb{C}$. Nedenfor gives nogle eksempler.

Talrum. Fra Matematik 1 er talrummene $\mathbb{R}^n$, hvor $n = 1, 2, 3, \ldots$, velkendte reelle vektorrum. Tilsvarende er de komplekse talrum

$$\mathbb{C}^n = \{(x_1, \ldots, x_n) | \, x_1, \ldots, x_n \in \mathbb{C}\}$$

eksempler på komplekse vektorrum, når addition og skalarmultiplikation defineres koordinatvis (som for $\mathbb{R}^n$), dvs.

$$\begin{aligned} (x_1, \ldots, x_n) + (y_1, \ldots, y_n) &= (x_1 + y_1, \ldots, x_n + y_n) \\ \lambda(x_1, \ldots, x_n) &= (\lambda x_1, \ldots, \lambda x_n) \, . \end{aligned}$$

Bemærk dog, at $\mathbb{C}^n$ også kan opfattes som et reelt vektorrum (idet λ ovenfor da blot antages at tilhøre $\mathbb{R}$ i stedet for $\mathbb{C}$), hvorimod $\mathbb{R}^n$ ikke kan opfattes som komplekst vektorrum (idet multiplikation af et reelt talsæt med et komplekst tal i almindelighed ikke tilhører $\mathbb{R}^n$).

Funktionsrum. Af større interesse for os i det følgende er vektorrum bestående af funktioner. Hvis eksempelvis $[a, b]$ er et interval, vil vi ligesom Messer bruge betegnelsen $F([a, b])$ for vektorrummet af funktioner defineret på $[a, b]$. Hvis ikke andet er anført, vil det være underforstået, at funktionerne antager komplekse værdier, hvorved $F([a, b])$ bliver et komplekst vektorrum. Hvis vi ønsker at understrege, at der er tale om reelle funktioner, henholdsvis komplekse funktioner, skriver vi $F([a, b], \mathbb{R})$, henholdsvis $F([a, b], \mathbb{C})$. Det er her naturligvis ikke afgørende, at funktionerne er definerede på et interval. Definitionsmængden kan være en vilkårlig mængde M, f.eks. $\mathbb{R}^2$ svarende til funktioner af to variable, idet addition og skalarmultiplikation jo defineres punktvis, dvs. for $f, g \in F(M)$ og $\lambda \in \mathbb{C}$ er funktionerne $f + g$ og λf givne ved

$$\begin{aligned} (f + g)(x) &= f(x) + g(x) \; , \quad x \in M \; , \\ (\lambda f)(x) &= \lambda f(x) \; , \quad x \in M \; . \end{aligned}$$

Som det vil fremgå, er det i konkrete problemstillinger sjældent rummene bestående af *alle* funktioner på en mængde, der er af interesse, men snarere delmængder heraf, som i sig selv er vektorrum (med samme addition og skalarmultiplikation), dvs. underrum af disse, jvf. Definition 1.10 i LA. Det kan f.eks. være rummet $C([a, b])$ bestående af kontinuerte funktioner på $[a, b]$. Denne er et underrum af $F([a, b])$, da summen af to kontinuerte funktioner er kontinuert, og ligeledes en konstant gange en kontinuert funktion, jvf. Theorem 1.11 i LA. Et andet eksempel er rummet $D([a,b])$ bestående af differentiable funktioner på $[a,b]$. Vi skal i Kapitel 2 indføre flere funktionsrum af interesse.

Følgerum. Vi får også lejlighed til at beskæftige os med vektorrum bestående af følger. Med $\ell(\mathbb{N})$ vil vi betegne mængden af alle komplekse talfølger med sædvanlig addition og skalarmultiplikation, dvs. hvis $(x_1, x_2, x_3, \dots)$ og $(y_1, y_2, y_3, \dots)$ er to talfølger og $\lambda \in \mathbb{C}$, sættes

$$\begin{aligned}(x_1, x_2, x_3, \dots) + (y_1, y_2, y_3, \dots) &= (x_1 + y_1, x_2 + y_2, x_3 + y_3, \dots) \;, \\ \lambda(x_1, x_2, x_3, \dots) &= (\lambda x_1, \lambda x_2, \lambda x_3, \dots) \;.\end{aligned}$$

Faktisk kan dette eksempel anskues som et specielt tilfælde af det foregående, svarende til $M = \mathbb{N}$, eftersom en følge $(x_1, x_2, x_3, \dots)$ kan opfattes som en funktion på $\mathbb{N}$, nemlig funktionen, der afbilder n i x_n, for $n \in \mathbb{N}$.

Som et eksempel på et underrum af $\ell(\mathbb{N})$ har vi rummet $\ell_0(\mathbb{N})$ bestående af følger, der er 0 fra et vist trin, dvs, følger med kun endelig mange elementer forskellige fra 0. Det overlades til læseren at overbevise sig om, at der her faktisk er tale om et underrum (jvf. Opgave 1.1).

1.2 Grundlæggende begreber

Lineær (u)afhængighed. Addition og skalarmultiplikation giver os mulighed for at danne *linearkombinationer* af vektorer i et vektorrum V. For $v_1, \dots, v_n \in V$ og skalarer $\lambda_1, \dots, \lambda_n$ kaldes udtrykket

$$\lambda_1 v_1 + \cdots + \lambda_n v_n$$

linearkombinationen af $v_1, \dots, v_n$ med koefficienter $\lambda_1, \dots, \lambda_n$. Da der gælder, at $0v = \underline{0}$ for alle vektorer v i V (hvor $\underline{0}$ betegner nulvektoren i V), er linearkombinationen af $v_1, \dots, v_n$, hvis koefficienter alle er 0, lig med $\underline{0}$. Såfremt alle andre linearkombinationer af $v_1, \dots, v_n$ er forskellige fra $\underline{0}$, siges $v_1, \dots, v_n$ at være *lineært uafhængige*. M.a.o. er $v_1, \dots, v_n$ lineært uafhængige, hvis der gælder

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = \underline{0} \quad \Rightarrow \quad \lambda_1 = \cdots = \lambda_n = 0 \;.$$

Vi siger i dette tilfælde også, at sættet $(v_1, \dots, v_n)$ er lineært uafhængigt.

Vi udvider begrebet lineær uafhængighed til også at omfatte følger af vektorer ved at definere følgen $(v_1, v_2, v_3, \dots)$ til at være lineært uafhængig, hvis sættet $(v_1, \dots, v_n)$ er lineært uafhængigt for alle $n \in \mathbb{N}$. Tit vil vi omtale følgen $(v_1, v_2, v_3, \dots)$ som et (uendeligt) sæt af vektorer i V. Når vi derfor omtaler *sæt af vektorer* i det følgende, kan der være tale om endelige eller uendelige sæt.

Et sæt af vektorer, der ikke er lineært uafhængigt siges selvfølgelig at være *lineært afhængigt*.

Span. Det er klart, at summen af en linearkombination af $v_1, \dots, v_n$ og en linearkombination af $w_1, \dots, w_m$ er en linearkombination af $v_1, \dots, v_n, w_1, \dots, w_m$, og at produktet af en skalar λ og en linearkombination af $v_1, \ldots, v_n$ er en linearkombination af $v_1, \ldots, v_n$. Heraf ses, at for en vilkårlig delmængde $A \subseteq V$ gælder, at mængden span$A$ bestående af alle linearkombinationer af vektorer fra A (se Definition 3.2 i LA), er et underrum af V (iflg. Theorem 1.11 i LA). Det kaldes *underrummet udspændt af* A og er faktisk det mindste underrum af V, der indeholder A, jvf. Opgave 1.2.

Basis og dimension. Et sæt $(v_1, \ldots, v_n)$ af vektorer i V siges at være en basis for V, hvis det er lineært uafhængigt og udspænder V, jvf. Definition 3.6 i LA. Det huskes fra Matematik 1, at dette betyder, at enhver vektor v i V på entydig måde kan skrives som en linearkombination af $v_1, \ldots, v_n$, og de tilsvarende koefficienter kaldes *koordinaterne* for v m.h.t. basen $(v_1, \ldots, v_n)$.

Hvis V har en basis $(v_1, \ldots, v_n)$, gælder, at et vilkårligt lineært uafhængigt sæt i V har højst n vektorer (Theorem 3.9 i LA), og et vilkårligt sæt, der udspænder V, har mindst n vektorer (Theorem 3.11 i LA). Heraf følger, at antallet af vektorer i enhver basis for V er netop n, som kaldes *dimensionen* af V og betegnes $\dim V$.

Som vi skal se, findes der i ovenfor omtalte funktionsrum og følgerum uendelige lineært uafhængige sæt, hvorfor rummenes dimension ikke kan være endelig.

Eksempel 1.1. a) Lad os se på vektorerne $v_1 = (1,2,1), v_2 = (0,2,3)$, $v_3 = (2,2,-1)$ i vektorrummet $\mathbb{C}^3$ og sætte

$$U = \text{span}\{v_1, v_2, v_3\} \, .$$

De tre vektorer er lineært afhængige, da

$$v_3 = 2v_1 - v_2 \, . \tag{1.1}$$

Altså er $v_3 \in \text{span}\{v_1, v_2\}$, og derfor er

$$U = \text{span}\{v_1, v_2\} \, ,$$

eftersom enhver linearkombination af v_1, v_2, v_3 også kan skrives som en linearkombination af v_1, v_2 ved brug af (1.1). Dette argument viser mere generelt, at der for vilkårlige delmængder A, B af et vektorrum V gælder, at

$$\text{span}(A \cup B) = \text{span}A \quad \text{hvis} \quad B \subseteq \text{span}A$$

(jvf. Opgave 1.3).

Vektorerne v_1 og v_2 ovenfor er lineært uafhængige, da de ikke er proportionale. Derfor udgør de en basis for U, så $\dim U = 2$.

b) Vi betragter vektorerne $v_1 = 1$, $v_2 = \cos x$, $v_3 = \cos 2x$ i vektorrummet $C([-\pi, \pi])$. Her betegner 1 som sædvanlig den konstante funktion 1 på $[-\pi, \pi]$.

Disse tre vektorer er lineært uafhængige. Antag nemlig, at

$$\lambda_1 + \lambda_2 \cos x + \lambda_3 \cos 2x = 0 \quad \text{for} \quad x \in [-\pi, \pi] \, .$$

Indsættes heri de tre værdier $x = 0$, $\frac{\pi}{2}$, π, fås de tre ligninger

$$\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 &= 0 \\ \lambda_1 + 0 \ - \lambda_3 &= 0 \\ \lambda_1 - \lambda_2 + \lambda_3 &= 0 \, , \end{aligned}$$

som let ses kun at have løsningen $\lambda_1 = \lambda_2 = \lambda_3 = 0$. (Af Eksempel 1.6 c) nedenfor vil det ligeledes fremgå, at v_1, v_2, v_3 er lineært uafhængige.) Altså har

$$U = \text{span}\{v_1, v_2, v_3\} \ .$$

dimension 3.

Det er klart, at $1 + \cos x \in U$. Idet vi genkalder formlen

$$\cos^2 \frac{1}{2}x = \frac{1}{2}(1 + \cos x) \ ,$$

følger det, at $\cos^2 \frac{1}{2}x \in U$. Tilsvarende ses, at $\cos^2 x \in U$. Derimod tilhører $\cos^3 x$ ikke U, jvf. Opgave 1.4.

c) I vektorrummet $\ell_0(\mathbb{N})$ sætter vi $\varepsilon_1 = (1, 0, 0, 0, \dots)$, $\varepsilon_2 = (0, 1, 0, 0, \dots)$, $\varepsilon_3 = (0, 0, 1, 0, \dots)$, $\dots$, dvs. ε_n er følgen, hvis elementer alle er 0 pånær det n'te, som er 1. Disse følger tilhører klart $\ell_0(\mathbb{N})$, og for $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ er

$$\lambda_1\varepsilon_1 + \cdots + \lambda_n\varepsilon_n = (\lambda_1, \lambda_2, \dots, \lambda_n, 0, 0, \dots) \ . \tag{1.2}$$

Dette viser, at enhver følge i $\ell_0(\mathbb{N})$ kan skrives som en linearkombination af vektorerne $\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots$, altså

$$\ell_0(\mathbb{N}) = \text{span}\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots\} \ .$$

Det ses også umiddelbart af (1.2), at sættet $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots)$ er lineært uafhængigt. Det er dermed et eksempel på en (uendelig) basis for $\ell_0(\mathbb{N})$.

1.3 Indre produkt rum

Indre produkt på reelle vektorrum er defineret i LA Definition 4.1. Vi udvider nu dette til også at omfatte komplekse vektorrum. Som det vil ses, består forskellen på det reelle og det komplekse tilfælde alene i, at kompleks konjugering af visse udtryk, specielt i (1.4) og (1.8) nedenfor, bortfalder i det reelle tilfælde.

Indre produkt. For hvert par af vektorer v, w i et komplekst vektorrum V er det *indre produkt* $\langle v, w\rangle$ et komplekst tal, som kræves at opfylde følgende fire betingelser:

$$\begin{aligned} &i) && \langle v, v\rangle > 0 \quad \text{for} \quad v \neq \underline{0} && (1.3)\\ &ii) && \langle v, w\rangle = \overline{\langle w, v\rangle} && (1.4)\\ &iii) && \langle v, w_1 + w_2\rangle = \langle v, w_1\rangle + \langle v, w_2\rangle && (1.5)\\ &iv) && \langle v, \lambda w\rangle = \lambda\langle v, w\rangle \ . && (1.6) \end{aligned}$$

Når der på et vektorrum er givet et sådant indre produkt, taler vi om et *indre produkt rum*.

Inden vi ser på nogle eksempler, er det værd at knytte nogle bemærkninger til definitionen:

- Bemærk det skarpe ulighedstegn i (1.3) for $v \neq \underline{0}$. Sættes $\lambda = 0$ i (1.6), fås

$$\langle v, \underline{0}\rangle = 0 \qquad \text{for alle } v \in V \ ,$$

og derfor specielt $\langle \underline{0}, \underline{0} \rangle = 0$. D.v.s. $\langle v, v \rangle = 0$, hvis og kun hvis $v = \underline{0}$.
• Betingelserne *iii*) og *iv*) udtrykker, at det indre produkt er *lineært i anden variabel* og er ækvivalente med kravet

$$\langle v, \lambda_1 w_1 + \lambda_2 w_2 \rangle = \lambda_1 \langle v, w_1 \rangle + \lambda_2 \langle v, w_2 \rangle$$

eller mere generelt

$$\langle v, \lambda_1 w_1 + \cdots + \lambda_n w_n \rangle = \lambda_1 \langle v, w_1 \rangle + \cdots + \lambda_n \langle v, w_n \rangle \tag{1.7}$$

(jvf. Opgave 1.5).
• Kombineres *ii*) med *iii*) og *iv*) fås (jvf. Opgave 1.5)

$$\begin{aligned} \langle v_1 + v_2, w \rangle &= \langle v_1, w \rangle + \langle v_2, w \rangle \\ \langle \lambda v, w \rangle &= \overline{\lambda} \langle v, w \rangle \end{aligned}$$

eller ækvivalent hermed

$$\langle \lambda_1 v_1 + \cdots + \lambda_n v_n, w \rangle = \overline{\lambda}_1 \langle v_1, w \rangle + \cdots + \overline{\lambda}_n \langle v_n, w \rangle \ . \tag{1.8}$$

Vi siger, at det indre produkt er *konjugeret lineært* i første variabel. I det reelle tilfælde bortfalder den komplekse konjugering, så det indre produkt er lineært i begge variable, som det er kendt fra Matematik 1.
• Det er ikke afgørende for den følgende teori, at vi har krævet linearitet i anden variabel i stedet for i første variabel. Faktisk er den sædvanlige konvention i matematiklitteratur linearitet i første variabel, hvorimod fysiklitteratur, specielt vedrørende kvantemekanik, fortrinsvis benytter den her valgte konvention.

Eksempel 1.2. a) *Det sædvanlige indre produkt på* $\mathbb{C}^n$. Dette defineres ved

$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = \overline{x_1} y_1 + \cdots + \overline{x_n} y_n \ . \tag{1.9}$$

Det overlades til læseren at eftervise, at kravene $i) - iv)$ er opfyldt (jvf. Opgave 1.6). Bemærk især den komplekse konjugering, der indgår i definitionen, som er afgørende for positivitet af

$$\langle (x_1, \dots, x_n), (x_1, \dots, x_n) \rangle = |x_1|^2 + \cdots + |x_n|^2 \ . \tag{1.10}$$

b) *Det sædvanlige indre produkt på* $C([a, b])$. For $f, g \in C([a, b])$ defineres dette ved

$$\langle f, g \rangle = \int_a^b \overline{f(x)} g(x) dx \ . \tag{1.11}$$

Også her er den komplekse konjugering afgørende for positivitet af

$$\langle f, f \rangle = \int_a^b |f(x)|^2 dx \ . \tag{1.12}$$

Bemærk, at det indre produkt (1.11) ikke giver mening for vilkårlige funktioner i $F([a, b])$, da integralet kun er defineret for kontinuerte funktioner. Bemærk også, at en kompleks funktion $f = f_1 + if_2$, hvor f_1 og f_2 betegner henholdsvis real- og imaginærdelen af f, er kontinuert, netop hvis både f_1 og f_2 er kontinuerte (reelle) funktioner, og integralet af f er da defineret ved at integrere real- og imaginærdel hver for sig, dvs.

$$\int_a^b f(x)dx = \int_a^b f_1(x)dx + i\int_a^b f_2(x)dx \ .$$

c) *Indre produkt på* $C([a,b])$ *med vægtfunktion.* For enhver positiv kontinuert funktion ρ på $[a,b]$ kan vi definere et indre produkt $\langle\cdot,\cdot\rangle_\rho$ på $C([a,b])$ ved

$$\langle f,g\rangle_\rho = \int_a^b \overline{f(x)}g(x)\rho(x)dx \ . \tag{1.13}$$

Dette kaldes det indre produkt med *vægtfunktion* ρ. For $\rho = 1$ fås det sædvanlige indre produkt. Det overlades til læseren at eftervise, at $\langle\cdot,\cdot\rangle_\rho$ opfylder kravene $i)-iv)$, jvf. Opgave 1.7.

I diskussionen af Bessel funktioner i AEM Theorem 2 p. 232 indgik det indre produkt $\langle\cdot,\cdot\rangle_\rho$ på $C([0,R])$, hvor $\rho(x) = x$, og det indre produkt (1.11) på $C([-1,1])$ optrådte i forbindelse med Legendre polynomier i Eksempel 4 p.237. Endvidere har vi set det sædvanlige indre produkt på $C([-\pi,\pi])$ i forbindelse med Fourier rækker i Afsnit 10.2 i AEM.

d) *Det sædvanlige indre produkt på* $\ell_0(\mathbb{N})$. Dette defineres ved

$$\langle(x_1,x_2,x_3,\dots),(y_1,y_2,y_3,\dots)\rangle = \sum_{n=1}^{\infty}\overline{x_n}y_n \ . \tag{1.14}$$

Her bemærkes, at rækken på højresiden er konvergent, da kun endelig mange af dens led er forskellige fra 0 ifølge definitionen af $\ell_0(\mathbb{N})$. Derimod kan (1.14) naturligvis ikke bruges som definition af et indre produkt på $\ell(\mathbb{N})$, da rækken generelt vil være divergent. Læseren bedes i Opgave 1.8 om at eftervise, at (1.14) definerer et indre produkt på $\ell_0(\mathbb{N})$. Vi skal i Kapitel 2 udvide dette til et større vektorrum.

Cauchy-Schwarz' ulighed. Et nøgleresultat vedrørende indre produkt rum er Cauchy-Schwarz' ulighed:

$$\boxed{|\langle v,w\rangle| \le \|v\|\|w\|} \tag{1.15}$$

hvor *normen* $\|v\|$ af $v \in V$ er defineret ved

$$\|v\| = \sqrt{\langle v,v\rangle} \ . \tag{1.16}$$

Denne ulighed findes i LA Theorem 4.8 for reelle vektorrum. For komplekse vektorrum kan den indses ved en mindre tilpasning af argumentet i LA, som vi giver her for fuldstændigheds skyld:
Hvis v eller w er $\underline{0}$, står der 0 på begge sider af uligheden, så vi antager, at $v \neq 0$ og $w \neq 0$. Da er også $\|v\| \neq 0$ og $\|w\| \neq 0$, og ved brug af $i) - iv)$ får vi

$$\begin{aligned} 0 &\le \langle \|v\|^{-1}v - \|w\|^{-1}w, \|v\|^{-1}v - \|w\|^{-1}w\rangle \\ &= \|v\|^{-2}\langle v,v\rangle + \|w\|^{-2}\langle w,w\rangle - \|v\|^{-1}\|w\|^{-1}(\langle v,w\rangle + \langle w,v\rangle) \\ &= 2 - 2\|v\|^{-1}\|w\|^{-1}\mathrm{Re}\langle v,w\rangle \ , \end{aligned}$$

og dermed
$$\mathrm{Re}\langle v,w\rangle \le \|v\|\|w\| \;, \tag{1.17}$$
hvor $\mathrm{Re}\, z$ står for realdelen af det komplekse tal z. Denne ulighed gælder altså for alle $v, w \in V \setminus \{\underline{0}\}$, så vi kan erstatte v med $e^{i\theta}v$, hvor θ er et argument til det komplekse tal $\langle v,w\rangle$, således at $\langle v,w\rangle = |\langle v,w\rangle|e^{i\theta}$. Da fås, at venstresiden af (1.17) bliver
$$\mathrm{Re}\langle e^{i\theta}v,w\rangle = \mathrm{Re}(e^{-i\theta}\langle v,w\rangle) = |\langle v,w\rangle| \;,$$
mens højresiden forbliver uændret $\|v\|\|w\|$, idet
$$\begin{aligned}\|e^{i\theta}v\| &= \sqrt{\langle e^{i\theta}v, e^{i\theta}v\rangle} = \sqrt{e^{-i\theta}e^{i\theta}\langle v,v\rangle} \\ &= \sqrt{\langle v,v\rangle} = \|v\| \;.\end{aligned}$$
Dette viser (1.15).

Eksempel 1.3. Cauchy-Schwarz' ulighed har et væld af nyttige konsekvenser ved anvendelse på passende vektorer i passende indre produkt rum:

- Lader vi f.eks. v og w være vektorerne x og $\sin x$ i $C([0,\pi])$, fås
$$\left(\int_0^\pi x\sin x dx\right)^2 \le \int_0^\pi x^2 dx \int_0^\pi \sin^2 x dx \;. \tag{1.18}$$
Her kan begge sider naturligvis udregnes, og man kan derved direkte verificere, at uligheden er opfyldt, jvf. Opgave 1.10.

- Af mere almen interesse er uligheden, man får ved at lade v være vektoren $(1,1,\ldots,1)$ i $\mathbb{C}^n$ og $w = (x_1,\ldots,x_n)$ en vilkårlig vektor i $\mathbb{C}^n$. Med det sædvanlige indre produkt (1.9) fås da, at $\langle v,w\rangle = x_1 + \cdots + x_n$ og $\|v\| = \sqrt{n}$. Iflg. (1.15) er derfor
$$|x_1 + \cdots + x_n|^2 \le n(|x_1|^2 + \cdots + |x_n|^2) \tag{1.19}$$
for vilkårlige komplekse tal $x_1,\ldots,x_n$.

- Lad endelig v være den konstante funktion 1 på $[a,b]$ og w være en vilkårlig kontinuert funktion f på $[a,b]$. Med det sædvanlige indre produkt (1.11) på $C([a,b])$ fås da, at $\langle v,w\rangle = \int_a^b f(x)dx$ og $\|v\| = \sqrt{b-a}$, og (1.15) giver
$$\left|\int_a^b f(x)dx\right|^2 \le (b-a)\int_a^b |f(x)|^2 dx \;. \tag{1.20}$$

Afstand og konvergens af følger. I LA Theorem 4.9 er det vist, at normen $\|\cdot\|$ givet ved (1.16) besidder følgende tre vigtige egenskaber:
$$\begin{aligned}\|v\| &> 0 \quad \text{for} \quad v \neq \underline{0} &(1.21)\\ \|\lambda v\| &= |\lambda|\|v\| &(1.22)\\ \|v+w\| &\le \|v\| + \|w\| \;, &(1.23)\end{aligned}$$
hvor (1.23) kaldes *trekantsuligheden*. Disse relationer gælder også for komplekse vektorrum og vises pånær mindre ændringer på samme måde som i LA p. 150-151 og overlades til læseren, jvf. Opgave 1.11.

Vi opfatter $\|v-w\|$ som et mål for afstanden mellem v og w. I talrummene $\mathbb{R}^n$ og $\mathbb{C}^n$ med sædvanligt indre produkt (1.9) er der tale om den sædvanlige euklidiske afstand mellem punkter $x=(x_1,\dots,x_n)$ og $y=(y_1,\dots,y_n)$ givet ved

$$\|x-y\| = \sqrt{|x_1-y_1|^2+\cdots+|x_n-y_n|^2}\ .$$

For funktioner f og g i $C([a,b])$ med sædvanligt indre produkt (1.11) er afstanden givet ved

$$\|f-g\| = \left(\int_a^b |f(x)-g(x)|^2 dx\right)^{\frac{1}{2}}\ .$$

Dette generaliserede afstandsbegreb sætter os i stand til at udvide konvergensbegrebet fra tal- eller punktfølger til følger af vektorer i et vilkårligt indre produkt rum på følgende vis.

Definition 1.4. En følge $(v_1,v_2,v_3,\dots)$ i et indre produkt rum V siges at konvergere imod vektoren $v\in V$, hvis der gælder, at afstanden mellem v_n og v går imod 0 for $n\to\infty$, altså hvis

$$\|v_n-v\|\to 0 \quad \text{for} \quad n\to\infty\ ,$$

og vi skriver da

$$\lim_{n\to\infty} v_n = v \qquad \text{eller} \qquad v_n\to v \quad \text{for} \quad n\to\infty\ .$$

Eksempel 1.5. Vi betragter følgen $(x,x^2,x^3,\dots)$ af funktioner i vektorrummet $C([0,1])$ med sædvanligt indre produkt. Denne følge konvergerer imod nulfunktionen, fordi

$$\|x^n-0\|^2 = \int_0^1 |x^n-0|^2dx = \int_0^1 x^{2n}dx = \frac{1}{2n+1} \to 0 \ \text{ for } \ n\to\infty\ .$$

Bemærk, at for hvert fast $x\in[0,1[$ konvergerer talfølgen $(x,x^2,x^3,\dots)$ imod 0, mens den for $x=1$ konvergerer imod (er konstant lig med) 1. Dvs. at funktionsfølgen $(x,x^2,x^3,\dots)$ ikke konvergerer imod 0 i alle punkter af $[0,1]$, men ikke desto mindre konvergerer den imod nulfunktionen i henhold til Definition 1.4.

Man kalder tit konvergens af en funktionsfølge $(f_1,f_2,f_3,\dots)$ i indre produkt rummet $C([a,b])$ for *konvergens i middel*. Er grænseværdien f, betyder dette altså, at

$$\int_a^b |f_n(x)-f(x)|^2dx \to 0 \quad \text{for} \quad n\to\infty\ ,$$

dvs. middelværdien af afstandskvadratet $|f_n(x)-f(x)|^2$ over intervallet $[a,b]$ skal gå imod 0.

Hvis der derimod gælder, at

$$f_n(x)\to f(x) \ \text{ for } \ n\to\infty \quad \text{for hvert } x\in[a,b]\ ,$$

taler vi om *punktvis konvergens* af $(f_1,f_2,f_3,\dots)$ imod f. Det er konvergens i denne betydning, der indgår i diskussionen af potensrækker i Matematik 1, og tillige i Theorem 1 p. 535 i AEM. Eksemplet ovenfor viser, at de to former for konvergens er forskellige (omend ikke helt uafhængige af hinanden).

Bemærk endelig, at vi i Afsnit 10.7 i AEM allerede har været inde på begrebet konvergens i middel, blot var det der ikke sat i forbindelse med indre produkt rum.

1.4 Ortogonalitet

Ortogonalsæt og Pythagoras' sætning. Som bekendt siges to vektorer v og w i et indre produkt rum V at være *ortogonale*, eller *vinkelrette på hinanden*, såfremt $\langle v, w\rangle = 0$. Et sæt $(v_1, v_2, v_3, \dots)$ af vektorer i V er et *ortogonalsæt*, hvis v_n og v_m er ortogonale for alle $n \neq m$. Hvis yderligere $\|v_n\| = 1$ for alle $n \in \mathbb{N}$, siges sættet at være et *ortonormalsæt*. Vi husker endvidere (jvf. LA Theorem 4.14), at et ortogonalsæt $(v_1, v_2, v_3, \dots)$, der ikke indeholder nulvektoren, og dermed i særdeleshed et ortonormalsæt, er lineært uafhængigt. Antag nemlig, at $\lambda_1 v_1 + \cdots + \lambda_n v_n = \underline{0}$. Tages indre produkt med v_i, hvor $i = 1, \dots, n$, på begge sider og benyttes (1.7), giver venstresiden

$$\lambda_1\langle v_i, v_1\rangle + \cdots + \lambda_n\langle v_i, v_n\rangle = \lambda_i\|v_i\|^2 ,$$

hvorimod højresiden giver $\langle v_i, \underline{0}\rangle = 0$. Altså er $\lambda_i\|v_i\|^2 = 0$ og dermed $\lambda_i = 0$, for hvert $i = 1, \dots, n$. Dette viser, at sættet $(v_1, \dots, v_n)$ er lineært uafhængigt for ethvert $n \in \mathbb{N}$, og dermed påstanden.

Vi noterer os med det samme følgende nyttige udgave af *Pythagoras' sætning*, som siger, at der for et endeligt ortogonalsæt $(v_1, \dots, v_n)$ i et indre produkt rum gælder, at

$$\|v_1 + \cdots + v_n\|^2 = \|v_1\|^2 + \cdots + \|v_n\|^2 . \tag{1.24}$$

Dette følger ved udregning af venstresiden:

$$\begin{aligned}
\langle v_1 + \cdots + v_n, v_1 + \cdots + v_n\rangle &= \langle \sum_{i=1}^n v_i, \sum_{j=1}^n v_j \rangle \\
&= \sum_{i=1}^n \sum_{j=1}^n \langle v_i, v_j\rangle \\
&= \sum_{i=1}^n \langle v_i, v_i\rangle \\
&= \|v_1\|^2 + \cdots + \|v_n\|^2 ,
\end{aligned}$$

hvor det tredie lighedstegn følger af ortogonaliteten, som indebærer, at kun leddene $\langle v_i, v_j\rangle$, hvor $i = j$ bidrager til summen.

Eksempel 1.6. a) I $\mathbb{C}^n$ er $((1,0,\dots,0),(0,1,0,\dots,0),\dots,(0,\dots,0,1))$ et ortonormalsæt. Dermed er det lineært uafhængigt (hvilket naturligvis også let kan indses på anden vis) og en basis for $\mathbb{C}^n$, da det klart udspænder $\mathbb{C}^n$.

b) Sættet $(\varepsilon_1, \varepsilon_2, \dots)$ indført i Eksempel 1.1 c) er et ortonormalsæt i $\ell_0(\mathbb{N})$. Overvej dette!

c) Sættet $(1, \cos x, \sin x, \cos 2x, \sin 2x, \dots)$ er et ortogonalsæt i $C([-\pi, \pi])$. Dette fremgår af udregningerne p.530-531 i AEM (jvf. Opgave 1.15), hvoraf det også femgår, at

$$\|\cos nx\| = \|\sin nx\| = \sqrt{\pi} \quad \text{for} \quad n \in \mathbb{N} .$$

Sættet er derfor ikke et ortonormalsæt. Ved normering (dvs. division med normen) af vektorerne i sættet fås ortonormalsættet

$$\left(\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos x, \frac{1}{\sqrt{\pi}}\sin x, \frac{1}{\sqrt{\pi}}\cos 2x, \frac{1}{\sqrt{\pi}}\sin 2x, \dots\right) . \tag{1.25}$$

Et andet ortonormalsæt i $C([-\pi,\pi])$ er

$$(\ldots, \frac{1}{\sqrt{2\pi}}e^{-i2x}, \frac{1}{\sqrt{2\pi}}e^{-ix}, \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{2\pi}}e^{ix}, \frac{1}{\sqrt{2\pi}}e^{2ix}, \ldots)\ , \tag{1.26}$$

hvor vi har nummereret vektorerne med $n \in \mathbb{Z}$ i stedet for $n \in \mathbb{N}$, således at den n'te vektor er funktionen

$$e_n(x) = \frac{1}{\sqrt{2\pi}}e^{inx}\ .$$

At disse vektorer er parvis ortogonale og har norm 1, fremgår af følgende udregning

$$\begin{aligned}
\langle e_n, e_m\rangle &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\overline{e^{inx}}e^{imx}dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-inx}e^{imx}dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}e^{i(m-n)x}dx \\
&= \begin{cases} \frac{1}{2\pi}\left[\frac{1}{i(m-n)}e^{i(m-n)x}\right]_{-\pi}^{\pi} = 0 & \text{for} \quad n \neq m \\ \frac{1}{2\pi}[x]_{-\pi}^{\pi} = 1 & \text{for} \quad n = m\ . \end{cases}
\end{aligned}$$

Det bemærkes, at dette ortonormalsæt optræder i Afsnit 10.5 i AEM.

d) Legendre polynomierne $P_0, P_1, P_2, \ldots$ udgør et ortogonalsæt i $C([-1,1])$, jvf. Eksempel 4 p.237 i AEM. Det er ikke et ortonormalsæt, idet

$$\|P_n\|^2 = \int_{-1}^{1}P_n(x)^2dx = \frac{2}{2n+1}$$

ifølge formel (8) p.242 i AEM.

Projektionssætningen. Vi afslutter dette kapitel med at diskutere den såkaldte *projektionssætning*, der i LA er anført som Theorem 4.19, og som giver en formel for projektionen af en vektor på et endeligdimensionalt underrum U af et indre produkt rum V. Inden vi formulerer sætningen, genkalder vi, at det *ortogonale komplement* $A^\perp$ til en delmængde $A \subseteq V$ består af de vektorer, der er vinkelrette på alle vektorer i A, dvs.

$$A^\perp = \{v \in V \mid \langle v, w\rangle = 0 \text{ for alle } w \in A\}\ .$$

Sætning 1.7. *Lad $(e_1, \ldots, e_n)$ være et endeligt ortonormalsæt i et indre produkt rum V, og lad*

$$U = \text{span}\{e_1, \ldots, e_n\}\ .$$

For en vilkårlig vektor $v \in V$ findes da netop een vektor $w \in U$, således at afstanden $\|v - w\|$ er mindst mulig (blandt alle afstande fra v til vektorer i U). Vektoren w kaldes projektionen af v på U og er givet ved

$$w = \langle e_1, v\rangle e_1 + \cdots + \langle e_n, v\rangle e_n\ , \tag{1.27}$$

og der gælder

$$v - w \in U^\perp\ .$$

Defineres nemlig w ved (1.27) og tages indre produkt med e_i på begge sider af denne ligning, giver venstresiden $\langle e_i, w\rangle$, mens højresiden ved en udregning som i begyndelsen af dette afsnit giver $\langle e_i, v\rangle$. Altså er $\langle e_i, v\rangle = \langle e_i, w\rangle$, d.v.s. $\langle e_i, v-w\rangle = 0$. Dette viser, at $v - w$ er vinkelret på alle vektorerne $e_1, \dots, e_n$ og dermed på alle linearkombinationer af dem ifølge (1.8). Derfor er $v - w \in U^\perp$ som påstået.

Vi mangler at indse, at w er den vektor i U, som har kortest afstand til v. Lad hertil w' være en vilkårlig vektor i U. Da er

$$\begin{aligned}\|v - w'\|^2 &= \|(v - w) + (w - w')\|^2\\ &= \|v - w\|^2 + \|w - w'\|^2 ,\end{aligned}$$

hvor vi har brugt Pythagoras, idet $v - w \in U^\perp$ og $w - w' \in U$. Men $\|w - w'\|^2$ er jo mindst mulig, nemlig 0, netop når $w' = w$, hvilket viser det ønskede.

Bemærk, at Pythagoras medfører dels

$$\begin{aligned}\|w\|^2 &= \|\langle e_1, v\rangle e_1\|^2 + \cdots + \|\langle e_n, v\rangle e_n\|^2\\ &= |\langle e_1, v\rangle|^2 + \cdots + |\langle e_n, v\rangle|^2\end{aligned}$$

og dels

$$\begin{aligned}\|v\|^2 &= \|(v - w) + w\|^2\\ &= \|v - w\|^2 + \|w\|^2\\ &\geq \|w\|^2 .\end{aligned}$$

Heraf følger den vigtige *Bessels ulighed*:

$$\sum_{i=1}^{n} |\langle e_i, v\rangle|^2 \leq \|v\|^2 \tag{1.28}$$

for enhver vektor $v \in V$ og ethvert endeligt ortonormalsæt $(e_1, \dots, e_n)$ i V.

Eksempel 1.8. Lad os bestemme projektionen af funktionen x i $C([-\pi, \pi])$ på underrummet $U = \operatorname{span}\{1, \cos x, \sin x\}$.

Sættes $e_1(x) = \frac{1}{\sqrt{2\pi}}$, $e_2(x) = \frac{1}{\sqrt{\pi}} \cos x$, $e_3(x) = \frac{1}{\sqrt{\pi}} \sin x$, er (e_1, e_2, e_3) et ortonormalsæt i $C([-\pi, \pi])$ og $U = \operatorname{span}\{e_1, e_2, e_3\}$, således at den søgte projektion er funktionen

$$\begin{aligned}g(y) &= \langle \frac{1}{\sqrt{2\pi}}, x\rangle \frac{1}{\sqrt{2\pi}} + \langle \frac{1}{\sqrt{\pi}} \cos x, x\rangle \frac{1}{\sqrt{\pi}} \cos y + \langle \frac{1}{\sqrt{\pi}} \sin x, x\rangle \frac{1}{\sqrt{\pi}} \sin y\\ &= \frac{1}{2\pi}\langle 1, x\rangle + \frac{1}{\pi}\langle \cos x, x\rangle \cos y + \frac{1}{\pi}\langle \sin x, x\rangle \sin y\\ &= 2 \sin y ,\end{aligned}$$

idet man ved udregning finder $\langle 1, x\rangle = \langle \cos x, x\rangle = 0$ og $\langle \sin x, x\rangle = 2\pi$.

Opgaver

Opgave 1.1 Eftervis, at mængden $\ell_0(\mathbb{N})$ bestående af følger med kun endelig mange elementer forskellige fra nul er et underrum af vektorrummet $\ell(\mathbb{N})$ af alle komplekse talfølger. (Benyt Theorem 1.11 i LA.)

Opgave 1.2 Begrund, at der for en delmængde A af et vektorrum V gælder, at $\mathrm{span}A$ er det mindste underrum af V, der indeholder A. (Overbevis dig selv om, at det, der kræves, er dels, at $\mathrm{span}A$ indeholder A, og dels, at ethvert underrum U, der indeholder A, også må indeholde enhver linearkombination af vektorer fra A.)

Opgave 1.3 Vis, at $\mathrm{span}A$ ikke ændres, hvis A udvides med en mængde B, der er indeholdt i $\mathrm{span}A$, altså at $\mathrm{span}(A \cup B) = \mathrm{span}A$, hvis $B \subseteq \mathrm{span}A$.

Opgave 1.4 Betragt underrummet $U = \mathrm{span}\{1, \cos x, \cos 2x\}$ af $C([-\pi, \pi])$.
Vis, at $\cos^2 x$ tilhører U, men at $\cos^3 x$ ikke tilhører U.

Opgave 1.5 Udled af de definerende egenskaber ved det indre produkt de to relationer (1.7) og (1.8):

$$\langle v, \lambda_1 w_1 + \cdots + \lambda_n w_n \rangle = \lambda_1 \langle v, w_1 \rangle + \cdots + \lambda_n \langle v, w_n \rangle$$

$$\langle \lambda_1 w_1 + \cdots + \lambda_n w_n, v \rangle = \overline{\lambda_1} \langle w_1, v \rangle + \cdots + \overline{\lambda_n} \langle w_n, v \rangle$$

Opgave 1.6 Verificer, at der ved

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = \overline{x_1} y_1 + \cdots + \overline{x_n} y_n$$

defineres et indre produkt på $\mathbb{C}^n$, altså at kravene *i)-iv)* til et indre produkt er opfyldt.

Opgave 1.7 Verificer, at hvis ρ er en positiv kontinuert funktion på $[a, b]$, da defineres der ved

$$\langle f, g \rangle_\rho = \int_a^b \overline{f(x)} g(x) \rho(x) dx$$

et indre produkt på $C([a, b])$.

Opgave 1.8 Verificer, at der ved

$$\langle (x_1, x_2, \ldots), (y_1, y_2, \ldots) \rangle = \sum_{n=1}^{\infty} \overline{x_n} y_n$$

defineres et indre produkt på $\ell_0(\mathbb{N})$.

Opgave 1.9 Lad V være et indre produkt rum. Brug regnereglerne for det indre produkt til at udlede *parallelogramidentiteten*:

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

og forklar, hvad denne identitet har med parallelogrammer at gøre.

Udled endvidere, forudsat V er et komplekst vektorrum, *polariseringsidentiteten*:

$$\langle v, w\rangle = \frac{1}{4}(\|v+w\|^2 - \|v-w\|^2 - i\|v+iw\|^2 + i\|v-iw\|^2)\ .$$

Opgave 1.10 Brug Cauchy-Schwarz' ulighed til at slutte, at

$$\left(\int_0^\pi x\sin x dx\right)^2 \le \int_0^\pi x^2 dx \int_0^\pi \sin^2 x dx\ .$$

Udregn begge sider af uligheden og konstater, at den er opfyldt.

Opgave 1.11 Eftervis de tre fundamentale egenskaber (1.21), (1.22) og (1.23) ved normen $\|\cdot\|$ på et indre produkt rum V, jvf. Theorem 4.9 i LA.

Opgave 1.12 Brug trekantsuligheden (1.23) til at udlede

$$|\ \|v\| - \|w\|\ | \le \|v-w\|\ ,$$

altså at afstanden (i $\mathbb{R}$) mellem normerne af v og w er mindre end eller lig med afstanden mellem v og w (i V). (Man siger, at normafbildningen $v \to \|v\|$ fra V ind i $\mathbb{R}$ er afstandsformindskende.)

Opgave 1.13 Brug uligheden fra Opgave 1.12 til at slutte, at der for en konvergent følge $(v_1, v_2, \dots)$ i et indre produkt rum gælder, at

$$\|\lim_{n\to\infty} v_n\| = \lim_{n\to\infty} \|v_n\|\ .$$

Opgave 1.14 Brug Cauchy-Schwarz' ulighed til at slutte, at der for en konvergent følge $(v_1, v_2, \dots)$ i et indre produkt rum gælder, at

$$\langle \lim_{n\to\infty} v_n, w\rangle = \lim_{n\to\infty} \langle v_n, w\rangle$$

og

$$\langle w, \lim_{n\to\infty} v_n\rangle = \lim_{n\to\infty} \langle w, v_n\rangle$$

for en vilkårlig vektor $w \in V$.

Opgave 1.15 Eftervis, at $(1, \cos x, \sin x, \cos 2x, \sin 2x, \dots)$ er et ortogonalsæt i $C([-\pi,\pi])$, enten ved direkte udregning eller ved at bruge, at sættet (1.26) er et ortonormalsæt sammen med Euler's formel $\exp ix = \cos x + i\sin x$.

Opgave 1.16 Vi ser på polynomierne $1, x, x^2, x^3, \dots$ i $C([-1,1])$ med sædvanligt indre produkt.

Bestem projektionen $Q_{n-1}(x)$ af x^n på $\operatorname{span}\{1, x, \dots, x^{n-1}\}$ for $n = 1, 2, 3$.

Sammenlign polynomierne $x^n - Q_{n-1}(x)$ for disse værdier af n med Legendre polynomierne p.208 i AEM. - Er du overrasket?

Opgave 1.17 Godtgør, at $(\sin x, \sin 2x, \sin 3x)$ er et ortogonalsæt i $C([0,\pi])$ med sædvanligt indre produkt.

Bestem $a_1, a_2, a_3 \in \mathbb{C}$, således at

$$\int_0^\pi |\cos x - \sum_{n=1}^{3} a_n \sin nx|^2 dx$$

antager den mindst mulige værdi.

Opgave 1.18 Lad vektorerne $\varepsilon_1, \varepsilon_2, \dots$ i $\ell_0(\mathbb{N})$ være givet som i Eksempel 1.1 c).

Eftervis, at sættet $(\varepsilon_1 + \varepsilon_2, \varepsilon_1 - \varepsilon_2, \varepsilon_3 + \varepsilon_4, \varepsilon_3 - \varepsilon_4, \dots)$ er et ortogonalsæt i $\ell_0(\mathbb{N})$ og bestem det tilsvarende ortonormalsæt ved normering.

Opgave 1.19 Eftervis, at funktionsfølgen $(f_1, f_2, \dots)$ i $C([0,1])$, hvor

$$f_n(x) = \frac{nx^2}{1+nx^2} \ ,$$

konvergerer i middel imod den konstante funktion 1.

Konvergerer den også punktvis imod funktionen 1?

Opgave 1.20 Find en funktionsfølge $(f_1, f_2, \dots)$ i $C([0,1])$, som konvergerer i middel imod nulfunktionen, og således at $f_n(1) \to \infty$ for $n \to \infty$.
Vink. $f_n(x) = a_n x^n$, hvor a_n er valgt passende, kan bruges.

2 Hilbert rum

2.1 Eksempler på Hilbert rum

Vi skal nu først forsøge at begrunde, at de indre produkt rum af funktioner eller følger, som blev indført i Kapitel 1, ikke er omfattende nok til vores formål. Husk på, med udgangspunkt i vores undersøgelser af Fourier rækker og partielle differentialligninger, at vi ønsker at kunne danne "uendelige linearkombinationer" eksempelvis af vektorer i ortonormalsættene omtalt i Eksempel 1.6 c) og d). Problemet er, at dette hyppigt vil give funktioner, der ikke er kontinuerte, dvs. som ikke ligger i vektorrummet. Tag for eksempel den "uendelige linearkombination"

$$\frac{4}{\pi}(\sin x + \tfrac{1}{3}\sin 3x + \tfrac{1}{5}\sin 5x + \cdots)$$

i $C([-\pi, \pi])$. Denne repræsenterer, som vi så i AEM Eksempel 1 p. 241–242 eller Eksempel 1.1 p. 532–534, funktionen f, som er 1 på $[0, \pi]$ og -1 på $[-\pi, 0[$, og som derfor ikke er kontinuert i 0.

Noget tilsvarende gør sig gældende for $\ell_0(\mathbb{N})$: Går vi formelt tilværks, ser vi, at den uendelige linearkombination

$$\lambda_1\varepsilon_1 + \lambda_2\varepsilon_2 + \lambda_3\varepsilon_3 + \cdots$$

er lig med følgen $(\lambda_1, \lambda_2, \lambda_3, \dots)$, jvf. Eksempel 1.1 c). Men denne tilhører kun $\ell_0(\mathbb{N})$, hvis højst endelig mange af koefficienterne λ_n er $\neq 0$, dvs. hvis linearkombinationen er endelig.

Vi kan undgå disse problemer ved at gøre de omtalte vektorrum større, uden at ændre på det indre produkt. Vi gennemfører dette for $\ell_0(\mathbb{N})$ først, og derefter for $C(I)$.

Vektorrummet $\ell_2(\mathbb{N})$. Bemærk til at begynde med, at normen på $\ell_0(\mathbb{N})$, som jo er givet ved

$$\|(x_1, x_2, \dots)\|^2 = \sum_{n=1}^{\infty} |x_n|^2 \tag{2.1}$$

er veldefineret for alle følger for hvilke højresiden er endelig. F.eks. tilhører følgen $(1, \frac{1}{2}, \frac{1}{3}, \dots)$ åbenbart ikke $\ell_0(\mathbb{N})$, men

$$\|(1, \tfrac{1}{2}, \tfrac{1}{3}, \dots)\|^2 = \sum_{n=1}^{\infty} \frac{1}{n^2}$$

er som bekendt endelig (faktisk lig med $\pi^2/6$). Vi kalder talfølger, for hvilke højresiden af (2.1) er endelig for *kvadratisk summable* følger og betegner mængden af disse med $\ell_2(\mathbb{N})$, altså

$$\ell_2(\mathbb{N}) = \{(x_1, x_2, \dots) \;\Big|\; \sum_{n=1}^{\infty} |x_n|^2 < \infty\}\,.$$

Dette er vores kandidat til udvidelsen af $\ell_0(\mathbb{N})$. I en vis forstand er dette den størst mulige udvidelse med den givne norm, idet vi simpelhen har medtaget alle følger med endelig norm. Spørgsmålet er nu dels, hvorvidt $\ell_2(\mathbb{N})$ faktisk er et vektorrum, og dels om det indre produkt givet ved (1.14) er veldefineret på $\ell_2(\mathbb{N})$, altså hvorvidt højresiden af (1.14) er konvergent for vilkårlige følger i $\ell_2(\mathbb{N})$. Svaret på begge spørgsmål er heldigvis ja, og det er ikke vanskeligt at indse, jvf. Opgave 2.1 og 2.2.

Vektorrummet $L_2([a,b])$. Situationen vedrørende $C([a,b])$ er tilsvarende. Vi foretager den maksimale udvidelse af $C([a,b])$ til mængden $L_2([a,b])$ bestående af alle *kvadratisk integrable* funktioner på $[a,b]$, d.v.s. funktioner hvis normkvadrat (1.12) er endeligt, altså

$$L_2([a,b]) = \{f : [a,b] \to \mathbb{C} \mid \int_a^b |f(x)|^2 dx < \infty\}\,. \tag{2.2}$$

Dette rum indeholder eksempelvis alle stykkevis kontinuerte funktioner på $[a,b]$. Husk, at en funktion f på $[a,b]$ er stykkevis kontinuert, hvis den er kontinuert i alle punkter, pånær muligvis i endelig mange punkter $a = t_0 < t_1 < t_2 < \cdots < t_N = b$, hvori dens grænseværdier fra venstre og højre kræves at eksistere. Alternativt betyder dette, at f på hvert af intervallerne $]t_{i-1}, t_i[$ kan udvides til en kontinuert funktion f_i på $[t_{i-1}, t_i]$, og integralet af f er da pr. definition givet ved

$$\int_a^b f(x)dx = \sum_{i=1}^N \int_{t_{i-1}}^{t_i} f_i(x)dx\,.$$

Men $L_2([a,b])$ indeholder flere funktioner. F.eks. er funktionen

$$f(x) = x^{-\alpha},\ \ 0 < x \le 1\ ,$$

hvor $\alpha > 0$, ikke stykkevis kontinuert på $[0,1]$, uanset hvilken værdi $f(0)$ tillægges. Ikke destomindre har vi, at

$$\int_0^1 |f(x)|^2 dx = \int_0^1 x^{-2\alpha} dx = \lim_{\epsilon\to 0}\int_\epsilon^1 x^{-2\alpha}dx = \lim_{\epsilon\to 0}\frac{1-\epsilon^{1-2\alpha}}{1-2\alpha} = \begin{cases} \frac{1}{1-2\alpha}\ , & \alpha < \frac{1}{2} \\ +\infty\ , & \alpha > \frac{1}{2}\ . \end{cases}$$

Altså er $f \in L_2([0,1])$ for $\alpha < \frac{1}{2}$.

Vi skal ikke her forsøge nøjere at indkredse arten af de funktioner, der tilhører $L_2([a,b])$. En udførlig diskussion ville kræve indførelse af det såkaldte Lebesgue integral, som er underforstået i (2.2), og som tillader integration af en større klasse af funktioner end det sædvanlige Riemann integral. I praksis får vi dog kun brug for at integrere funktioner, for hvilke en simpel udvidelse af Riemann integralet, som i de lige nævnte eksempler, kan anvendes.

Man viser nu på lignende vis som for $\ell_2(\mathbb{N})$, at $L_2([a,b])$ er et vektorrum, og at integralet, der indgår i definitionen af det indre produkt i (1.11), er endeligt for $f, g \in L_2([a,b])$, jvf. Opgave 2.3. At eftervise de fire krav $i)-iv)$ er ikke anderledes end for $C([a,b])$, pånær et enkelt punkt, nemlig det skarpe ulighedstegn i $i)$. Tag f.eks. en funktion f som er nul overalt i $[a,b]$ undtagen i et enkelt punkt (eller endelig mange punkter). Denne funktion er stykkevis kontinuert og dens norm er øjensynlig 0. Der findes altså funktioner i $L_2([a,b])$ forskellige fra nulfunktionen, som har norm

0. Der findes en standard måde at takle dette problem på, nemlig ved at identificere funktioner, hvis differens har norm 0. Dette betyder, at man i stedet for at definere $L_2([a,b])$ som et rum af funktioner opfatter det som et rum af klasser (eller mængder) af funktioner, således at to funktioner f og g tilhører samme klasse netop hvis

$$\int_a^b |f(x) - g(x)|^2 dx = 0\,. \tag{2.3}$$

Herved får alle funktioner i en klasse samme norm (jvf. Opgave 2.4), som derfor kan defineres som klassens norm. Tilsvarende vil det indre produkt (1.11) af to funktioner ikke ændres, hvis de byttes ud med andre funktioner fra deres respektive klasser (jvf. Opgave 2.4), hvilket derfor kan defineres som det indre produkt af klasserne. Denne ændring af synspunkt har ikke nævneværdig praktisk betydning. Det er fortsat nyttigt at tænke på elementerne i $L_2([a,b])$ som funktioner, men blot være opmærksom på, at en ligning i vektorrummet $L_2([a,b])$, hvor der på venstresiden står en funktion f og på højresiden en funktion g, blot betyder at (2.3) er opfyldt, og ikke at funktionerne stemmer overens i alle punkter af $[a,b]$. Man siger i stedet at f er lig med g *næsten overalt* i $[a,b]$, eller at $f(x) = g(x)$ *for næsten alle* $x \in [a,b]$.

Vektorrummet $L_2(I,\rho)$. Vi kan tilsvarende indføre vektorrummet $L_2(I)$, hvor I er et vilkårligt (ikke nødvendigvis endeligt eller lukket) interval, og specielt $L_2(\mathbb{R})$. Mere generelt kan vi for en positiv vægtfunktion ρ på I indføre vektorrummet

$$L_2(I,\rho) = \{f : I \to \mathbb{C} \mid \int_I |f(x)|^2 \rho(x) dx < \infty\} \tag{2.4}$$

med indre produkt $\langle \cdot\,,\,\cdot \rangle_\rho$ givet ved (1.13).

$L_2(I)$ og kvantemekanik. Rummene $L_2(I)$ er af særlig vigtighed i den kvantemekaniske formalisme. Et af grundpostulaterne i denne formalisme er nemlig, at den fysiske tilstand af en partikel, hvis bevægelsesfrihed er begrænset til intervallet I, er givet ved en *bølgefunktion* ψ, som er en funktion i $L_2(I)$, og at $|\psi(x)|^2$ da er sandsynlighedstætheden for partiklens position. Dette betyder, at sandsynligheden for, at partiklen befinder sig i intervallet $[x, x+\Delta x] \subseteq I$, er

$$\int_x^{x+\Delta x} |\psi(x)|^2 dx\,. \tag{2.5}$$

Bemærk, at den ovenfor diskuterede identifikation af funktioner i samme klasse er ganske naturlig ud fra et kvantemekanisk synspunkt, eftersom to funktioner ψ og ϕ, som opfylder (2.3), regnes for at beskrive den samme kvantemekaniske tilstand. Specielt er sandsynlighederne givet ved (2.5) de samme for ψ og ϕ (jvf. Opgave 2.4).

Et andet grundelement i den kvantemekaniske formalisme med relation til det indre produkt er måleprocessen. En partikel, der vides at befinde sig i en tilstand ψ, vil normalt efter en måleproces befinde sig i en anden tilstand, eller anderledes udtrykt: Måleprocessen forstyrrer det fysiske system. Mere præcist er det sådan, at hvis man foretager en måling, der som resultat enten bekræfter eller afkræfter, at partiklen (efter målingen) befinder sig i tilstanden ϕ, da er sandsynligheden for et bekræftende svar givet ved

$$|\langle \psi, \phi \rangle|^2\,.$$

Specielt gælder, at tilstanden efter målingen ikke kan være vinkelret på ψ. Man kalder tit $\langle \psi, \phi \rangle$ for *overlappet* mellem tilstandene ψ og ϕ.

Da den totale sandsynlighed for at finde partiklen i I er 1, er kvantemekaniske tilstande repræsenteret af enhedsvektorer i $L_2(I)$, dvs. de har norm lig med 1. Ikke destomindre siger man tit, at en partikel befinder sig i en tilstand ψ, som ikke er en enhedsvektor. Hermed menes i såfald, at den befinder sig i den tilsvarende *normerede tilstand* $\psi/\|\psi\|$. Med denne konvention kan man danne linearkombinationer af kvantemekaniske tilstande. Denne egenskab ved tilstandsrummet for et kvantemekanisk system, som ikke har nogen analogi for klassiske mekaniske systemer, kaldes *superpositionsprincippet*. Dette er altså en direkte konsekvens af vektorrumsstrukturen af $L_2(I)$ og giver anledning til de for kvantemekanik så karakteristiske interferensfænomener.

Hilbert rum. Vi afslutter dette afsnit med spørgsmålet om, hvilken karakteristisk egenskab, der adskiller de udvidede indre produkt rum fra dem, der var vores udgangspunkt, ud over de lidt løse betragtninger vedrørende "uendelige linearkombinationer", vi har været inde på. Det store dyr i åbenbaringen her hedder *fuldstændighed*. Denne egenskab er en smart måde at udtrykke maksimaliteten, der indgik i konstruktionen af $\ell_2(\mathbb{N})$ og $L_2(I)$ – altså, at vi medtog *alle* vektorer, for hvilke normen (og det indre produkt) var definerede. Man siger også, at $\ell_2(\mathbb{N})$ og $L_2([a,b])$ er fremkommet ved fuldstændiggørelse af $\ell_0(\mathbb{N})$, henholdsvis $C([a,b])$.

For at kunne formulere denne egenskab har vi brug for først at definere begrebet *Cauchy følge* (eller *fundamentalfølge*): Vi siger, at en følge (v_n) i et indre produkt rum er en Cauchy følge, hvis der for hvert $\varepsilon > 0$ findes et $N \in \mathbb{N}$, således at

$$\|v_n - v_m\| < \varepsilon \text{ for } n, m > N\,,$$

dvs. afstanden mellem to af følgens elementer v_n og v_m kan gøres vilkårlig lille ved at vælge n og m store nok.

Det er klart, at enhver konvergent følge er en Cauchy følge, fordi hvis $v_n \to v$ for $n \to \infty$, da er v_n og v_m begge tæt på v og derfor tæt på hinanden, når n og m er store nok. Eller mere præcist: For et vilkårligt $\varepsilon > 0$ findes $N \in \mathbb{N}$, så $\|v_n - v\| < \frac{\varepsilon}{2}$ for $n > N$, og derfor

$$\|v_n - v_m\| = \|(v_n - v) + (v - v_m)\| \leq \|v_n - v\| + \|v - v_m\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for $n, m > N$.

Fuldstændighed er kravet om, at det omvendte også er sandt:

Definition 2.1. Et indre produkt rum H er fuldstændigt, hvis enhver Cauchy følge er konvergent. I så fald siges H at være et Hilbert rum.

For de reelle tal $\mathbb{R}$, betragtet som indre produkt rum med sædvanligt indre produkt, er denne egenskab ækvivalent med supremumsegenskaben, omtalt i Matematik 1. Dette diskuteres mere udførligt i appendikset, hvori fuldstændighedsbegrebet søges yderligere belyst. Det følger heraf, at $\mathbb{R}^n$ og $\mathbb{C}^n$ er Hilbert rum og mere generelt, at *ethvert endeligdimensionalt indre produkt rum er et Hilbert rum* (se Opgave 2.5). For endeligdimensionale indre produkt rum er kravet om fuldstændighed derfor overflødigt. Helt anderledes forholder det sig med uendeligdimensionale vektorrum. F.eks. er hverken $\ell_0(\mathbb{N})$ eller $C([a,b])$ fuldstændige. Derimod har vi følgende

Sætning 2.2. *Der gælder, at $\ell_2(\mathbb{N})$ og $L_2([a,b])$ og mere generelt $L_2(I,\rho)$ er Hilbert rum.*

Dette resultat er helt afgørende for disse vektorrums anvendelighed i analysen. Det er ikke svært at eftervise for $\ell_2(\mathbb{N})$, men mere besværligt for $L_2(I,\rho)$, idet det kræver kendskab til det før omtalte udvidede integralbegreb, hvorfor vi afholder os fra at gennemføre det her.

Det bør bemærkes, at tilstandsrummene for kvantemekaniske systemer generelt er Hilbert rum. F.eks. består tilstandsrummet for flerpartikelsystemer uden spin af kvadratisk integrable funktioner af flere variable, nemlig koordinaterne for de enkelte partikler. Vi skal dog ikke omtale disse yderligere i disse noter.

2.2 Ortonormaludvikling

Vi skal i dette afsnit generalisere de centrale resultater vedrørende endeligdimensionale indre produkt rum til vilkårlige Hilbert rum, herunder specielt vise projektionssætningen for uendelige ortonormalsæt.

Konvergens af rækker i Hilbert rum. Vi begynder med at karakterisere præcist hvilke "uendelige linearkombinationer" af vektorerne i et ortonormalsæt, der giver mening. Hertil er det nødvendigt først at definere konvergens af rækker i et indre produkt rum. Dette gøres på samme måde som for talrækker:

Rækken $\sum\limits_{n=1}^{\infty} v_n$, hvis led v_n er vektorer i et indre produkt rum V, siges at være konvergent med sum $v \in V$, hvis

$$\lim_{N\to\infty} \sum_{n=1}^{N} v_n = v\,,$$

altså hvis *afsnitsfølgen* $(s_1, s_2, \dots)$, hvor

$$s_N = \sum_{n=1}^{N} v_n\,,$$

er konvergent i V med grænseværdi v.

Eksempel 2.3. a) Lad os se på rækken $\sum_{n=1}^{\infty} \frac{1}{n}\varepsilon_n$ i $\ell_2(\mathbb{N})$, hvor ortonormalsættet $(\varepsilon_1, \varepsilon_2, \dots)$ er som i Eksempel 1.1 c). Her er

$$s_N = \sum_{n=1}^{N} \frac{1}{n}\varepsilon_n = \left(1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{N}, 0, 0, \dots\right).$$

Vi har tidligere bemærket, at følgen $x = (1, \frac{1}{2}, \frac{1}{3}, \dots)$ tilhører $\ell_2(\mathbb{N})$. Da

$$x - s_N = \left(0, 0, \dots, 0, \frac{1}{N+1}, \frac{1}{N+2}, \dots\right)$$

fås

$$\|x - s_N\|^2 = \sum_{n=N+1}^{\infty} \frac{1}{n^2} \to 0 \text{ for } N \to \infty\,,$$

fordi $\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ er konvergent, så "halen" $\sum\limits_{n=N+1}^{\infty} \frac{1}{n^2}$ konvergerer imod 0 for $N \to \infty$. Altså har vi indset, at

$$\sum_{n=1}^{\infty} \frac{1}{n}\varepsilon_n = (1, \tfrac{1}{2}, \tfrac{1}{3}, \dots)\,.$$

b) Lad os mere generelt undersøge rækken $\sum\limits_{n=1}^{\infty} \lambda_n \varepsilon_n$. Her er

$$s_N = \sum_{n=1}^{N} \lambda_n \varepsilon_n = (\lambda_1, \lambda_2, \dots, \lambda_N, 0, 0, \dots)\,.$$

Følgen $x = (\lambda_1, \lambda_2, \lambda_3, \dots)$ tilhører $\ell_2(\mathbb{N})$ pr. definition netop hvis $\sum\limits_{n=1}^{\infty} |\lambda_n|^2 < \infty$. I dette tilfælde fås

$$x - s_N = (0, 0, \dots, 0, \lambda_{N+1}, \lambda_{N+2}, \dots)$$

og derfor

$$\|x - s_N\|^2 = \sum_{n=N+1}^{\infty} |\lambda_n|^2 \to 0 \text{ for } N \to \infty\,.$$

Altså er

$$\sum_{n=1}^{\infty} \lambda_n \varepsilon_n = (\lambda_1, \lambda_2, \dots)\,,$$

når højresiden tilhører $\ell_2(\mathbb{N})$. Ellers er rækken divergent. (Overvej dette! Det følger også af Lemma 2.4 nedenfor.)

Ved brug af følgende hjælperesultat kan vi generalisere overvejelserne i det foregående eksempel til vilkårlige ortonormalsystemer i et vilkårligt Hilbert rum.

Lemma 2.4. *Lad $(e_1, e_2, \dots)$ være et ortonormalsæt i et Hilbert rum H. Da er rækken $\sum\limits_{n=1}^{\infty} \lambda_n e_n$ konvergent i H hvis og kun hvis $\sum\limits_{n=1}^{\infty} |\lambda_n|^2$ er konvergent i $\mathbb{R}$, altså netop hvis talfølgen $(\lambda_1, \lambda_2, \dots)$ tilhører $\ell_2(\mathbb{N})$. I bekræftende fald gælder*

$$\Big\| \sum_{n=1}^{\infty} \lambda_n e_n \Big\|^2 = \sum_{n=1}^{\infty} |\lambda_n|^2\,. \tag{2.6}$$

Dette resultat afhænger på afgørende vis af fuldstændigheden af H, ifølge hvilken rækken $\sum\limits_{n=1}^{\infty} \lambda_n e_n$ er konvergent netop hvis afsnitsfølgen $(s_1, s_2, \dots)$ er en Cauchy følge. Men da

$$s_N - s_M = \sum_{n=M+1}^{N} \lambda_n e_n\,,$$

for $N > M$, følger af Pythagoras, at

$$\|s_N - s_M\|^2 = \sum_{n=M+1}^{N} |\lambda_n|^2 \,.$$

Kaldes nu afsnitsfølgen for talrækken $\sum_{n=1}^{\infty} |\lambda_n|^2$ for $(\gamma_1, \gamma_2, \dots)$, viser dette at

$$\|s_N - s_M\|^2 = \gamma_N - \gamma_M = |\gamma_N - \gamma_M| \,.$$

Heraf ses, at $(s_1, s_2, s_3, \dots)$ er en Cauchy følge (i H) netop hvis $(\gamma_1, \gamma_2, \gamma_3, \dots)$ er en Cauchy følge (i $\mathbb{R}$). Men for fuldstændige rum er Cauchy følge og konvergent følge det samme, og da både H og $\mathbb{R}$ er fuldstændige, følger den første påstand i lemmaet heraf.

Endelig følger ligningen (2.6) af, at

$$\Big\| \sum_{n=1}^{N} \lambda_n e_n \Big\|^2 = \sum_{n=1}^{N} |\lambda_n|^2$$

iflg. Pythagoras, og derfor

$$\sum_{n=1}^{\infty} |\lambda_n|^2 = \lim_{N\to\infty} \sum_{n=1}^{N} |\lambda_n|^2 = \lim_{N\to\infty} \Big\| \sum_{n=1}^{N} \lambda_n e_n \Big\|^2 = \Big\| \sum_{n=1}^{\infty} \lambda_n e_n \Big\|^2 ,$$

hvor det sidste lighedstegn følger af Opgave 1.13.

Projektionssætningen. Udstyret med Lemma 2.4 kan vi nu definere det *afsluttede underrum udspændt af et ortonomalsæt* $(e_1, e_2, \dots)$ ved

$$\overline{\operatorname{span}}\{e_1, e_2, \dots\} = \left\{ \sum_{n=1}^{\infty} \lambda_n e_n \;\middle|\; \sum_{n=1}^{\infty} |\lambda_n|^2 < \infty \right\} .$$

Dvs. span $\{e_1, e_2, \dots\}$ og $\overline{\operatorname{span}}\{e_1, e_2, e_3, \dots\}$ afviger fra hinanden ved, at der i det første kun indgår sædvanlige (endelige) linearkombinationer, mens der i det andet indgår alle tilladte, dvs. konvergente, "uendelige linearkombinationer".

Den generelle projektionssætning lyder nu således:

Sætning 2.5. *Lad* $(e_1, e_2, \dots)$ *være et ortonormalsæt i Hilbert rummet* H *og lad*

$$U = \overline{\operatorname{span}}\{e_1, e_2, \dots\} \,.$$

For en vilkårlig vektor $v \in H$ *findes da netop én vektor* $w \in U$*, således at afstanden* $\|v - w\|$ *er mindst mulig. Vektoren* w *kaldes projektionen af* v *på* U *og er givet ved*

$$w = \sum_{n=1}^{\infty} \langle e_n, v \rangle e_n \,, \tag{2.7}$$

og der gælder

$$v - w \in U^{\perp} \,. \tag{2.8}$$

Eftervisningen af dette foregår, pånær enkelte tilføjelser af teknisk art, efter samme recept som for Sætning 1.7:
Det første, vi bemærker, er, at rækken i (2.7) er konvergent. Ifølge Lemma 2.4 er dette nemlig ensbetydende med, at rækken

$$\sum_{n=1}^{\infty} |\langle e_n, v\rangle|^2$$

er konvergent. Men det er den, fordi ifølge Bessels ulighed (1.28) er $\sum\limits_{n=1}^{N} |\langle e_n, v\rangle|^2 \leq \|v\|^2$ for alle $N \in \mathbb{N}$ og derfor

$$\sum_{n=1}^{\infty} |\langle e_n, v\rangle|^2 \leq \|v\|^2 . \tag{2.9}$$

Vi kan altså definere w ved (2.7) og får da

$$\begin{aligned}\langle e_m, w\rangle &= \left\langle e_m, \sum_{n=1}^{\infty} \langle e_n, v\rangle e_n \right\rangle \\ &= \sum_{n=1}^{\infty} \langle e_m, \langle e_n, v\rangle e_n\rangle \\ &= \sum_{n=1}^{\infty} \langle e_n, v\rangle \langle e_m, e_n\rangle \\ &= \langle e_m, v\rangle ,\end{aligned}$$

hvoraf

$$\langle e_m, v - w\rangle = 0$$

for alle $m \in \mathbb{N}$. Men heraf følger, at $v - w \in U^{\perp}$, fordi der for enhver vektor $\sum_{m=1}^{\infty} \lambda_m e_m$ i U gælder, at

$$\left\langle \sum_{m=1}^{\infty} \lambda_m e_m, v - w \right\rangle = \sum_{m=1}^{\infty} \overline{\lambda_m} \langle e_m, v - w\rangle = 0 .$$

Resten følger nu lige som i beviset for Sætning 1.7.

Vi har ovenfor to gange trukket et summationstegn $\sum\limits_{n=1}^{\infty}$ uden for et indre produkt. Dette er tilladt ifølge Opgave 2.6.

Ortonormaludvikling, Parsevals ligning og Fourier rækker. For et endeligdimensionalt indre produkt rum er en ortonormalbasis en basis, som også er et ortonormalsæt. Med den generaliserede projektionssætning til rådighed kan vi indføre en brugbar generalisering af dette begreb til vilkårlige Hilbert rum.

Definition 2.6. En *ortonormalbasis* for et Hilbert rum H er et ortonormalsæt $(e_1, e_2, e_3, \dots)$, således at

$$\{e_1, e_2, e_3, \dots\}^{\perp} = \{\underline{0}\} .$$

Det følger direkte af projektionssætningen, at der for en ortonormalbasis gælder, at $v - w = 0$, dvs. $v = w$, hvor w er givet ved (2.7). Altså

$$\boxed{v = \sum_{n=1}^{\infty} \langle e_n, v \rangle e_n \, .} \tag{2.10}$$

Bemærk specielt, at det følger, at

$$\overline{\text{span}}\{e_1, e_2, e_3, \dots\} = H \; .$$

Rækken i ligningen (2.10) kaldes *ortonormaludviklingen* af v m.h.t. ortonormalbasen $(e_1, e_2, \dots)$. Af Lemma 2.4 følger da *Parsevals ligning*

$$\boxed{\|v\|^2 = \sum_{n=1}^{\infty} |\langle e_n, v \rangle|^2 \, .} \tag{2.11}$$

Eksempel 2.7. Af Eksempel 2.2 b) fremgår, at sættet $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots)$ er en ortonormalbasis for $\ell_2(\mathbb{N})$. For $x = (x_1, x_2, \dots) \in \ell_2(\mathbb{N})$ gælder, at $\langle \varepsilon_n, x \rangle = x_n$ således at (2.10) i dette tilfælde siger, at

$$(x_1, x_2, \dots) = \sum_{n=1}^{\infty} x_n \varepsilon_n \, ,$$

hvilket er identisk med den sidste ligning i Eksempel 2.3. Endvidere ses (2.11) at være identisk med definitionen af normen på $\ell_2(\mathbb{N})$.

Af betydelig større interesse end dette noget trivielle eksempel er, at de to ortonormalsæt i Eksempel 1.6 c) er ortonormalbaser for $L_2([-\pi, \pi])$. Dette faktum er allerede blevet formuleret i AEM Kapitel 10, omend noget indirekte, og udgør i en vis forstand grundlaget for Fourierrækketeorien, som det nu skal forklares.

Lad os se på sættet (1.26) eller, mere generelt,

$$\left(\dots, \tfrac{1}{\sqrt{2L}} e^{-i2\pi \frac{x}{L}}, \tfrac{1}{\sqrt{2L}} e^{-i\pi \frac{x}{L}}, \tfrac{1}{\sqrt{2L}}, \tfrac{1}{\sqrt{2L}} e^{i\pi \frac{x}{L}}, \tfrac{1}{\sqrt{2L}} e^{i2\pi \frac{x}{L}}, \dots \right) \; . \tag{2.12}$$

Dette udgør en ortonormalbasis for $L_2([-L, L])$. I henhold til Definition 2.6 betyder dette, at den eneste kvadratisk integrable funktion f på $[-L, L]$, der opfylder

$$\int_{-L}^{L} f(x) e^{i\pi n \frac{x}{L}} dx = 0$$

for alle $n \in \mathbb{Z}$, er nulfunktionen. Vi skal ikke vise dette her, men tage det for givet i det følgende. Ortonormaludviklingen (2.10) af en kvadratisk integrabel funktion f m.h.t. denne ortonormalbasis antager nu formen

$$f(x) = \sum_{-\infty}^{\infty} c_n e^{i\pi n \frac{x}{L}} \tag{2.13}$$

for næsten alle $x \in [-L, L]$, hvor

$$c_n = \frac{1}{\sqrt{2L}} \langle \frac{1}{\sqrt{2L}} e^{i\pi n \frac{x}{L}}, f \rangle = \frac{1}{2L} \int_{-L}^{L} f(x) e^{-i\pi n \frac{x}{L}} dx \; .$$

Vi ser altså, at denne ortonormaludvikling netop er den komplekse version af Fourier rækkeudviklingen omtalt i AEM Afsnit 10.5 (se især formel (9) p. 548). Bemærk dog, at der i (2.13) er tale om *konvergens i middel* , d.v.s. at

$$\int_{-L}^{L} |f(x) - \sum_{-N}^{N} c_n e^{i\pi n \frac{x}{L}}|^2 dx \to 0 \tag{2.14}$$

for $N \to \infty$, hvorimod Theorem 1 p.535 i AEM udtaler sig om *punktvis konvergens* d.v.s. konvergens af talrækken i (2.13) for hvert fastholdt x. Disse resultater vedrørende konvergens af Fourier rækker kan sammenfattes som følger.

Sætning 2.8. *For enhver kvadratisk integrabel funktion f på $[-L, L]$ er Fourier rækken konvergent i middel og dens sum er f, d.v.s.* (2.14) *gælder.*

Hvis f ydermere er stykkevis glat, er Fourier rækken konvergent i hvert punkt $x \in \mathbb{R}$, og dens sum er lig med $\frac{1}{2}(\tilde{f}(x+) + \tilde{f}(x-))$, hvor $\tilde{f}$ betegner den periodiske udvidelse af f med periode $2L$, og $\tilde{f}(x+)$ og $\tilde{f}(x-)$ er dens grænseværdier fra h.h.v. højre og venstre i x. Specielt er summen lig med $\tilde{f}(x)$, hvis $\tilde{f}$ også er kontinuert i x.

For en kvadratisk integrabel funktion f antager Parsevals ligning nu formen

$$\int_{-L}^{L} |f(x)|^2 dx = 2L \sum_{-\infty}^{\infty} |c_n|^2 \; . \tag{2.15}$$

Denne version af ligningen optræder ikke direkte i AEM, men er analog med formel (8) p. 536.

På tilsvarende vis som ovefor fås (se Opgave 2.11), at ortonormaludvikling m.h.t. ortonormalbasen (1.25) netop er Fourier rækkeudvikling som defineret i AEM Afsnit 10.3 (for $L = \pi$), og at Parsevals ligning i dette tilfælde netop er formel (8) p. 536 i AEM.

Endvidere kan Fourier-Legendre og Fourier-Bessel rækkeudviklingerne som indført i AEM pp. 242-243 opfattes som ortonormaludviklinger i henholdsvis $L_2([-1, 1])$ og i $L_2(]0, R], x)$. Specielt svarer Fourier-Legendre udvikling til ortonormaludvikling m.h.t. ortonormalbasen $(p_0, p_1, p_2, \dots)$ i $L_2([-1, 1])$, hvor

$$p_n = \sqrt{\frac{2n+1}{2}} P_n \; ,$$

og P_n som sædvanlig betegner det n'te Legendre polynomium.

Eksempel 2.9. Vi går tilbage til Eksempel 1 p.241-242 og Eksempel 1 p.532-534 i AEM og betragter rækken

$$\frac{4}{\pi} \left(\sin x + \tfrac{1}{3} \sin 3x + \tfrac{1}{5} \sin 5x + \cdots \right) \; . \tag{2.16}$$

Da $\left(\frac{1}{\sqrt{\pi}}\sin x, \frac{1}{\sqrt{\pi}}\sin 2x, \dots\right)$ er et ortonormalsæt i $L_2([-\pi,\pi])$, jvf. Eksempel 1.6 c), og rækken

$$1^2 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots = \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2}$$

vides at være konvergent, da er (2.16) konvergent i $L_2([-\pi,\pi])$ ifølge Lemma 2.4. Vi kalder summen for f, dvs.

$$f(x) = \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{1}{2n-1} \sin(2n-1)x \tag{2.17}$$

for næsten alle $x \in [-\pi,\pi]$. Højresiden i (2.17) er faktisk ortonormaludviklingen af f m.h.t. ortonormalbasen (1.25), idet

$$\begin{aligned}
\left\langle \frac{1}{\sqrt{\pi}} \sin mx, f \right\rangle &= \left\langle \frac{1}{\sqrt{\pi}} \sin mx, \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{1}{2n-1} \sin(2n-1)x \right\rangle \\
&= \frac{4}{\sqrt{\pi}} \sum_{n=1}^{\infty} \frac{1}{2n-1} \left\langle \frac{1}{\sqrt{\pi}} \sin mx, \frac{1}{\sqrt{\pi}} \sin(2n-1)x \right\rangle \\
&= \begin{cases} \frac{4}{\sqrt{\pi}} \frac{1}{m} & \text{hvis } m \text{ er ulige} \\ 0 & \text{hvis } m \text{ er lige} \end{cases}
\end{aligned}$$

for $m = 1, 2, \dots$, og tilsvarende ses at $\left\langle \frac{1}{\sqrt{2\pi}}, f \right\rangle = \left\langle \frac{1}{\sqrt{\pi}} \cos mx, f \right\rangle = 0$.

Men højresiden i (2.17) er ifølge de to nævnte eksempler i AEM også ortonormaludviklingen af funktionen

$$g(x) = \begin{cases} -1 & \text{for } x \in [-\pi, 0[\\ 1 & \text{for } x \in [0, \pi]\ . \end{cases}$$

Vi kan derfor slutte, at $f(x) = g(x)$ for næsten alle $x \in [-\pi,\pi]$. Faktisk ved vi p.gr.a. Theorem 1 p. 535 i AEM, at rækken (2.17) er konvergent for hvert $x \in [-\pi,\pi]$ og at

$$f(x) = \begin{cases} -1 & \text{for } x \in]-\pi, 0[\\ 1 & \text{for } x \in]0, \pi[\\ 0 & \text{for } x = -\pi, 0, \pi\ . \end{cases}$$

Opgaver

Opgave 2.1 Lad $x = (x_1, x_2, \dots)$ og $y = (y_1, y_2, \dots)$ være to talfølger. Benyt uligheden $|x_n + y_n|^2 \leq 2(|x_n|^2 + |y_n|^2)$ (se (1.19)) til at vise, at $x + y$ tilhører $\ell_2(\mathbb{N})$, hvis x og y tilhører $\ell_2(\mathbb{N})$.

Vis, at $\ell_2(\mathbb{N})$ er et underrum af $\ell(\mathbb{N})$.

Opgave 2.2 Lad $x = (x_1, x_2, \dots)$ og $y = (y_1, y_2, \dots)$ være to talfølger. Benyt uligheden $|\overline{x_n} y_n| \leq \frac{1}{2}(|x_n|^2 + |y_n|^2)$, som følger af $(|x_n| - |y_n|)^2 \geq 0$, til at slutte, at

$$\sum_{n=1}^{\infty} |\overline{x_n} y_n| < \infty \quad \text{hvis} \quad x, y \in \ell_2(\mathbb{N}) \ ,$$

og at det indre produkt givet ved (1.14) er veldefineret på $\ell_2(\mathbb{N})$.

Opgave 2.3 Eftervis, at $L_2([a,b])$ er et underrum af $F([a,b])$, og at det indre produkt givet ved (1.11) er veldefineret på $L_2([a,b])$. (Benyt lignende fremgangsmåde som i Opgave 2.1 og 2.2.)

Opgave 2.4 Benyt uligheden fra Opgave 1.12 til at indse, at normen af en funktion f i $L_2([a,b])$ kun afhænger af f's klasse, d.v.s. at $\|f\| = \|f'\|$, hvis f og f' tilhører samme klasse.

Brug endvidere Cauchy-Schwarz til at indse, at det indre produkt af to funktioner i $L_2([a,b])$ kun afhænger af de klasser, som f og g tilhører.

Opgave 2.5 1) Begrund, at en følge i $\mathbb{R}^n$ er konvergent, hhv. en Cauchy følge, netop hvis de tilsvarende n koordinatfølger er konvergente, hhv. Cauchy følger, i $\mathbb{R}$, og benyt dette til at indse, at $\mathbb{R}^n$ er fuldstændigt, og ligeså $\mathbb{C}^n$.

2) Lad H være et endeligdimensionalt indre produkt rum og lad $(e_1, \dots, e_n)$ være en ortonormalbasis for H.

Brug Pythagoras til at godtgøre, at afstanden mellem to vektorer i H er lig med afstanden i $\mathbb{R}^n$ eller $\mathbb{C}^n$ mellem deres koordinatsæt m.h.t. basen, og benyt dette til at slutte, at H er fuldstændigt.

Opgave 2.6 Lad $\sum_{n=1}^{\infty} v_n$ være en konvergent række i et indre produkt rum V.

Benyt Opgave 1.14 til at slutte, at

$$\langle \sum_{n=1}^{\infty} v_n, w \rangle = \sum_{n=1}^{\infty} \langle v_n, w \rangle$$

og

$$\langle w, \sum_{n=1}^{\infty} v_n \rangle = \sum_{n=1}^{\infty} \langle w, v_n \rangle$$

for en vilkårlig vektor $w \in V$.

Opgave 2.7 Eftervis, at sættet (2.12) er et ortonormalsæt i $C([-L, L])$.

Udnyt, at sættet (1.26) er en ortonormalbasis for $C([-\pi, \pi])$ til at vise, at (2.12) er en ortonormalbasis for $C([-L, L])$. (Vink: Benyt et simpelt variabelskift til at udtrykke det indre produkt på $C([-L, L])$ ved det indre produkt på $C([-\pi, \pi])$.)

Opgave 2.8 Eftervis, at

$$\left(\sqrt{\tfrac{2}{L}} \sin \tfrac{\pi x}{L}, \sqrt{\tfrac{2}{L}} \sin \tfrac{2\pi x}{L}, (\sqrt{\tfrac{2}{L}} \sin \tfrac{3\pi x}{L}, \dots \right)$$

og

$$\left(\tfrac{1}{\sqrt{L}}, \sqrt{\tfrac{2}{L}}\cos\tfrac{\pi x}{L}, \sqrt{\tfrac{2}{L}}\cos\tfrac{2\pi x}{L}, \sqrt{\tfrac{2}{L}}\cos\tfrac{3\pi x}{L}, \dots\right)$$

er ortonormalsæt i $L_2([0,L])$. Disse er begge ortonormalbaser for $L_2([0,L])$.

Opgave 2.9 Eftervis, at

$$\left(\sin\tfrac{\pi x}{2L}, \sin\tfrac{3\pi x}{2L}, \sin\tfrac{5\pi x}{2L}, \dots\right)$$

er et ortogonalsæt i $C([0,L])$, og opskriv det tilsvarende ortonormalsæt efter normering. Dette er en ortonormalbasis for $L_2([0,L])$.

Opgave 2.10 Lad $(e_1, e_2, \dots)$ være et ortonormalsæt i Hilbert rummet H.

1) Vis, at rækken $\sum_{n=1}^{\infty} \frac{1}{n} e_n$ er konvergent i H, og afgør for hvilke $\alpha \in \mathbb{R}$ rækken $\sum_{n=1}^{\infty} n^{\alpha} e_n$ er konvergent i H.

2) Bestem projektionen af vektorerne $e_1 \pm 2e_2$ på (underrummet udspændt af) vektoren $\sum_{n=1}^{\infty} n^{-1} e_n$.
(Man kan benytte, at $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.)

Opgave 2.11 Diskuter sammenhængen mellem ortonormaludvikling m.h.t. ortonormalbaserne (1.25) og Fourier rækkerne i Afsnit 10.3 i AEM.

Opgave 2.12 Lad a være et komplekst tal. Bestem de værdier af a, for hvilke følgen $(1, a, a^2, a^3, \dots)$ tilhører $\ell_2(\mathbb{N})$.

Opgave 2.13 Lad $(e_1, e_2, e_3, \dots)$ være en ortonormalbasis for Hilbert rummet H. Vis, at der for alle $v, w \in H$ gælder:

$$\langle v, w\rangle = \sum_{i=1}^{\infty} \langle v, e_i\rangle\langle e_i, w\rangle \ .$$

Opgave 2.14 1) Afgør for hvilke $\alpha \in \mathbb{R}$ funktionerne

$$f_\alpha(x) = \frac{1}{x^\alpha} \qquad \text{og} \qquad g_\alpha(x) = e^{-\alpha x}$$

tilhører $L_2([1,\infty[)$.

b) Eftervis, at følgerne $(f_1, f_2, f_3, \dots)$ og $(g_1, g_2, g_3, \dots)$ konvergerer (i middel) imod nulfunktionen i $L_2([1,\infty[)$.

Opgave 2.15 a) Afgør, for hvilke $\alpha \in \mathbb{R}$ funktionerne

$$f_\alpha(x) = \frac{1}{(1+|x|)^\alpha} \qquad \text{og} \qquad g_\alpha(x) = \frac{1}{|x|^\alpha}$$

tilhører $L_2(\mathbb{R})$.

b) Eftervis, at følgen $(f_1, f_2, f_3, \dots)$ konvergerer (i middel) imod nulfunktionen i $L_2(\mathbb{R})$.

Opgave 2.16 a) Afgør, for hvilke $\alpha \in \mathbb{R}$ funktionen

$$f_\alpha(x) = \frac{1}{1 + |x|^\alpha}$$

tilhører $L_2(\mathbb{R})$.

b) Eftervis, at følgen $(f_1, f_2, f_3, \dots)$ konvergerer (i middel) imod funktionen f i $L_2(\mathbb{R})$, der er lig med 1 på intervallet $[-1, 1]$ og 0 udenfor.

Opgave 2.17 Begrund, at alle funktioner af formen

$$P(x)e^{-\alpha x^2} ,$$

hvor $\alpha > 0$ og P er et polynomium, tilhører $L_2(\mathbb{R})$.

Opgave 2.18 Lad $(e_1, e_2, e_3, \dots)$ være en ortonormalbasis for $L_2(\mathbb{R})$ og antag, at en kvantemekanisk partikel på $\mathbb{R}$ befinder sig i tilstanden

$$\psi = \sum_{n=1}^{\infty} a_n e_n .$$

Begrund, at sandsynligheden for, at en måling vil resultere i tilstanden e_n, er lig med $|a_n|^2$ samt, at der gælder

$$\sum_{n=1}^{\infty} |a_n|^2 = 1 .$$

3 Operatorer på Hilbert rum

3.1 Eksempler på operatorer på Hilbert rum

Vi vil i det følgende benytte betegnelsen *lineær operator* (eller blot operator) synonymt med lineær afbildning, specielt når der er tale om lineære afbildninger mellem uendeligdimensionale vektorrum, som har vores hovedinteresse i disse noter. En lineær operator fra et Hilbert rum H_1 til et andet H_2 er således en afbildning A, der opfylder

$$A(v+w) = Av + Aw \quad \text{og} \quad A(\lambda v) = \lambda(Av) \tag{3.1}$$

for alle v, w i definitionsmængden $D(A) \subseteq H_1$ og alle skalarer λ. Ækvivalent hermed er, at A bevarer linearkombinationer, dvs. for $v_1, \ldots, v_n$ i $D(A)$ og skalarer $\lambda_1, \ldots, \lambda_n$ gælder

$$A(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \lambda_1 A v_1 + \cdots + \lambda_n A v_n \,, \tag{3.2}$$

jvf. Opgave 3.1. Vi har her brugt betegnelsen Ax i stedet for $A(x)$, som det er sædvane for operatorer. Når vi har indført en særlig betegnelse for definitionsmængden $D(A)$, også kaldet *domænet* for A, skyldes det, at en række af de operatorer, der er af størst relevans for fysik, ikke kan defineres på hele Hilbert rummet H_1, men kun på et passende "stort" underrum $D(A)$. Med "stort" mener vi her mere præcist, at

$$D(A)^\perp = \{\underline{0}\} \,, \tag{3.3}$$

altså at der ikke er andre vektorer end nulvektoren, der står vinkelret på alle vektorer i $D(A)$. Lad os se på nogle centrale eksempler.

Eksempel 3.1. a) *Differentiation.* Det er velkendt, at differentiation er en lineær operation, dvs. $(f+g)' = f' + g'$ og $(\lambda f)' = \lambda f'$, hvor f og g er differentiable funktioner på et interval $[a, b]$. Derfor er afbildningen

$$\frac{d}{dx} : f \to f'$$

en lineær operator fra $L_2([a, b])$ til $L_2([a, b])$, som dog ikke er defineret på hele $L_2([a, b])$. Valget af definitionsmængde er langt fra entydigt. Et muligt valg, og måske det umiddelbart mest nærliggende, er underrummet $C^1([a, b])$ bestående af de kontinuert differentiable funktioner. En anden mulighed er underrummet bestående af stykkevis glatte funktioner. Men der findes andre og, viser det sig, mere relevante domæner.

Til illustration ser vi på tilfældet $[a, b] = [-\pi, \pi]$, hvor vi har ortonormalbasen (1.26), hvis n'te vektor er funktionen

$$e_n(x) = \frac{1}{\sqrt{2\pi}} e^{\mathrm{i}nx} \,, \qquad n \in \mathbb{Z} \,. \tag{3.4}$$

Disse funktioner er vilkårligt ofte differentiable, og

$$e_n'(x) = \frac{\mathrm{i}n}{\sqrt{2\pi}} e^{\mathrm{i}nx} = \mathrm{i}n e_n(x) \,. \tag{3.5}$$

For en (endelig) linearkombination $\lambda_{-m}e_{-m} + \cdots + \lambda_n e_n$ af vektorerne fra ortonormalbasen har vi

$$\frac{d}{dx}(\lambda_{-m}e_{-m} + \cdots + \lambda_n e_n) = -\mathrm{i}m\lambda_{-m}e_m + \cdots + \mathrm{i}n\lambda_n e_n \, . \tag{3.6}$$

Hermed er $\frac{d}{dx}$ defineret på $\operatorname{span}\{e_n \mid n \in \mathbb{Z}\}$.

Da enhver funktion f i $L_2([-\pi,\pi])$ har en ortonormaludvikling $f = \sum_{n\in\mathbb{Z}} \lambda_n e_n$, hvor $\lambda_n = \langle e_n, f\rangle$, er det naturligt at spørge, om (3.6) kan udvides til også at gælde for "uendelige linearkombinationer", dvs. for ortonormaludviklinger, således at $\frac{d}{dx}$ defineres ved

$$\frac{d}{dx}\left(\sum_{n\in\mathbb{Z}} \lambda_n e_n\right) = \sum_{n\in\mathbb{Z}} \mathrm{i}n\lambda_n e_n \, . \tag{3.7}$$

For at denne definition skal give mening kræves, at højresiden af (3.7) er konvergent i $L_2([-\pi,\pi])$, og ifølge Lemma 2.3 er dette tilfældet netop, hvis

$$\sum_{n\in\mathbb{Z}} |n\,\lambda_n|^2 < \infty \, .$$

Men dette er ikke sandt for alle funktioner i $L_2([-\pi,\pi])$. Tag f.eks. funktionen f givet ved

$$f = \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{1}{n}\, e_n \, ,$$

hvor rækken er konvergent, da $\sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{1}{n^2} < \infty$. For denne funktion bliver højresiden af (3.7) lig med rækken $\sum_{n\in\mathbb{Z}\setminus\{0\}} i\, e_n$, som er divergent, da $\sum_{n\in\mathbb{Z}\setminus\{0\}} 1$ er divergent. Vi slutter derfor, at vi ikke ved (3.7) kan udvide operatoren $\frac{d}{dx}$ til hele $L_2([-\pi,\pi])$, men kun til domænet

$$D\left(\frac{d}{dx}\right) = \left\{\sum_{n\in\mathbb{Z}} \lambda_n e_n \mid \sum_{n\in\mathbb{Z}} |n\,\lambda_n|^2 < \infty\right\} \tag{3.8}$$

ved formlen (3.7). Dette er, som det ses, det maksimale underrum, hvorpå $\frac{d}{dx}$ kan defineres ved (3.7), og det kan vises, at det indeholder alle kontinuert differentiable funktioner f, for hvilke $f(\pi) = f(-\pi)$, jvf. Opgave 3.2. På den anden side indeholder $D(\frac{d}{dx})$ ikke hele $C^1([-\pi,\pi])$, da alle funktioner f i $D(\frac{d}{dx})$ ses at opfylde $f(-\pi) = f(\pi)$, mens der naturligvis findes funktioner i $C^1([-\pi,\pi])$, som ikke opfylder denne randbetingelse.

Det viser sig, at fysiske problemstillinger tit foreskriver det relevante domæne for en lineær operator i form af randbetingelser kombineret med krav om at operatoren skal være selvadjungeret, et centralt begreb vi skal definere i næste afsnit. I nærværende eksempel kunne vi således have lagt ud med at betragte operatoren $i\frac{d}{dx}$ på underrummet af alle C^1-funktioner f på $[-\pi,\pi]$, for hvilke $f(\pi) = f(-\pi)$. Der findes da en entydig selvadjungeret udvidelse af denne til et større domæne, nemlig $D(\frac{d}{dx})$ givet ved (3.8). Uden randbetingelse findes der ikke nogen selvadjungeret udvidelse. Noget tilsvarende gør sig gældende for mere generelle differentialoperatorer, der indføres i d) nedenfor, og diskuteres yderligere i Eksempel 3.5 d-e) og i Eksemplerne 3.17-18.

b) *Diagonaliserbare operatorer.* Det foregående eksempel er et specialtilfælde af følgende mere generelle situation. Antag, at H er et Hilbert rum, og at $(e_1, e_2, \dots)$ er en ortonormalbasis for H. Lad endvidere $(a_1, a_2, \dots)$ være en følge af skalarer. Vi kan da definere en lineær operator A ved fastsættelsen

$$A\, e_n = a_n e_n\,, \qquad n \in \mathbb{N}\,, \tag{3.9}$$

og dermed ifølge (3.2)

$$A(\lambda_1 e_1 + \cdots + \lambda_n e_n) = \lambda_1 a_1 e_1 + \cdots + \lambda_n a_n e_n\,. \tag{3.10}$$

Herved er A defineret på $\operatorname{span}\{e_1, e_2, \dots\}$. Vi kan endda udvide definitionen til domænet

$$D(A) = \{v \in H \mid \sum_{n=1}^{\infty} |a_n \langle e_n, v\rangle|^2 < \infty\} \tag{3.11}$$

ved at udvide (3.10) til

$$Av = A\left(\sum_{n=1}^{\infty} \langle e_n, v\rangle e_n\right) = \sum_{n=1}^{\infty} a_n \langle e_n, v\rangle e_n\,, \tag{3.12}$$

idet den sidste række i (3.12) er konvergent netop når $v \in D(A)$. Det er ikke svært at indse, at $D(A)$ er et underrum af H, som opfylder (3.3), jvf. Opgave 3.3.

Det er værd at bemærke, at hvis $(a_1, a_2, \dots)$ er en *begrænset* talfølge, dvs. hvis der findes et $C \geq 0$ så

$$|a_n| \leq C \text{ for alle } n = 1, 2, 3, \dots\,,$$

da er

$$\|Av\| \leq C\|v\| \tag{3.13}$$

for alle $v \in H$ (jvf. Opgave 3.3). Speciels er da $D(A) = H$, således at A i dette tilfælde er defineret på hele H. Bemærk også, at hvis $a_n = 1$ for alle n, da er A lig med *identitetsoperatoren* I_H på H, idet (3.12) giver

$$Av = \sum_{n=1}^{\infty} \langle e_n, v\rangle e_n = v = I_H v \quad \text{for alle } v \in H\,.$$

Operatorer af formen (3.12) kaldes diagonaliserbare (se også Definition 3.12). Nærmere begrundelse for denne betegnelse gives i Afsnit 3.3.

c) *Multiplikationsoperatorer på $\ell_2(\mathbb{N})$.* Hvis vi i det foregående eksempel benytter $H = \ell_2(\mathbb{N})$ med den naturlige ortonormalbasis $(\varepsilon_1, \varepsilon_2, \dots)$ antager (3.12) formen

$$A(x_1, x_2, \dots) = (a_1 x_1, a_2 x_2, \dots) \tag{3.14}$$

(overvej dette!) og domænet $D(A)$ bliver

$$D(A) = \left\{ (x_1, x_2, \dots) \mid \sum_{n=1}^{\infty} |a_n x_n|^2 < \infty \right\} . \tag{3.15}$$

Operatoren A kaldes i dette tilfælde multiplikationsoperatoren svarende til følgen $(a_1, a_2, \dots)$ og betegnes $M_{(a_1, a_2, \dots)}$.

d) *Multiplikationsoperatorer på* $L_2(I)$. Lad f være en (f.eks. kontinuert) funktion på intervallet I. Multiplikationsoperatoren M_f på $L_2(I)$ hørende til funktionen f er da givet ved

$$M_f g = f \cdot g \,, \tag{3.16}$$

hvor vi må sikre os, at højresiden tilhører $L_2(I)$, dvs. domænet for M_f er

$$D(M_f) = \{ g \in L_2(I) \mid \int_I |f(x)g(x)|^2 dx < \infty \} . \tag{3.17}$$

Igen bemærkes det (jvf. Opgave 3.4), at $D(M_f)$ er et underrum af $L_2(I)$, og at hvis f er begrænset, dvs. hvis der findes $C \geq 0$ så

$$|f(x)| \leq C \text{ for alle } x \in I \,,$$

da er

$$\|M_f g\| \leq C \|g\| \,, \tag{3.18}$$

for alle $g \in L_2(I)$. Specielt er da $D(M_f) = L_2(I)$.

I tilfældet $I = \mathbb{R}$ og $f(x) = x$, kaldes $M_f = M_x$ for *stedoperatoren* for en kvantemekanisk partikel i det endimensionale rum $\mathbb{R}$, idet middelværdien af dens stedkoordinat x i tilstanden ψ er givet ved

$$\langle \psi, M_x \psi \rangle = \int_I x |\psi(x)|^2 dx \,.$$

e) *Differentialoperatorer*. Ved sammensætning af differentiation og multiplikationsoperatorer på $L_2([a,b])$ fås mere generelle såkaldte differentialoperatorer. Vi nøjes på dette sted med at nævne anden ordens operatoren

$$B = r(x) \frac{d^2}{dx^2} + s(x) \frac{d}{dx} + t(x) \,,$$

hvor r, s, t er kontinuerte funktioner på $[a, b]$. Denne operator afbilder altså en to gange kontinuert differentiabel funktion f i funktionen

$$Bf = rf'' + sf' + tf \,.$$

Vi er allerede stødt på operatorer af denne slags, f.eks. i forbindelse med Sturm-Liouville teorien i Afsnit 4.7 i AEM, og vender tilbage hertil i Eksempel 3.5 e) og Eksempel 3.17 nedenfor.

f) *Integraloperatorer.* Modstykket til en differentialoperator er en integraloperator. En sådan er givet ved en såkaldt *integralkerne* $k(x,y)$, som er en kontinuert funktion af to variable $x, y \in [a,b]$. Den tilsvarende integraloperator K på $L_2([a,b])$ er da defineret ved

$$(Kf)(x) = \int_a^b k(x,y)f(y)dy\,. \tag{3.19}$$

Den er defineret på hele $L_2([a,b])$, idet vi kan indse at højresiden tilhører $L_2([a,b])$ for alle $f \in L_2([a,b])$ på følgende vis. For fastholdt $x \in [a,b]$ kan højresiden af (3.19) opfattes som det indre produkt af funktionerne $y \to \overline{k}(x,y)$ og f. Cauchy-Schwarz giver derfor

$$\Big|\int_a^b k(x,y)f(y)dy\Big|^2 \leq \int_a^b |k(x,y)|^2 dy \int_a^b |f(y)|^2 dy\,,$$

som også kan skrives

$$|(Kf)(x)|^2 \leq \int_a^b |k(x,y)|^2 dy \cdot \|f\|^2\,.$$

Integreres m.h.t. x fås heraf

$$\|Kf\|^2 \leq \int_a^b \int_a^b |k(x,y)|^2 dydx \cdot \|f\|^2 < \infty\,, \tag{3.20}$$

da dobbeltintegralet

$$\int_a^b \int_a^b |k(x,y)|^2 dydx \tag{3.21}$$

er endeligt (og lig med volumenet under grafen for $|k(x,y)|^2$), da k er kontinuert på kvadratet $[a,b] \times [a,b]$. Dette viser, at $Kf \in L_2([a,b])$ for $f \in L_2([a,b])$, og altså at K er defineret på hele $L_2([a,b])$.

Ovenstående kan åbenbart udvides til et vilkårligt (begrænset eller ubegrænset) interval I, hvis blot

$$\int_I \int_I |k(x,y)|^2 dydx < \infty\,.$$

Et konkret eksempel på en integraloperator så vi i forbindelse med varmeledningsligningen for en uendelig lang stang, jvf. Afsnit 11.6 i AEM. Vi fandt der, at løsningen $u(x,t)$ til varmeledningsligningen er givet ved

$$u(x,t) = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{4t}} f(y)dy$$

for $t > 0$, hvor f angiver begyndelsestemperaturfordelingen. Dvs. vi har, at

$$u(x,t) = (K_t f)(x)\ ,$$

hvor K_t er integraloperatoren på $L_2(\mathbb{R})$, hvis kerne er lig med *varmekernen*

$$k_t(x,y) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{(x-y)^2}{4t}}\ .$$

Begrænsede operatorer. I (3.20) ovenfor fandt vi for integraloperatoren K, at

$$\|Kf\| \leq C\|f\|, \quad f \in L_2([a,b]),$$

hvor konstanten C var givet ved kvadratroden af integralet (3.21). Tilsvarende uligheder fandt vi i (3.13) og (3.18) for multiplikationsoperatorer svarende til begrænsede talfølger eller funktioner. Operatorer, der tilfredsstiller uligheder af denne slags kaldes *begrænsede operatorer*.

Definition 3.2. En lineær operator A fra et Hilbert rum H_1 til et andet H_2 er begrænset, hvis dens domæne er hele H_1, og der findes en konstant $C \geq 0$, så

$$\|Av\| \leq C\|v\| \text{ for alle } v \in H_1. \tag{3.22}$$

Den mindste konstant C, der opfylder (3.22) kaldes *normen af* A og betegnes $\|A\|$. (Vi har her benyttet samme betegnelse $\|\cdot\|$ for normerne på H_1 og H_2.)

At A er begrænset er p.gr.a. lineariteten ækvivalent med at A er *kontinuert*, dvs. at der gælder

$$A(\lim_{n\to\infty} v_n) = \lim_{n\to\infty} Av_n \tag{3.23}$$

for enhver konvergent følge $(v_1, v_2, \dots)$ i H_1. Antag nemlig, at $v_n \to v$ for $n \to \infty$, og at (3.22) gælder (hvor vi kan antage $C > 0$). For givet $\varepsilon > 0$ findes da $N \in \mathbb{N}$, således at $\|v_n - v\| < \frac{\varepsilon}{C}$ for $n > N$, og derfor

$$\|Av_n - Av\| = \|A(v_n - v)\| \leq C\|v_n - v\| \leq C \cdot \frac{\varepsilon}{C} = \varepsilon$$

for $n > N$. Dette viser, at $Av_n \to Av$ for $n \to \infty$, altså at (3.23) holder. Det omvendte, altså at (3.23) medfører (3.22), bedes du om at eftervise i Opgave 3.6.

Generelt er begrænsede lineære operatorer lettere at arbejde med end *ubegrænsede operatorer* (dvs. operatorer, der ikke opfylder kravet i Definition 3.2). Mange af de operatorer, der har interesse i fysik eller i andre sammenhænge er da også begrænsede. En vigtig undtagelse herfra er imidlertid differentialoperatorerne. F.eks. har vi med betegnelser som i Eksempel 3.1 a) ovenfor, at

$$\left\|\frac{d}{dx} e_n\right\| = \|\mathrm{i}ne_n\| = |n|\|e_n\|$$

for alle $n \in \mathbb{N}$, hvilket klart viser, at operatoren $\frac{d}{dx}$ ikke opfylder (3.22) for nogen konstant C (og i øvrigt heller ikke er defineret på hele $L_2([a,b])$). I mange tilfælde har, som vi også skal se, en differentialoperator en invers afbildning, som er en begrænset operator, f.eks. en integraloperator som vi skal se i Eksempel 3.19 nedenfor, se også Opgave 3.7. Man kan derfor tit anvende resultater vedrørende begrænsede operatorer til at analysere eksempelvis differentialoperatorer. Vi vil derfor i det følgende ofte begrænse os til at formulere generelle resultater for begrænsede operatorer alene.

Endelig bemærkes, at lineære operatorer defineret på endeligdimensionale Hilbert rum automatisk er begrænsede, jvf. Opgave 3.8. Som følge heraf forekommer eksempelvis problemstillinger vedrørende kontinuitet og domæner for operatorer kun i det uendeligdimensionale tilfælde.

3.2 Selvadjungerede og unitære operatorer

De vigtigste klasser af lineære operatorer, der er brug for i de fysiske fag, er de *selvadjungerede*, også kaldt *Hermiteske*, operatorer og de *unitære* operatorer. For at kunne definere disse er det nødvendigt først at indføre begrebet *adjungeret operator*.

Adjungeret operator. Lad først $A : H_1 \to H_2$ være en lineær operator mellem endeligdimensionale Hilbert rum og lad $(e_1, \ldots, e_n)$ og $(f_1, \ldots, f_m)$ være ortonormalbaser for henholdsvis H_1 og H_2. Med hensyn til disse repræsenteres A af en $m \times n$-matrix $\underline{\underline{A}}$, hvis matrixelementer a_{ij} er givet ved (se Afsnit 6.3 i LA)

$$a_{ij} = \langle f_i, Ae_j \rangle \tag{3.24}$$

for $i = 1, \ldots, m$, $j = 1, \ldots, n$. Den adjungerede operator til A kan da defineres som operatoren $A^* : H_2 \to H_1$, som m.h.t. baserne $(f_1, \ldots, f_m)$ og $(e_1, \ldots, e_n)$ repræsenteres af den transponerede og kompleks konjungerede matrix $\underline{\underline{A}}^*$, også kaldet den *Hermitesk konjungerede* matrix til $\underline{\underline{A}}$. $\underline{\underline{A}}^*$ er altså en $n \times m$-matrix, hvis matrixelementer a^*_{ji} er givet ved

$$a^*_{ji} = \overline{a_{ij}} \tag{3.25}$$

for $i = 1, \ldots, m$, $j = 1, \ldots, n$. Da $a^*_{ji} = \langle e_j, A^* f_i \rangle$, er (3.24) ensbetydende med

$$\langle f_i, Ae_j \rangle = \overline{\langle e_j, A^* f_i \rangle} = \langle A^* f_i, e_j \rangle .$$

For to vektorer $v = \lambda_1 e_1 + \cdots + \lambda_n e_n \in H_1$ og $w = \mu_1 f_1 + \cdots + \mu_m f_m \in H_2$ giver dette ved direkte udregning (jvf. Opgave 3.11), at

$$\langle w, Av \rangle = \langle A^* w, v \rangle . \tag{3.26}$$

(Bemærk, at det indre produkt på venstresiden refererer til H_2, mens det på højresiden henviser til H_1, men vi bruger som sædvanlig samme betegnelse for de to indre produkter.) Det følger af (3.26), at operatoren A^* er uafhængig af valget af de to ortonormalbaser (Overvej dette!). I det generelle tilfælde, hvor H_1 og/eller H_2 er uendeligdimensionalt, er det mest bekvemt at tage ligningen (3.26) som udgangspunkt for definitionen af A^*. Dette gøres i følgende sætning.

Sætning 3.3. *For enhver begrænset lineær operator $A : H_1 \to H_2$ findes en entydig operator A^*, som opfylder* (3.26) *for alle $v \in H_1$, og $w \in H_2$. A^* kaldes den adjungerede operator til A. Den er begrænset og*

$$\|A^*\| = \|A\| .$$

At A^* er entydig følger let: Hvis nemlig A^* og $A^\dagger$ begge opfylder (3.26) må der gælde, at $\langle (A^* - A^\dagger)w, v \rangle = 0$ for alle $v \in H_1$ og $w \in H_2$. Altså må $(A^* - A^\dagger)w \in H_1^\perp = \{\underline{0}\}$ og dermed $A^* w = A^\dagger w$, for alle $w \in H_2$. Dette betyder jo at $A^* = A^\dagger$.

At A^* overhovedet findes skyldes et faktum, som bruges overalt i kvantemekanik, og som kan forklares som følger. For en given vektor $u \in H_1$, er afbildningen $\ell_u : v \to \langle u, v \rangle$ som bekendt lineær fra H_1 til $\mathbb{C}$ og desuden begrænset som følge af Cauchy-Schwarz

$$|\ell_n(v)| \leq \|u\| \, \|v\| .$$

Begrænsede lineære afbildninger fra H_1 til $\mathbb{C}$ kaldes *linearformer* på H_1 og mængden af disse kaldes det *duale rum* til H_1. Ovennævnte faktum går så ud på, at *enhver linearform på H_1 er af formen ℓ_u for en entydig vektor $u \in H_1$*. Dette resultat er ikke overmåde svært at vise, men vi springer argumentet over her. Har man først dette resultat, følger eksistensen af A^*w ved at bemærke, at venstresiden af (3.26) for fastholdt $w \in H_2$ er, som funktion af v, en begrænset linearform på H_1, nemlig lig med sammensætningen $\ell_w \circ A$, som er lineær og begrænset i følge Opgave 3.6. Derfor findes en entydig vektor i H_1, som vi kalder A^*w, så (3.26) er opfyldt. Hermed er A^* defineret på H_2. I Opgave 3.12 skal det vises, at A^* er lineær og begrænset og at $\|A\| = \|A^*\|$.

For to begrænsede lineære operatorer A og B fra H_1 til H_2 og $\lambda \in C$ gælder, at

$$(\lambda A)^* = \overline{\lambda} A^* \quad \text{og} \quad (A+B)^* = A^* + B^* \ .$$

Hvis endvidere C er en begrænset lineær operator fra H_2 til H_3, er

$$(CA)^* = A^*C^* \ ,$$

hvor vi, som det er sædvane for lineære operatorer, benytter betegnelsen CA i stedet for $C \circ A$. De to første relationer overlades til læseren at verificere. Den sidste ses af at

$$\langle v, CAw \rangle = \langle C^*v, Aw \rangle = \langle A^*C^*v, w \rangle$$

for alle $v \in H_3$ og alle $w \in H_1$.

Man kan også definere adjungerede til ubegrænsede lineære operatorer, men vi undlader at behandle det generelle tilfælde her og nøjes i stedet med specielle tilfælde, når vi får brug for det.

Lidt om Dirac notation. P.A.M. Dirac indførte i sin klassiker "The Principles of Quantum Mechanics" en særlig notation for linearformen ℓ_v, nemlig $\langle v|$, og kaldte linearformer for "bra-vektorer". Tilsvarende betegnede han vektorer med $|v\rangle$ i stedet for bare v og kaldte vektorerne for "ket-vektorer". Her henviser bra, hhv. ket, til de første tre, hhv. sidste tre, bogstaver i ordet bracket, som bruges for det indre produkt, der betegnes $\langle \cdot | \cdot \rangle$ i stedet for $\langle \cdot\,, \cdot \rangle$. I denne notation betegner $\langle u|v\rangle$ derfor både det indre produkt af u og v og værdien af linearformen $\langle u|$ i $|v\rangle$, altså $\langle u|(|v\rangle)$.

Tilsvarende skrives $\langle u|A|v\rangle$ i stedet for $\langle u, Av\rangle$ (eller $\langle A^*u, v\rangle$). I den kvantemekaniske formalisme er fysisk målbare størrelser repræsenteret ved selvadjungerede operatorer, (jvf. f.eks. stedoperatoren indført i Eksempel 3.1 c), og tallet

$$\langle v|A|v\rangle$$

angiver *middelværdien* (eller *forventningsværdien*) af den fysiske størrelse svarende til A i tilstanden $|v\rangle$. Dette uddybes yderligere sidst i dette afsnit.

Selvadjungerede operatorer. Vi er nu rede til at indføre de omtalte selvadjungerede operatorer.

Definition 3.4. En begrænset lineær operator $A : H \to H$ på et Hilbert rum H siges at være selvadjungeret, hvis $A = A^*$.

For at frembringe nogle eksempler på selvadjungerede operatorer ser vi først på, hvordan den adjungerede kan bestemmes i nogle konkrete tilfælde.

Eksempel 3.5. a) Den adjungerede til multiplikationsoperatoren M_f defineret i (3.16) kan aflæses af følgende regning, hvor $g, h \in L_2(I)$:

$$\begin{aligned}\langle g, M_f h\rangle &= \langle g, f \cdot h\rangle \\ &= \int_I \overline{g(x)} f(x) h(x) dx \\ &= \int_I \overline{\overline{f(x)} \cdot g(x)} h(x) dx \\ &= \langle M_{\overline{f}} g, h\rangle ,\end{aligned}$$

hvor $\overline{f}$ er den kompleks konjungerede funktion. Heraf ses, at

$$M_f^* = M_{\overline{f}} . \tag{3.27}$$

Specielt er M_f selvadjungeret, netop hvis f er en reel funktion. Det bemærkes, at hvis f også er en ubegrænset funktion (dvs, antager vilkårlig store positive eller negative værdier), da er M_f en ubegrænset selvadjungeret operator.

b) Tilsvarende fås (jvf. Opgave 3.15) for multiplikationsoperatoren $M_{(a_1,a_2,\dots)}$ defineret ved (3.14) på $\ell_2(\mathbb{N})$, at

$$M^*_{(a_1,a_2,\dots)} = M_{(\overline{a}_1,\overline{a}_2,\dots)}$$

Specielt er $M_{(a_1,a_2,\dots)}$ selvadjungeret, netop hvis talfølgen $(a_1, a_2, \dots)$ er reel.

c) For integraloperatoren K defineret ved (3.19) gælder

$$\begin{aligned}\langle f, Kg\rangle &= \int_a^b \overline{f(x)} (Kg)(x) dx \\ &= \int_a^b \overline{f(x)} \left(\int_a^b k(x,y) g(y) dy\right) dx \\ &= \int_a^b \int_a^b \overline{f(x)} k(x,y) g(y) dy dx \\ &= \int_a^b \int_a^b \overline{\overline{k(x,y)} f(x)} g(y) dy dx \\ &= \int_a^b \overline{\left(\int_a^b \overline{k(x,y)} f(x) dx\right)} g(y) dy ,\end{aligned}$$

hvor vi i sidste skridt har benyttet, at ombytning af rækkefølgen af integrationerne er tilladt. Heraf kan vi nu aflæse, at K^* er givet ved

$$(K^* f)(y) = \int_a^b \overline{k}(x,y) f(x) dx ,$$

dvs. K^* er integraloperatoren med kerne

$$k^*(x,y) = \overline{k}(y,x) .$$

Specielt er K selvadjungeret, netop hvis kernen k opfylder

$$k(x,y) = \overline{k}(y,x)$$

for $x, y \in [a,b]$.

d) For differentialoperatoren $\frac{d}{dx}$ i Eksempel 3.1 a) får vi, når f og g er kontinuert differentiable funktioner på $[-\pi, \pi]$,

$$\begin{aligned}\left\langle f, \frac{d}{dx}g \right\rangle &= \int_{-\pi}^{\pi} \overline{f}(x) g'(x) dx \\ &= \overline{f}(\pi)g(\pi) - \overline{f}(-\pi)g(-\pi) \\ &- \int_{-\pi}^{\pi} \overline{f}'(x) g(x) dx \,,\end{aligned}$$

hvor vi i sidste skridt har brugt partiel integration. Antages at $f(\pi) = f(-\pi)$ og $g(\pi) = g(-\pi)$ udgår de første to led og vi får

$$\left\langle f, \frac{d}{dx}g \right\rangle = -\left\langle \frac{d}{dx}f, g \right\rangle \,.$$

Ganges med i på begge sider fås

$$\left\langle f, i\frac{d}{dx}g \right\rangle = \left\langle i\frac{d}{dx}f, g \right\rangle \,. \tag{3.28}$$

Da basisfunktionerne e_n, der indgår i (3.7) har samme værdi i π og $-\pi$ gælder (3.28) faktisk for alle f og g tilhørende domænet $D\left(\frac{d}{dx}\right)$ givet ved (3.8), og kan også verificeres direkte ved brug af (3.7), jvf. Opgave 3.16. I denne forstand er operatoren $i\frac{d}{dx}$ et eksempel på en ubegrænset selvadjungeret operator som allered nævnt i Eksempel 3.1 a). Vi understreger igen vigtigheden af randbetingelsen, for at operatoren er selvadjungeret, et fænomen som gentager sig i det følgende eksempel. Som nævnt er disse randbetingelser ofte dikteret af den fysiske problemstilling.

e) I Sturm-Liouville teorien diskuterede vi differentialligninger af formen

$$(rf')' + q \cdot f = -\lambda p \cdot f \,, \tag{3.29}$$

hvor r er en differentiabel funktion og q og p kontinuerte funktioner på intervallet $[a,b]$. Ydermere krævedes randbetingelser af formen

$$\left.\begin{aligned} k_1 f(a) + k_2 f'(a) = 0 \\ \ell_1 f(b) + \ell_2 f'(b) = 0 \end{aligned}\right\} \tag{3.30}$$

opfyldt.

Vi ser på den tilhørende *Sturm-Liouville operator* L givet ved

$$Lf = -\frac{1}{p}\left[(rf')' + q \cdot f\right] \tag{3.31}$$

for funktioner f i Hilbert rummet $L_2([a,b], p)$, som er to gange kontinuert differentiable, og hvor vi antager, at $p > 0$ på $[a,b]$. Bemærk, at L er defineret, således at (3.29) kan skrives på formen

$$Lf = \lambda f \,. \tag{3.32}$$

Det er klart, at der er tale om en anden ordens differentialoperator. Ved partiel integration finder vi

$$\begin{aligned}\langle f, Lg\rangle_p &= \int_a^b \overline{f}(x)\,[-(r(x)g'(x))' + q(x)g(x)]\,dx \\ &= r(a)\overline{f}(a)g'(a) - r(b)\overline{f}(b)g'(b) \\ &\quad + \int_a^b \left[r(x)\overline{f'}(x)g'(x) + q(x)\overline{f}(x)g(x)\right] dx\,.\end{aligned}$$

Tilsvarende finder man

$$\begin{aligned}\langle Lf, g\rangle_p &= r(a)\overline{f'}(a)g(a) - r(b)\overline{f'}(b)g(b) \\ &\quad + \int_a^b \left[r(x)\overline{f'}(x)g(x) + q(x)\overline{f}(x)g(x)\right] dx\,.\end{aligned}$$

Antages yderligere, at f og g opfylder randbetingelserne (3.30) ses, jvf. Opgave 3.17, at randleddene på højresiderne af disse ligninger stemmer overens, således at

$$\langle f, Lg\rangle_p = \langle Lf, g\rangle_p$$

for funktioner f, g, der opfylder randbetingelserne. I denne forstand er L en selvadjungeret (ubegrænset) operator på $L_2([a,b],p)$. Vi vender tilbage til Sturm-Liouville problemet i Eksempel 3.17.

Unitære operatorer. Disse er generelt af en helt anden karakter end de selvadjungerede og defineres som følger.

Definition 3.6. En lineær operator $U : H_1 \to H_2$ mellem Hilbert rum er unitær, hvis den er begrænset og dens inverse eksisterer og er lig med U^*, dvs. hvis

$$U^*U = I_{H_1} \quad \text{og} \quad UU^* = I_{H_2}\,. \tag{3.33}$$

Eksempel 3.7. a) Lad os afgøre, for hvilke kontinuerte funktioner f på intervallet I multiplikationsoperatoren M_f defineret ved (3.16) er unitær. Hertil bemærkes først, at

$$M_f\, M_f^* = M_f M_{\overline{f}} = M_{|f|^2}\ ,$$

hvor vi har benyttet at $M_f^* = M_{\overline{f}}$ og at

$$M_f M_g = M_{f\cdot g}\,,$$

jvf. Opgave 3.18, for vilkårlige kontinuerte funktioner f og g på I. Tilsvarende er

$$M_f^* M_f = M_{|f|^2}\,.$$

Altså er M_f unitær netop hvis $M_{|f|^2}$ er lig med identitetsoperatoren på $L_2(I)$, hvilket er tilfældet hvis og kun hvis $|f|^2 = 1$, dvs.

$$|f(x)| = 1 \text{ for alle } x \in I\,,$$

jvf. Opgave 3.18.

2) Tilsvarende finder man, at multiplikationsoperatoren $M_{(a_1,a_2,\dots)}$ defineret ved (3.14) er unitær netop hvis

$$|a_n| = 1 \textit{ for alle } n \in \mathbb{N}\,.$$

Flere eksempler på unitære operatorer følger nedenfor, men det er nyttigt først at se på nogle af deres generelle egenskaber. Den måske vigtigste af disse er, at en unitær operator $U : H_1 \to H_2$ *bevarer det indre produkt* i den forstand, at

$$\langle Uv, Uw \rangle = \langle v, w \rangle \tag{3.34}$$

for alle $v, w \in H_1$. Dette følger af (3.33), idet

$$\langle Uv, Uw \rangle = \langle U^*Uv, w \rangle = \langle I_{H_1} v, w \rangle = \langle v, w \rangle .$$

Sættes $v = w$ i (3.34) fås specielt

$$\|Uv\| = \|v\| , \qquad v \in H , \tag{3.35}$$

dvs. U er *normbevarende* og dermed også *afstandsbevarende*, hvilket man også udtrykker ved at sige, at U er *isometrisk*.

Af (3.32-33) følger klart, at hvis $(e_1, e_2, \dots)$ er et ortonormalsæt i H_1, da er $(Ue_1, Ue_2, \dots)$ et ortonormalsæt i H_2. Der gælder endda, at hvis $(e_1, e_2, \dots)$ er en ortonormalbasis for H_1, da er $(Ue_1, Ue_2, \dots)$ en ortonormalbasis for H_2. Antag nemlig, at $y \in \{Ue_1, Ue_2, \dots\}^\perp$, dvs. at

$$\langle y, Ue_n \rangle = 0$$

for alle $n \in \mathbb{N}$. Vi skal så indse at $y = \underline{0}$. Sættes $x = U^*y$, er $Ux = UU^*y = I_{H_2} y = y$ ifølge (3.33). Altså er

$$\langle Ux, Ue_n \rangle = \langle y, Ue_n \rangle = 0$$

for alle $n \in \mathbb{N}$. Men da $\langle Ux, Ue_n \rangle = \langle x, e_n \rangle$, følger heraf at $x \in \{e_1, e_2, \dots\}^\perp$, og derfor at $x = 0$ ifølge Definition 2.5 af ortonormalbasis. Men så er $y = Ux = U\underline{0} = \underline{0}$, som vi skulle vise.

Vi har altså, at unitære operatorer afbilder ortonormalbaser i ortonormalbaser.

Det foregående kan sammenfattes mere præcist som følger:

Sætning 3.8. *For en begrænset operator $U : H_1 \to H_2$ er følgende fire betingelser ækvivalente:*

i) U er unitær.

ii) U er bijektiv og bevarer det indre produkt.

iii) U er bijektiv og isometrisk.

iv) U afbilder ortonormalbaser i ortonormalbaser.

Vi har ovenfor indset i) $\Rightarrow$ ii) $\Rightarrow$ iii) og i) $\Rightarrow$ iv), eftersom kravet om bijektivitet netop sikrer at $U^{-1} : H_2 \to H_1$ eksisterer. At eftervise de omvendte implikationer er overladt til Opgave 3.19.

Eksempel 3.9. Lad $(e_1, e_2, \dots)$ være en ortonormalbasis for et Hilbert rum H. Da er $(e_2, e_1, e_4, e_3, \dots)$ oplagt også en ortonormalbasis for H. Der defineres da en unitær operator U på H ved fastsættelsen

$$Ue_1 = e_2\,,\; Ue_2 = e_1\,,\; Ue_3 = e_4\,,\; Ue_4 = e_3, \dots$$

Dens virkning på en vilkårlig vektor $v = \sum_{n=1}^{\infty} \langle e_n, v\rangle e_n$ i H er givet ved

$$\begin{aligned} Uv &= \sum_{n=1}^{\infty} \langle e_n, v\rangle U e_n \\ &= \langle e_1, v\rangle e_2 + \langle e_2, v\rangle e_1 + \langle e_3, v\rangle e_4 + \langle e_4, v\rangle e_3 + \dots\,, \end{aligned}$$

hvor det ses, at højresiden er konvergent ifølge Lemma 2.3 med normkvadrat

$$\|Uv\|^2 = \sum_{n=1}^{\infty} |\langle e_n, v\rangle|^2 = \|v\|^2\,.$$

Eksempel 3.10. For en given ortonormalbasis $(e_1, e_2, \dots)$ for et Hilbert rum H haves som bekendt ifølge (2.10-11)

$$v = \sum_{n=1}^{\infty} \langle e_n, v\rangle e_n\,,$$

og

$$\sum_{n=1}^{\infty} |\langle e_n, v\rangle|^2 = \|v\|^2\,.$$

Følgen $(\langle e_1, v\rangle, \langle e_2, v\rangle, \dots)$ i $\ell_2(\mathbb{N})$ kaldes passende for *koordinatfølgen* for v m.h.t. ortonormalbasen $(e_1, e_2, \dots)$. Afbildningen $U : H \to \ell_2(\mathbb{N})$, der afbilder en vektor v i dens koordinatfølge, altså

$$Uv = (\langle e_1, v\rangle, \langle e_2, v\rangle, \dots)$$

er unitær, fordi den afbilder ortonormalbasen $(e_1, e_2, \dots)$ for H i den naturlige ortonormalbasis $(\varepsilon_1, \varepsilon_2, \dots)$ for $\ell_2(\mathbb{N})$ (overvej dette!).

Eksempel 3.11. Fourier transformationen $\mathcal{F}$ på $\mathbb{R}$ defineredes i Afsnit 10.10 i AEM ved formlen

$$(\mathcal{F}f)(w) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-iwx} dx \tag{3.36}$$

for integrable funktioner f. Det er klart at $\mathcal{F}f$ er lineær i f, se AEM p. 572 Theorem 1. Der gælder, hvis f også er kvadratisk integrabel, at

$$\int_{-\infty}^{\infty} |\mathcal{F}f(w)|^2 dw = \int_{-\infty}^{\infty} |f(x)|^2 dx\,, \tag{3.37}$$

altså at

$$\|\mathcal{F}f\| = \|f\|\,, \tag{3.38}$$

hvor $\|\cdot\|$ betegner normen på $L_2(\mathbb{R})$. (Desværre står (3.37) ikke eksplicit i AEM, men vi begrundede den ved brug af Parsevals ligning for Fourier rækker (2.15)). Af (3.38) fremgår, at $\mathcal{F}$ er en isometri på mængden af funktioner på $\mathbb{R}$, der både er integrable (for at (3.36) er veldefineret) og kvadratisk integrable (for at integralerne i (3.37) er endelige). Der er en standard måde, som vi ikke skal redegøre for her, at udvide $\mathcal{F}$ til hele $L_2(\mathbb{R})$, hvorved $\mathcal{F}$ bliver en unitær operator på $L_2(\mathbb{R})$.

Vi bemærker, at hvis ψ er bølgefunktionen for en kvantemekanisk partikel i det endimensionale rum $\mathbb{R}$, da er $\mathcal{F}\psi(w)$ bølgefunktionen i impulsvariablen w, således at $|\mathcal{F}\psi(w)|^2$ er sandsynlighedstætheden for impulsen w.

Eksempel 3.12. Vi skal nu se et eksempel på en operator, der er isometrisk uden at være unitær, idet den ikke opfylder bijektivitetskravet i Sætning 3.8 iii).

Lad $(e_1, e_2, \dots)$ være en ortonormalbasis for Hilbert rummet H og definer operatoren $S : H \to H$, kaldet *skiftoperatoren*, ved

$$S\,e_n = e_{n+1} \text{ for } n \in \mathbb{N}\,.$$

Dette betyder for en vektor $v \in H$, at

$$Sv = \sum_{n=1}^{\infty} \langle e_n, v\rangle e_{n+1}\,.$$

Her er rækken på højresiden konvergent ifølge Lemma 2.3 og

$$\|Sv\|^2 = \sum_{n=1}^{\infty} |\langle e_n, v\rangle|^2 = \|v\|^2\,.$$

Altså er S en isometri, men er ikke bijektiv, idet f.eks. vektoren e_1 ikke er billede af nogen vektor i H.

Dette fænomen kan ikke forekomme i det endeligdimensionale tilfælde, jvf. Opgave 3.20.

Lineære operatorer i kvantemekanik. Relevansen af selvadjungerede operatorer i kvantemekanik skyldes, at de repræsenterer fysisk målbare størrelser, idet $\langle\psi, A\psi\rangle$ er lig med middelværdien af den fysiske størrelse svarende til operatoren A, når partiklen befinder sig i tilstanden ψ. Dette er endnu et af kvantemekanikkens postulater.

Et eksempel er stedoperatoren M_x nævnt i Eksempel 3.1 d). Et andet vigtigt eksempel er *impulsoperatoren* $-i\frac{d}{dx}$, defineret på domænet bestående af funktioner i $L_2(\mathbb{R})$, hvis afledede er kvadratisk integrabel. I tilstanden ψ er dens middelværdi givet ved

$$\langle\psi, -i\frac{d}{dx}\psi\rangle = \langle\mathcal{F}\psi, -i\mathcal{F}\psi'\rangle = \langle\mathcal{F}\psi, w\mathcal{F}\psi\rangle = \int_{-\infty}^{\infty} w|\mathcal{F}\psi(w)|^2 dw\ ,$$

hvor vi i første lighedstegn har benyttet, at $\mathcal{F}$ bevarer det indre produkt, og andet lighedshedstegn følger af

$$(\mathcal{F}\psi')(w) = iw(\mathcal{F}\psi)(w)\ ,$$

jvf. AEM p.573. Altså er middelværdien af impulsen i tilstanden ψ lig med middelværdien af w i tilstanden $\mathcal{F}\psi$ i overensstemmelse med bemærkningerne sidst i Eksempel 3.11

Både sted- og impulsoperatoren er eksempler på ubegrænsede, selvadjungerede operatorer. Generelt medfører $A = A^*$, at

$$\langle \psi, A\psi \rangle = \langle A\psi, \psi \rangle = \overline{\langle \psi, A\psi \rangle}\,,$$

dvs. at middelværdien $\langle \psi, A\psi \rangle$ er reel, som den skal være for enhver fysisk størrelse.

Unitære operatorer er derimod tit relaterede til symmetrioperationer, såsom rotationer eller translationer, af det kvantemekaniske system. Dette skyldes især isometriegenskaben i Sætning 3.8 iii). Hvis således et kvantemekanisk system eksempelvis roteres eller translateres, vil dets oprindelige tilstand ψ ændres til en ny tilstand $U\psi$. Men som følge af sandsynlighedsfortolkningen af bølgefunktionens numeriske kvadrat er normen af enhver fysisk realiserbar tilstand lig med 1 (= den totale sandsynlighed). Derfor må

$$\|U\psi\| = \|\psi\| = 1\,.$$

Man kan argumentere for, at afbildningen U må være lineær på Hilbert rummet af (ikke nødvendigvis normerede) tilstande (dette gjorde E.P. Wigner i 1939). Det følger da, at

$$\|U\psi\| = \|\psi\|$$

gælder for alle tilstande ψ. U må også være bijektiv, fordi den inverse afbildning U^{-1} svarer til den inverse operation (rotation eller translation) på systemet. Dermed er U unitær ifølge Sætning 3.8.

3.3 Egenværdier, egenvektorer og diagonalisering

For lineære afbildninger på endeligdimensionale vektorrum er begreberne egenværdi og egenvektor velkendte fra Matematik 1 og overføres uændret til det uendeligdimensionale tilfælde. Således siges en skalar a at være *egenværdi* for en operator A på Hilbert rummet H, hvis der findes en vektor $v \neq \underline{0}$ i domænet for A, således at

$$Av = av\,, \tag{3.39}$$

og vektorer $v \neq \underline{0}$, der opfylder (3.39), kaldes *egenvektorer* for A hørende til egenværdien a. Sammen med nulvektoren udgør disse *egenrummet* hørende til a.

Egenværdier og egentilstande i kvantemekanik. Ovenfor nævnte vi, at der i den kvantemekaniske formalisme svarer en selvadjungeret operator A på Hilbert rummet af tilstande til hver fysisk målelig størrelse. Vi kan nu præcisere dette yderligere, idet de mulige udfald af en måling af størrelsen svarende til A netop er A's egenværdier. Ydermere gælder, at hvis resultatet af en måling er egenværdien a, da befinder systemet sig efter målingen i en egentilstand, dvs. tilstanden er en egenvektor, hørende til a, og når systemet befinder sig i denne egentilstand vil en måling af størrelsen med sandsynlighed 1 give resultatet a. Ifølge bemærkningerne om måleprocessen og det indre produkt i Afsnit 2.2 skulle dette betyde, at egenvektorer hørende til egenværdier forskellige fra a må være vinkelrette på egenvektorer hørende til a. Desuden skulle egenværdierne for A gerne være reelle, da fysisk målbare størrelser (f.eks. stedkoordinater og impuls) er reelle. Begge disse egenskaber sikres af, at A er selvadjungeret, som følgende sætning viser.

Sætning 3.13. *Hvis A er en selvadjungeret operator på et Hilbert rum H, da er alle dens egenværdier reelle og egenvektorer hørende til forskellige egenværdier er ortogonale.*

Antag nemlig at $Av = av$, hvor $v \neq 0$. Da er

$$a\langle v, v\rangle = \langle v, av\rangle = \langle v, Av\rangle = \langle Av, v\rangle = \langle av, v\rangle = \overline{a}\langle v, v\rangle\,,$$

hvor vi i tredie lighedstegn har brugt at $A = A^*$. Da $\langle v, v\rangle = \|v\|^2 \neq 0$ følger heraf at $a = \overline{a}$, altså at a er et reelt tal.

Antag dernæst, at $Av_1 = a_1 v_1$ og $Av_2 = a_2 v_2$, hvor nu a_1 og a_2 er reelle. Da er

$$\begin{aligned} a_2\langle v_1, v_2\rangle &= \langle v_1, a_2 v_2\rangle = \langle v_1, Av_2\rangle = \langle Av_1, v_2\rangle \\ &= \langle a_1 v_1, v_2\rangle = a_1\langle v_1, v_2\rangle\,, \end{aligned}$$

hvoraf følger at $(a_1 - a_2)\langle v_1, v_2\rangle = 0$. Hvis $a_1 \neq a_2$, må derfor $\langle v_1, v_2\rangle = 0$, som påstået.

Diagonaliserbare operatorer. De foregående bemærkninger illustrerer den vigtige rolle, som egenværdier og egenvektorer spiller i kvantemekanik. Uafhængig heraf er det p.gr.a. en operators simple virkning på sine egenvektorer generelt af stor nytte med henblik på at gennemskue operatorens egenskaber at kende dens egenvektorer og egenværdier, og jo flere des bedre. Det mest gunstige tilfælde i så henseende er, hvis operatoren er *diagonaliserbar* i følgende forstand.

Definition 3.14. En lineær operator A på et Hilbert rum H er diagonaliserbar, hvis der findes en ortonormalbasis for H bestående af egenvektorer for A. Den omtalte ortonormalbasis siges da at diagonalisere A.

Lad os for at se, hvad denne definition indebærer, antage, at A er en diagonaliserbar operator på H, og at $(e_1, e_2, \dots)$ er en ortonormalbasis, der diagonaliserer A. Betegnes de tilsvarende egenværdier med $a_1, a_2, \dots$, er altså $Ae_n = a_n e_n$ for $n \in \mathbb{N}$. Vi ser, at dette netop er det i Eksempel 3.1 b) beskrevne tilfælde. Hvis A er en begrænset operator, er følgen af egenværdier $a_1, a_2, \dots$ også begrænset, og der gælder for $v \in H$, at

$$Av = A\left(\sum_{n=1}^{\infty}\langle e_n, v\rangle e_n\right) = \sum_{n=1}^{\infty}\langle e_n, v\rangle Ae_n = \sum_{n=1}^{\infty} a_n\langle e_n, v\rangle e_n\,,$$

jvf. Opgave 3.21, altså at A er givet ved formlen (3.12). Dette viser, at A kan udtrykkes alene ved sine egenvektorer og egenværdier. Også for de ubegrænsede diagonaliserbare operatorer A af fysisk relevans i det følgende gælder, at de er givne ved (3.12) på domænet $D(A)$ defineret i (3.11). Specielt følger af Eksempel 3.1 a), at operatoren $\frac{d}{dx}$ på $D\left(\frac{d}{dx}\right)$ givet ved (3.8), er diagonaliserbar med egenværdier in, $n \in \mathbb{Z}$.

Hvis specielt H er endeligdimensionalt, og $A : H \to H$ er diagonaliserbar, da gælder, at matricen $\underline{\underline{A}}$, der repræsenterer A m.h.t. en diagonaliserende ortonormalbasis $(e_1, \dots, e_n)$ ifølge (3.24) har matrixelementer a_{ij} givet ved

$$a_{ij} = \langle e_i, Ae_j \rangle = \langle e_i, a_j e_j \rangle = a_j \langle e_i, e_j \rangle = \begin{cases} a_j & \text{hvis } i = j \\ 0 & \text{hvis } i \neq j \,, \end{cases}$$

hvor $a_1, \dots, a_n$ betegner egenværdierne. Altså er $\underline{\underline{A}}$ en diagonalmatrix med egenværdierne stående i diagonalen. Dette er naturligvis årsagen til betegnelsen "diagonaliserbar".

Eksempel 3.15. *Ortogonal projektion.* Lad $(e_1, e_2, \dots)$være et ortonormalsæt i et Hilbert rum H og lad $U = \overline{\text{span}}\{e_1, e_2, \dots\}$ være det afsluttede underrum udspændt af sættets vektorer. Ved den ortogonale projektion, eller blot *projektionen*, på U forstås afbildningen P, der afbilder en vektor $v \in H$ i dens projektion på U. Dvs. ifølge (2.7)

$$Pv = \sum_{n=1}^{\infty} \langle e_n, v \rangle e_n \,. \tag{3.40}$$

Det er oplagt, at P er lineær, og ifølge Bessels ulighed (2.9) er $\|Pv\| \leq \|v\|$, så P er begrænset med norm ≤ 1. Faktisk er $\|P\| = 1$, da jo

$$Pv = v \text{ for alle } v \in U \,.$$

Desuden er

$$Pv = 0 \text{ for alle } v \in U^{\perp}$$

ifølge (3.40), da $\langle e_n, v \rangle = 0$ for alle $n \in \mathbb{N}$ i dette tilfælde. Af de to sidste ligninger følger, at alle vektorer i $U \setminus \{\underline{0}\}$ er egenvektorer for P med egenværdi 1, mens vektorerne i $U^{\perp} \setminus \{\underline{0}\}$ er egenvektorer med egenværdi 0. At der ikke findes andre egenvektorer skal du vise i Opgave 3.26. Vi kan altid supplere $(e_1, e_2, \dots)$ med et ortonormalsæt i $U^{\perp}$, så de tilsammen udgør en ortonormalbasis for H. I konkrete tilfælde er det som regel klart, hvordan dette skal gøres, så vi undlader at give et generelt bevis for denne påstand. Dette viser, at P er diagonaliserbar, og (3.40) er i dette tilfælde identisk med (3.12), idet bidragene svarende til egenværdien 0 i (3.12) udgår.

Diagonaliserbarhed af selvadjungerede operatorer. Lad nu A være en selvadjungeret operator svarende til en fysisk målbar størrelse, eksempelvis energien af et kvantemekanisk system, og lad os antage, at de mulige udfald af en måling af energien, dvs. A's egenværdier, er $\lambda_1, \lambda_2, \dots$, samt, for simpelheds skyld, at alle tilhørende egenrum er endimensionale. Lader vi e_n være en normeret egentilstand svarende til egenværdien λ_n, da er $(e_1, e_2, \dots)$ et ortonormalsæt ifølge Sætning 3.13. Hvis dette ikke var en ortonormalbasis, fandtes der en tilstand ψ, så $\langle e_n, \psi \rangle = 0$ for alle n. Men dette betyder, at sandsynligheden $|\langle e_n, \psi \rangle|^2$ for, at systemet i tilstanden ψ har energien λ_n, er lig med 0 for alle n, hvilket er i modstrid med, at energien kan måles. Altså må egenvektorerne for A udgøre en ortonormalbasis, dvs. A må være diagonaliserbar. Denne observation begrunder relevansen af at kunne afgøre, hvilke selvadjungerede operatorer der er diagonaliserbare. Som den følgende diskussion viser, er svaret simpelt i det endeligdimensionale tilfælde, mens det uendeligdimensionale tilfælde er mere kompliceret.

Et af hovedresultaterne fra første års lineære algebra var Theorem 8.24 i LA, som siger, at enhver selvadjungeret lineær afbildning på et reelt endeligdimensionalt indre produkt rum er diagonaliserbar. (I LA bruges betegnelsen symmetrisk i stedet for selvadjungeret, jvf. Definition 8.20 i LA.) Dette resultat gælder med næsten uændret bevis også for komplekse vektorrum:

Sætning 3.16. *Enhver selvadjungeret operator på et endeligdimensionalt reelt eller komplekst Hilbert rum er diagonaliserbar.*

Eksempel 3.17. Lad H være et todimensionalt komplekst Hilbert rum og lad (e_1, e_2) være en vilkårlig ortonormalbasis for H. Med A betegner vi den lineære afbildning på H, der m.h.t. basen (e_1, e_2) repræsenteres af matricen

$$\underline{\underline{A}} = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}.$$

Da $\underline{\underline{A}}^* = \underline{\underline{A}}$ er A selvadjungeret. Vi finder A's egenværdier ved at løse ligningen $\det(\underline{\underline{A}} - \lambda \underline{\underline{I}}) = 0$ (hvor $\underline{\underline{I}}$ er enhedsmatricen), idet dette ligesom for reelle matricer betyder, at ligningssystemet $(\underline{\underline{A}} - \lambda \underline{\underline{I}})\underline{x} = 0$ har en ikke-triviel løsning $\underline{x} = \binom{x_1}{x_2}$, hvilket igen er ensbetydende med, at vektoren $x = x_1 e_1 + x_2 e_2$ opfylder $Ax = \lambda x$.

Ligningen lyder

$$\det \begin{pmatrix} 1-\lambda & i \\ -i & 1-\lambda \end{pmatrix} = (1-\lambda)^2 - 1 = 0\,,$$

hvilket giver $\lambda = 0$ eller $\lambda = 2$. For $\lambda = 0$ fås at ligningen $(\underline{\underline{A}} - \lambda \underline{\underline{I}})\binom{x_1}{x_2} = 0$ er ækvivalent med $x_1 + ix_2 = 0$, hvis løsninger er af formen $t\binom{1}{i}$, hvor $t \in \mathbb{C}$. For $\lambda = 2$ fås tilsvarende løsningerne $t\binom{1}{-i}$, $t \in \mathbb{C}$.

Egenvektorerne for A er altså af formen $t(e_1 + ie_2)$ eller $t(e_1 - ie_2)$, hvor $t \in \mathbb{C}\backslash\{0\}$. Ved normering fås ortonormalbasen $\left(\frac{1}{\sqrt{2}}(e_1 + ie_2), \frac{1}{\sqrt{2}}(e_1 - ie_2)\right)$, m.h.t. hvilken A repræsenteres af diagonalmatricen

$$\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

Man kan med god grund spørge om Sætning 3.16 mon også er gældende i det uendeligdimensionale tilfælde. Svaret herpå er et kvalificeret ja. Man kan faktisk formulere den såkaldte *spektralsætning* om selvadjungerede operatorer på et vilkårligt Hilbert rum, som generaliserer Sætning 3.16, men det vil kræve en generalisering af begrebet diagonaliserbarhed i forhold til Definition 3.14, som vi vil afholde os fra i dette kursus. Men lidt mere herom senere.

Der findes dog vigtige klasser af operatorer, for hvilke en simpel generalisering af Sætning 3.16 er mulig. En nyttig generalisering er følgende, hvis bevis vi ikke skal komme ind på her (se f.eks. §5 i [3]).

Sætning 3.18. *Enhver integraloperator K på $L_2(I)$, hvor I er et interval, med integralkerne $k(x, y)$, som opfylder (jvf. Eksempel 3.1 e) og Eksempel 3.5 c))*

$$k(x, y) = \overline{k}(y, x)\,,$$

og

$$\int_I \int_I |k(x, y)|^2 dxdy < \infty\,,$$

er diagonaliserbar.

Denne sætning kan bl.a. bruges til udledning af de fundamentale resultater vedrørende regulære Sturm–Liouville problemer, idet det viser sig, at Sturm-Liouville operatoren L, jvf. Eksempel 3.5 e), har en invers operator K som i Sætning 3.18. Dens egenvektorer er netop egenfunktionerne for Sturm-Liouville problemet, mens dens egenværdier er de inverse af egenværdierne for Sturm-Liouville problemet. Dette illustreres af følgende eksempel.

Eksempel 3.19. Vi ser på følgende Sturm-Liouville problem på $[0, \pi]$:

$$\begin{aligned} u'' - u &= -2\lambda u \\ u(0) = u(\pi) &= 0\,. \end{aligned}$$

Den tilhørende Sturm-Liouville operator er (se Eksempel 3.5 e))

$$L = -\tfrac{1}{2}\frac{d^2}{dx^2} + \tfrac{1}{2}\,.$$

Denne opfylder $\langle u, Lv\rangle = \langle Lu, v\rangle$, når u, v er to gange kontinuert differentiable funktioner på $[0, \pi]$, som opfylder randbetingelserne, og λ er en egenværdi for Sturm-Liouville problemet med tilhørende egenfunktion u netop hvis λ er egenværdi for L med tilhørende egenvektor u ifølge (3.32). Ved at løse problemet finder man (gør dette!), at egenværdierne er

$$\lambda_n = \tfrac{1}{2}(1 + n^2)\,, \qquad n \in \mathbb{N}$$

med tilhørende normaliserede egenfunktioner

$$e_n(x) = \sqrt{\frac{2}{\pi}}\sin nx\,,$$

i $L_2([0, \pi])$. Disse udgør en ortonormalbasis for $L_2([0, \pi])$, jvf. Opgave 2.7. Altså er L diagonaliserbar og

$$Lu = \sum_{n=1}^{\infty} \tfrac{1}{2}(1 + n^2)\langle e_n, u\rangle e_n\,. \tag{3.41}$$

Ligesom for $\frac{d}{dx}$ i Eksempel 3.1 a) fremgår heraf, at L er ubegrænset og defineres ved (3.41) på domænet $D(L)$ bestående af de funktioner u, for hvilke højresiden er konvergent.

Lad nu K være operatoren på $L_2([0,\pi])$ der har samme egenvektorer som L, men inverse egenværdier, dvs.

$$Ku = \sum_{n=1}^{\infty} \frac{2}{1+n^2}\langle e_n, u\rangle e_n\,. \tag{3.42}$$

Denne er åbenbart begrænset i modsætning til L, da $0 < \frac{2}{1+n^2} \leq 1$ for $n \in \mathbb{N}$. Desuden gælder

$$\begin{aligned} KLu &= \sum_{n=1}^{\infty} \tfrac{1}{2}(1+n^2)\langle e_n, u\rangle K e_n \\ &= \sum_{n=1}^{\infty} \langle e_n, u\rangle e_n = u \end{aligned}$$

for $u \in D(L)$, og tilsvarende

$$LKu = u \text{ for } u \in L_2([0,\pi])\,.$$

Altså er K den inverse operator til L.

Vi kan endda indse, at K er en integral operator, idet (3.42) ved udregning giver

$$\begin{aligned} (Ku)(x) &= \frac{2}{\pi}\sum_{n=1}^{\infty} \frac{2}{1+n^2}\left(\int_0^{\pi} u(y)\sin ny\,dy\right)\sin nx \\ &= \frac{4}{\pi}\sum_{n=1}^{\infty}\int_0^{\pi} \frac{1}{1+n^2}\sin nx \sin ny\, u(y)dy \\ &= \frac{4}{\pi}\int_0^{\pi}\left(\sum_{n=1}^{\infty} \frac{1}{1+n^2}\sin nx\sin ny\right)u(y)dy\,, \end{aligned}$$

hvor vi i sidste skridt har tilladt os at ombytte integration og summation. Hertil er det vigtigt at bemærke, at rækken

$$k(x,y) = \frac{4}{\pi}\sum_{n=1}^{\infty} \frac{1}{1+n^2}\sin nx\sin ny \tag{3.43}$$

er konvergent for alle x og y, da $|\sin nx \sin ny| \leq 1$ og $\sum\limits_{n=1}^{\infty} \frac{1}{1+n^2}$ er konvergent. Dette viser, at K er en integraloperator med kerne $k(x,y)$, som kan vises at være kontinuert og er reel og symmetrisk i x og y.

Bemærk, at vi ikke kunne have gennemført det tilsvarende for L, da rækken i (3.43) da havde indeholdt $(1+n^2)$ i stedet for $(1+n^2)^{-1}$, som ville give anledning til konvergensproblemer.

Som man vil kunne gætte sig til fra dette eksempel, gælder for et generelt regulært Sturm-Liouville problem med egenværdier $\lambda_1, \lambda_2, \ldots$ og tilhørende normaliserede egenfunktioner $e_1, e_2, \ldots$, at integralkernen for den inverse til Sturm-Liouville operatoren er givet ved

$$k(x,y) = \sum_{n=1}^{\infty} \frac{1}{\lambda_n} e_n(x)e_n(y)\,, \tag{3.44}$$

hvor det for konvergensen er afgørende, at $\lambda_n \to \infty$ for $n \to \infty$. (Der gælder faktisk at $\lambda_n \geq c \cdot n^2$ for en konstant $c > 0$.)

Eksempel 3.20. Det forrige eksempel er tæt beslægtet med beskrivelsen af en kvantemekanisk partikel, der er begrænset til intervallet $[0, \pi]$ af uendelig høje potentialbarrierer udenfor dette interval. Den til energien hørende operator (Hamilton operatoren) er

$$H = -\tfrac{1}{2}\frac{d^2}{dx^2}$$

(pånær faktoren $\hbar^2/m$, som vi ser bort fra). Da bølgefunktionen ψ skal være nul udenfor $[0, \pi]$ pålægges randbetingelserne $\psi(0) = \psi(\pi) = 0$, hvorved vi finder, at egenværdierne for H (dvs. de mulige energiniveauer) bliver

$$E_n = \tfrac{1}{2}n^2\,, \quad n \in \mathbb{N}\,,$$

med tilhørende egenfunktioner

$$\psi_n(x) = \sqrt{\frac{2}{\pi}}\sin nx\,.$$

Ser vi i stedet på operatoren svarende til stedkoordinaten x, dvs. multiplikationsoperatoren M_x, da er denne selvadjungeret (og begrænset), men den har øjensynlig ingen egenværdier. At $M_x\psi = a\psi$ betyder nemlig

$$x\psi(x) = a\psi(x)\;, \;\; \text{dvs.} \quad (x-a)\psi(x) = 0\;,$$

for (næsten alle) $x \in [a, b]$. Men da $x - a \neq 0$ for $x \neq a$ betyder dette, at $\psi(x) = 0$ for (næsten alle) $x \in [0, \pi]$, dvs. $\psi = 0$. Altså er a ikke egenværdi for noget $a \in \mathbb{C}$.

I en vis forstand er alle $a \in [0, \pi]$ dog tæt på at være egenværdier. Vælges nemlig en bølgefunktion, der er koncentreret nær $a \in [0, \pi]$, dvs. $\psi(x) = 0$ for $|x - a| \geq \varepsilon$ for et lille $\varepsilon > 0$, da er

$$\|M_x\psi - a\psi\|^2 = \int_0^\pi |x-a|^2|\psi(x)|^2 dx \leq \varepsilon^2 \int_0^\pi |\psi(x)|^2 dx = \varepsilon^2\,,$$

så afstanden mellem $M_x\psi$ og $a\psi$ kan gøres vilkårlig lille for passende normaliserede vektorer ψ. Hvis a var en egenværdi, kunne denne afstand gøres til 0 ved at vælge ψ til at være en egentilstand.

Generelt siges $a \in \mathbb{C}$ at tilhøre *spektret* $\sigma(A)$ for en selvadjungeret operator A på et Hilbert rum, hvis der for hvert $\varepsilon > 0$ findes en vektor $v \in H$ med $\|v\| = 1$ så $\|Av - av\| \leq \varepsilon$. Spektret består således dels af egenværdierne, som siges at udgøre *punktspekret* og de "generaliserede egenværdier", som udgør det *kontinuerte spektrum*. Således har M_x kun et kontinuert spektrum, nemlig $[0, \pi]$. Man kan vise, at $\sigma(A) \subseteq \mathbb{R}$ for enhver selvadjungeret operator A. Generelt skal det kontinuerte spektrum inkluderes i de mulige værdier, som den til operatoren svarende fysiske størrelse kan antage. Det er netop forekomsten af det kontinuerte spektrum, som nødvendiggør den tidligere omtalte generalisering af diagonaliserbarhedsbegrebet for at formulere den generelle spektralsætning, som dog ligger udenfor programmet i dette kursus.

Opgaver

Opgave 3.1. Eftervis, at linearitetskravene (3.1) og (3.2) er ækvivalente.

Opgave 3.2. Antag, at f er en kontinuert differentiabel funktion på $[-\pi, \pi]$, dvs. f er differentiabel med kontinuert differentialkvotient f', og således at $f(\pi) = f(-\pi)$.

Brug partiel integration til at vise, at der gælder

$$\mathrm{i}n\langle e^{\mathrm{i}nx}, f\rangle = \langle e^{\mathrm{i}nx}, f'\rangle \ ,$$

og brug Bessels ulighed for f' og Lemma 2.4 til at slutte, at $f \in D\left(\frac{d}{dx}\right)$ givet ved (3.8).

Opgave 3.3. Begrund, at der for en talfølge $(a_1, a_2, \dots)$ og en ortonormalbasis $(e_1, e_2, \dots)$ for et Hilbert rum H gælder, at

$$D(A) = \{v \in H \mid \sum_{n=1}^{\infty} |a_n \langle e_n, v\rangle|^2 < \infty\}$$

er et underrum af H. Vis (3.13), hvis følgen $(a_1, a_2, \dots)$ er begrænset.

Opgave 3.4. Begrund, at der for en kontinuert funktion på intervallet I gælder, at

$$D(M_f) = \{g \in L_2(I) \mid \int_I |f(x)g(x)|^2 dx < \infty\}$$

er et underrum af $L_2(I)$. Vis (3.18), hvis f er en begrænset funktion.

Opgave 3.5. Eftervis, at der for enhver kontinuert funktion f på intervallet $[a, b]$ gælder, at

$$u(x) = \int_a^b e^{-|x-y|} f(y) dy = \int_a^x e^{y-x} f(y) dy + \int_x^b e^{x-y} f(y) dy$$

er en løsning til differentialligningen

$$\left(-\tfrac{1}{2}\frac{d^2}{dx^2} + \tfrac{1}{2}\right) u = f$$

og opfylder randbetingelserne

$$u(a) - u'(a) = 0 \, , \quad u(b) + u'(b) = 0 \, .$$

(Det er tilladt at differentiere m.h.t. x under integraltegnet ovenfor.)

Opgave 3.6. Antag, at den lineære operator $A : H_1 \to H_2$ er kontinuert i $\underline{0}$, dvs. at A opfylder (3.22) for alle følger $(v_1, v_2, \dots)$ i H_1, som konvergerer imod $\underline{0}$. Ækvivalent hermed er, idet $A\underline{0} = \underline{0}$, at der for ethvert $\varepsilon > 0$ findes et $\delta > 0$, så $\|Av\| < \varepsilon$ for $\|v\| < \delta$.

Vælg her $\varepsilon = 1$ og et tilhørende $\delta = \delta_1$, og vis at

$$\|Av\| \leq \delta_1^{-1}\|v\|$$

for alle $v \in H_1$, og dermed at A er begrænset. Vis dernæst, at hvis $A : H_1 \to H_2$ og $B : H_2 \to H_3$ er to begrænsede operatorer, da er $BA : H_1 \to H_3$ også begrænset og

$$\|BA\| \leq \|B\|\,\|A\|\,.$$

Opgave 3.7. Lad den lineære operator A på Hilbert rummet H være givet ved (3.12) på domænet $D(A)$ defineret ved (3.11), hvor vi antager at $a_n \neq 0$ for alle $n \in \mathbb{N}$. Lad tilsvarende A^{-1} være defineret på $D(A^{-1})$ ved at erstatte a_n med a_n^{-1} for alle $n \in \mathbb{N}$.

Eftervis, at A^{-1} faktisk er den inverse afbildning til A, dvs. at

$$A^{-1}Af = f \text{ og } AA^{-1}g = g$$

for $f \in D(A)$ og $g \in D(A^{-1})$.

Giv et eksempel på en ubegrænset operator, der har en begrænset invers, og på en ubegrænset operator, der har en ubegrænset invers.

Opgave 3.8. Lad $A : H_1 \to H_2$ være en lineær afbildning mellem to endeligdimensionale Hilbert rum.

Vælg en ortonormalbasis $(e_1, \dots, e_n)$ for H_1 og vis, at

$$\|Av\| \leq (\|Ae_1\| + \cdots + \|Ae_n\|)\|v\|$$

for alle $v \in H_1$, og dermed at A er begrænset.

Opgave 3.9. På Hilbert rummet $\ell_2(\mathbb{N})$ (med sædvanligt indre produkt) betragtes afbildningen A givet ved

$$A(x_1, x_2, \dots) = (ix_1, i^2x_2, i^3x_3, \dots)\,.$$

a) Gør rede for, at $\|Ax\| = \|x\|$ for alle $x \in \ell_2(\mathbb{N})$.

b) Vis, at A er en begrænset lineær operator og bestem $\|A\|$.

c) Gør rede for at $A^4 = I_H$, (hvor A^4 som sædvanlig betegner $A \circ A \circ A \circ A$).

Opgave 3.10. Eftervis, at matrix-elementerne a_{ij} i matricen, der repræsenterer en lineær afbildning $A : H_1 \to H_2$ mellem to endeligdimensionale Hilbert rum m.h.t. to baser $(e_1, \dots, e_n)$ og $(f_1, \dots, f_m)$ er givet ved (3.24):

$$a_{ij} = \langle f_i, Ae_j \rangle$$

(se Afsnit 6.3 i LA).

Opgave 3.11. Gennemfør udregningen, der leder til (3.26).

Opgave 3.12. Vis, for en begrænset operator $A : H_1 \to H_2$, at

$$\|A\| = \sup\{\|Av\| \mid \|v\| \leq 1\}$$

og

$$\|A\| = \sup\{|\langle u, Av\rangle \mid |\|u\| \leq 1\,,\ \|v\| \leq 1\}\,,$$

og brug det sidste udtryk til at indse, at $\|A\| = \|A^*\|$.

Opgave 3.13. Lad H være et 3-dimensionalt komplekst Hilbert rum og lad A være operatoren, der m.h.t. en given ortonormalbasis (e_1, e_2, e_3) repræsenteres af matricen

$$\begin{pmatrix} 2 & 3+i & 2i \\ 0 & 2 & 0 \\ i & 1-i & 2 \end{pmatrix}.$$

Bestem matricen, der repræsenterer A^* m.h.t. samme basis og beregn $A^*(e_1+ie_2-e_3)$.

Opgave 3.14. Lad A være operatoren på $\mathbb{C}^3$, der m.h.t. den naturlige basis repræsenteres af matricen

$$\begin{pmatrix} 2 & 0 & -i \\ 0 & 2 & 0 \\ i & 0 & 2 \end{pmatrix}$$

Begrund, at A er selvadjungeret, bestem dens egenværdier samt en ortonormalbasis for $\mathbb{C}^3$, der diagonaliserer A.

Opgave 3.15. Eftervis, at der for multiplikationsoperatoren $M_{(a_1,a_2,\dots)}$ på $\ell_2(\mathbb{N})$ gælder, at

$$M^*_{(a_1,a_2,\dots)} = M_{(\bar{a}_1,\bar{a}_2,\dots)}\,.$$

Opgave 3.16. Eftervis identiteten (3.28) direkte ved brug af (3.7).

Opgave 3.17. Fuldfør udregningerne i Eksempel 3.5 e) ved at bruge randbetingelserne i Sturm-Liouville problemet til at vise, at

$$\langle f, Lg\rangle_p = \langle Lf, g\rangle_p\,,$$

hvor L er Sturm-Liouville operatoren (3.31).

Opgave 3.18. Vis, at der for multiplikationsoperatorer M_f og M_g på $L_2(I)$ gælder, at

$$M_f M_g = M_{fg}\ ,$$

og at M_f er unitær, netop hvis $|f(x)| = 1$ for alle $x \in I$.

Opgave 3.19. Eftervis implikationerne iii) $\Rightarrow$ ii) $\Rightarrow$ i) og iv) $\Rightarrow$ i) i Sætning 3.8. Til den første kan man bruge polariseringsidentiteten fra Opgave 1.9.

Opgave 3.20. Begrund, at bijektivitetskravet i ii) og iii) i Sætning 3.8 er overflødigt i det endeligdimensionale tilfælde, forudsat $\dim H_1 = \dim H_2$.

Opgave 3.21. Godtgør, at hvis $A : H_1 \to H_2$ er en begrænset operator og $\sum_{n=1}^{\infty} v_n$ er en konvergent række i H_1, da er

$$A\left(\sum_{n=1}^{\infty} v_n\right) = \sum_{n=1}^{\infty} Av_n \,.$$

Opgave 3.22. Vis, at afbildningen $U : L_2(I, \rho) \to L_2(I)$ givet ved

$$Uf = \rho^{1/2} f \,, \quad f \in L_2(I, \rho) \,,$$

er unitær.

Opgave 3.23. Godtgør, at hvis A er en diagonaliserbar operator på H_1 og $U : H_1 \to H_2$ er unitær, da er operatoren $U\,A\,U^*$ også diagonaliserbar med samme egenværdier som A.

Opgave 3.24. Eftervis, at der for enhver egenværdi a for en unitær operator $U : H \to H$ gælder, at $|a| = 1$, samt at egenvektorer hørende til forskellige egenværdier er ortogonale.

Opgave 3.25. En lineær operator $A : H \to H$ siges at være *positiv*, hvis

$$\langle f, Af \rangle > 0$$

for $f \neq \underline{0}$ i domænet for A.

a) Eftervis, at enhver egenværdi for en positiv operator er positiv.

b) Begrund, at en diagonaliserbar operator er positiv, netop hvis alle dens egenværdier er positive.

c) Vis, at egenværdierne for det regulære Sturm-Liouville problem, dvs. for $p > 0$, $q \geq 0$, $r > 0$ og $k_1 k_2 \leq 0$, $\ell_1 \ell_2 \geq 0$, alle er positive, pånær hvis $k_1 = l_1 = 0$, i hvilket tilfælde 0 også forekommer som egenværdi med tilhørende egenrum bestående af de konstante funktioner.

(Brug udregningerne i Eksempel 3.5 e) til at indse, at Sturm-Liouville operatoren er positiv, når bortses fra sidstnævnte tilfælde.)

Opgave 3.26. Vis, at projektionen P i Eksempel 3.15 ikke har andre egenværdier end 0 og 1.

Opgave 3.27 Bestem den adjungerede S^* til skiftoperatoren i Eksempel 3.12, og beregn S^*S og SS^*.

Opgave 3.28 Vis, at hvis $U : H_1 \to H_2$ og $V : H_2 \to H_3$ er unitære operatorer, da er $VU : H_1 \to H_3$ også unitær.

Giv et eksempel på (begrænsede) selvadjungerede operatorer A og B på et Hilbert rum H, således at AB ikke er selvadjungeret.

4 Appendiks

Reelle tal, fysiske størrelser og fuldstændighed

Formålet med dette appendiks er at begrunde de reelle tals (og ikke kun de rationale tals) bydende nødvendighed i (kvantitativ) naturvidenskab som følge af dens eksperimentelle grundlag, og at de grundlæggende egenskaber ved de reelle tal kan forstås i lyset heraf, herunder specielt de reelle tals fuldstændighed. Denne synsvinkel leder på naturlig vis til den generalisering af fuldstændighedsbegrebet, som indgår i Kapitel 2, og som der knyttes yderligere kommentarer til sidst i appendikset.

Vi begynder med at konstatere, at resultatet af enhver måling angives med et (eller flere) tal a samt en måleusikkerhed ε, hvilket betyder, at den sande værdi s af den målte størrelse ligger i intervallet $[a-\varepsilon, a+\varepsilon]$. Der findes naturligvis målinger, hvor resultatet ikke er behæftet med nogen usikkerhed. F.eks. kan vi med sikkerhed angive antallet af måner omkring jorden til 1, og vi kan i princippet tælle antallet af stjerner i Mælkevejen eller antallet af molekyler i en gasbeholder. Dette beror på, at vi ved, at resultatet af målingen skal være et helt tal og dermed eksakt, hvis måleusikkerheden er mindre end 1. Vi ser bort fra sådanne tælle-eksperimenter i det følgende og retter opmærksomheden mod målinger af størrelser, der såvidt vides ikke er kvantiserede, f. eks. måling af tyngdeaccelerationen g eller af lysets hastighed c.

Tilstedeværelsen af en måleusikkerhed $\varepsilon > 0$ giver en vis frihed i valget af angivelse af resultatet. Mest almindeligt er nok, som ovenfor, at angive et måleresultat som en endelig decimalbrøk. Dette vil vi også antage gjort i det følgende, og benytter betegnelsen $\mathbb{Q}_0$ for mængden af endelige decimalbrøker. Mere generelt kunne vi erstatte $\mathbb{Q}_0$ med mængden $\mathbb{Q}$ af samtlige rationale tal, men det er ikke vigtigt for det følgende, om vi vælger den ene eller den anden konvention. En vigtig grund til netop disse valg er, udover at rationale tal kan angives ved endelige udtryk, at vi føler os fortrolige med regning med rationale tal, eftersom regnereglerne for disse er direkte afledt af regnereglerne for de hele tal. En hovedpointe i det følgende er, at de imidlertid ikke er tilstrækkelige til beskrivelse af den fysiske virkelighed, som vi kender den.

Lad os tænke os, at lysets hastighed i et forsøg (I) bestemmes til 299.7 (i enheder af 10^6m/s) med en usikkerhed på 0.1. Den sande værdi c ligger da i intervallet $[299.6, 299.8]$. I et andet forsøg (II) findes den måske til at være 299.82 med usikkerhed 0.09, hvilket er konsistent med det første, idet de to intervaller $[299.6, 299.8]$ og $[299.73, 299.91]$ overlapper hinanden, og den sande værdi c må tilhøre fællesmængden $[299.73, 299.8]$. Man forbedrer så måleusikkerheden i forsøg (I) til 0.02 og finder værdien 299.81, hvilket igen er konsistent med de tidligere resultater, da intervallet $[299.79, 299.83]$ overlapper fællesmængden af de to første, og den sande værdi c må findes i dette overlap. Dette eksempel illustrerer to vigtige ting:

Ved tilstedeværelsen af usikkerhed ved enhver måling af en størrelse kan denne for det første ikke bestemmes ved en endelig serie af målinger. I den idealiserede situation, hvor man forestiller sig, at en eksperimentator udfører en uendelig serie af målinger med usikkerheder $\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots$, der aftager imod 0, er resultatet en uendelig følge $c_1, c_2, c_3, \dots$ af værdier, som er konsistente inden for usikkerhederne, hvilket vil sige, at overlappet, dvs. fællesmængden, O_N af intervallerne $[c_1 - \varepsilon_1, c_1 + \varepsilon_1]$,

$[c_2-\varepsilon_2, c_2+\varepsilon_2], \ldots, [c_N-\varepsilon_N, c_N+\varepsilon_N]$, må være ikke-tom for hvert $N \in \mathbb{N}$, dvs.

$$O_N = \bigcap_{i=1}^{N} [c_i-\varepsilon_i, c_i+\varepsilon_i] \neq \emptyset\ . \tag{A1}$$

For det andet kan en anden forsøgsudøver ved målinger af samme størrelse med usikkerheder $\varepsilon_1', \varepsilon_2', \varepsilon_3', \ldots$ nå frem til en anden følge $c_1', c_2', c_3', \ldots$ af værdier, uden at de er uenige om resultatet, hvis blot de stemmer overens inden for de angivne usikkerheder, hvilket igen betyder, at

$$O_N \cap O_N' = \bigcap_{i=1}^{N} ([c_i-\varepsilon_i, c_i+\varepsilon_i] \cap [c_i'-\varepsilon_i', c_i'+\varepsilon_i']) \neq \emptyset \tag{A2}$$

for hvert $N \in \mathbb{N}$. Dette udtrykker jo blot, at hvis man fletter de to følger af måleresultater ind imellem hinanden, da udgør de tilsammen en følge af målinger af den samme størrelse.

Vi kalder en talfølge $(c_1, c_2, c_3, \ldots)$ med tilhørende usikkerheder $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots)$, der aftager imod 0, og således at (A1) er opfyldt, for en *måleserie*. Bag sådan en måleserie ligger naturligvis forestillingen om, at de målte værdier tilnærmer den sande værdi c af den målte størrelse, altså at følgen $(c_1, c_2, c_3, \ldots)$ konvergerer mod c. Spørgsmålet, der interesserer os her, er, hvorvidt vi kan være sikre på, at følgen overhovedet konvergerer mod noget tal. Da fysikeren ikke kan levere andet end en måleserie, må al vor viden om tallet c, hvis det findes, være indeholdt i måleserien, og det er nærliggende at slutte, at værdien c af den målte størrelse skal identificeres med den fundne måleserie, og derfor også, at to måleserier, der opfylder (A2), er at betragte som ækvivalente. Overraskende nok er dette netop essensen i en af matematikere yndet definition af de reelle tal: Man tager de velkendte rationale tal, eller bare alle endelige decimalbrøker, og konstruerer så de reelle tal som måleserier, således at to måleserier, der opfylder (A2), identificeres. En mere præcis måde at udtrykke dette er følgende: Betingelsen (A2) giver anledning til en inddeling af mængden af alle måleserier i såkaldte ækvivalensklasser, hvor to måleserier tilhører samme ækvivalensklasse, netop hvis de opfylder (A2). Et reelt tal er da pr. definition en sådan ækvivalensklasse, og en måleserie i denne siges at repræsentere det pågældende reelle tal.

Spørger man nu, hvordan den sædvanlige decimalfremstilling hænger sammen med denne opfattelse af de reelle tal, er svaret simpelt: For at bestemme de første n decimaler i tallet svarende til følgen $(c_1, c_2, \ldots)$ vælges et N så stort, at usikkerheden ε_N er $\leq \frac{1}{2} \cdot 10^{-n-1}$. Da intervallet O_N har længde $\leq 2\varepsilon_N$ følger det, at hvis O_N indeholder en endelig decimalbrøk, hvis $(n+1)$'te decimal er forskellig fra 0 og 9, da har alle endelige decimalbrøker i O_N de første n decimaler tilfælles. Da $n \in \mathbb{N}$ var vilkårligt og $O_{N_1} \subseteq O_N$ for $N_1 \geq N$, bestemmes på denne måde en entydig uendelig decimalbrøk, medmindre der for enhver decimalplads n efter et vist trin n_0 gælder, at det tilsvarende O_N kun indeholder endelige decimalbrøker, der har 0 eller 9 på n'te plads. I dette sidste tilfælde får man (efter lidt nærmere overvejelser) to mulige decimalfremstillinger, den ene ender på lutter nuller og den anden på lutter nitaller fra et vist trin. Dette er den sædvanlige (og eneste) flertydighed i decimalfremstillinger. F.eks. repræsenterer $1.0000\ldots$ og $0.9999\ldots$ samme tal, idet de ikke kan skelnes ved målinger: Følgerne $(1.0,\ 1.00,\ 1.000, \dots)$ og $(0.9,\ 0.99,\ 0.999, \dots)$ med måleusikkerheder $(0.2,\ 0.02,\ 0.002, \dots)$ repræsenterer samme fysiske størrelse, fordi i dette tilfælde er $O_N = [1-2\cdot 10^{-N}, 1+2\cdot 10^{-N}]$ og $O'_N = [1-3\cdot 10^{-N}, 1+10^{-N}]$ og $O_N \cap O'_N = [1-2\cdot 10^{-N}, 1+10^{-N}] \neq \emptyset$.

Det er naturligvis afgørende for nyttigheden af decimalfremstillinger af reelle tal, at ækvivalente måleserier får samme decimalfremstilling p.gr.a. (A2), altså at den pånær ovennævnte særtilfælde giver en entydig repræsentation af en målelig størrelse, som vi kender det.

Selvom vi hermed har argumenteret for det naturlige i at opfatte reelle tal som repræsenteret ved måleserier, kan det måske forekomme som en omstændelig synsvinkel. F.eks. forekommer det ikke indlysende praktisk at betragte det rationale tal $\frac{1}{3}$ som repræsenteret ved talfølgen $(0.3,\ 0.33,\ 0.333,\ \dots)$ med usikkerheder $(0.1,\ 0.01,\ 0.001,\ \dots)$. Ikke destomindre har denne opfattelse sine fortrin frem for decimalfremstillingen. For at tage et eksempel lad os se på, hvordan man adderer reelle tal. Med to vilkårlige uendelige decimalbrøker x og y er dette uklart, da der ikke er noget sted at begynde. Tager man istedet to følger af endelige decimalbrøker $(x_1, x_2, \dots)$ og $(y_1, y_2, \dots)$, der repræsenterer de to tal med usikkerheder hhv. $(\varepsilon_1, \varepsilon_2, \dots)$ og $(\varepsilon'_1, \varepsilon'_2, \dots)$, da er det velkendt, at summen og differencen af de to størrelser er bestemte ved følgerne $(x_1 \pm y_1,\ x_2 \pm y_2, \dots)$ med usikkerheder $(\varepsilon_1 + \varepsilon'_1, \varepsilon_2 + \varepsilon'_2, \dots)$. Man skal så naturligvis sikre sig, at man ender med samme måleserie pånær ækvivalens, uanset hvilke to måleserier man repræsenterer de to tal med. Dette er intuitivt klart og overlades til læseren at overbevise sig om (man får brug for trekantsuligheden). Tilsvarende er afstanden $|x-y|$ mellem x og y lig med det reelle tal repræsenteret af $(|x_1 - y_1|, |x_2 - y_2|, \dots)$ med usikkerheder $(\varepsilon_1 + \varepsilon'_1, \varepsilon_2 + \varepsilon'_2, \dots)$. Således kan alle relationer mellem reelle tal fås ved at overføre de tilsvarende relationer for rationale tal til følger af sådanne, naturligvis med passende omtanke, idet der skal tages hensyn til måleusikkerhederne. Lad os i denne forbindelse lige nævne, at f.eks. $x \leq y$ ikke betyder, at $x_1 \leq y_1,\ x_2 \leq y_2, \dots$ for *vilkårlige* måleserier $(x_1, x_2, \dots)$ og $(y_1, y_2, \dots)$ repræsenterende x og y, men kun at der *findes* måleserier, for hvilke det gælder (overvej dette!).

Som bekendt repræsenteres et rationalt tal $\frac{m}{n}$ af en decimalbrøk, der er periodisk fra et vist trin, hvilket følger af, at der ved division med n kun kan forekomme n mulige rester, så den omtalte periode må være $\leq n$. F.eks. er $\frac{1}{7} = 0,142857142857\dots$. Det er intuitivt klart, at periodiske decimalbrøker snarere hører til undtagelsen end reglen blandt alle decimalbrøker: Kaotiske strukturer er mere hyppige end regelmæssige. Man kan faktisk vise, at ved udtagelse af en tilfældig uendelig decimalbrøk er sandsynligheden 0 for, at den er periodisk. I denne forstand udgør de rationale tal en forsvindende delmængde af de reelle tal. Så det er en højst ikke-triviel udvidelse af mængden $\mathbb{Q}_0$ (eller $\mathbb{Q}$), der er foretaget ovenfor, og der er af samme årsag ikke meget belæg for at forvente, at talværdien af fysiske størrelser i almindelighed er rationale tal. Et andet udtryk for de irrationale tals overmægtighed i forhold til de rationale, er følgende: Betragtes kun tallene i $[0,1]$, kan vi stille de rationale op i rækkefølge ved først at tage dem, der har nævner 2, dernæst dem der har nævner 3, o.s.v., dvs. $\frac{0}{2}, \frac{1}{2}, \frac{2}{2}, \frac{0}{3}, \frac{1}{3}, \frac{2}{3}, \frac{3}{3}, \cdots$, og derefter stryge alle gentagelser (dvs. $\frac{0}{3}, \frac{3}{3}, \frac{0}{4}, \frac{2}{4}, \frac{4}{4}, \cdots$). Tilbage står en talfølge som udgør samtlige rationale tal i $[0,1]$. Da disse således kan nummereres med naturlige tal, er der lige mange af dem, som der er af naturlige tal. Vi siger, at de rationale tal er tællelige, eller har samme mægtighed som $\mathbb{N}$. Man kan ret let vise, at noget tilsvarende ikke kan lade sig gøre for samtlige reelle tal i $[0,1]$. (Dette blev først gjort af G. Cantor i 1874). – Disse bemærkninger understreger forhåbentlig både omfanget og vigtigheden af det reelle talområde.

Man kan nu med rette spørge, om det ikke efter samme recept er muligt at udvide de reelle tal yderligere ved at betragte måleserier af reelle tal i stedet for måleserier af tal i $\mathbb{Q}_0$ (eller $\mathbb{Q}$). Svaret herpå er nej af den simple grund, at ethvert reelt tal kan tilnærmes vilkårlig godt med en endelig decimalbrøk (hvilket man jo netop gør ved brug af måleserier), hvoraf følger, at enhver reel måleserie er ækvivalent med en måleserie i $\mathbb{Q}_0$. Man får derfor ikke flere tal end de reelle. Dette er netop et udtryk for de reelle tals fuldstændighed, som vi nu skal forklare nærmere og dermed sætte den foregående diskussion i mere præcis forbindelse med diskussionen i Kapitel 2.

Det afgørende at bemærke er, at en måleserie ikke er andet end, hvad vi har kaldt en Cauchy følge i Afsnit 2.1, eller mere præcist: En talfølge $(x_1, x_2, \dots)$ er en Cauchy følge, netop hvis der findes en følge $(\varepsilon_1, \varepsilon_2, \dots)$ af usikkerheder, der aftager imod 0, således at de to tilsammen udgør en måleserie. Antag nemlig, at følgen $(x_1, x_2, \dots)$ med usikkerheder $(\varepsilon_1, \varepsilon_2, \dots)$ opfylder (A1). For et givet $\varepsilon > 0$ vælges så et $N \in \mathbb{N}$, så $\varepsilon_N \leq \frac{1}{2}\varepsilon$. For $n, m > N$, hvor vi kan antage at $n < m$, vælges et $a \in O_m$. Da har a højst afstand ε_n til x_n og højst afstand ε_m til x_m, så x_n og x_m har en indbyrdes afstand på højst $\varepsilon_n + \varepsilon_m$. Men da $\varepsilon_m \leq \varepsilon_n \leq \varepsilon_N$, er denne afstand $\leq 2\varepsilon_N \leq \varepsilon$. Dette viser, at $(x_1, x_2, \dots)$ er en Cauchy følge.

Hvis omvendt $(x_1, x_2, \dots)$ er en Cauchy følge, definerer vi os måleusikkerheder $(\varepsilon_1, \varepsilon_2, \dots)$ ved at lade ε_n være det største af tallene $|x_k - x_\ell|$, hvor $k, \ell \geq n$, medmindre disse alle er 0, i hvilket tilfælde vi sætter $\varepsilon = \frac{1}{n}$. Det bør bemærkes, at antagelsen om, at $(x_1, x_2, \dots)$ er en Cauchy følge, sikrer, at der er et største blandt de nævnte tal: Hvis et af tallene $|x_{k_0} - x_{\ell_0}|$ er $\neq 0$, da findes $N \in \mathbb{N}$ så $|x_k - x_\ell| \leq |x_{k_0} - x_{\ell_0}|$ for $k, \ell \geq N$, således at det største af tallene skal søges blandt de endelig mange tal $|x_k - x_\ell|$, hvor $n \leq k, \ell < N$, eller $|x_{k_0} - x_{\ell_0}|$. Med det givne valg af ε_n, $n \in \mathbb{N}$, er det klart, at følgen $(\varepsilon_1, \varepsilon_2, \dots)$ aftager imod 0, som påkrævet, samt at (A1) er opfyldt, fordi $x_N \in O_N$ for hvert $N \in \mathbb{N}$. Dette viser, at enhver Cauchy følge giver anledning til en måleserie.

Mens det for eksempelvis en fysiker er mere naturligt at tale om måleserier end om Cauchy følger, er sidstnævnte begreb dog enklere i den forstand, at en Cauchy følge alene er bestemt ved selve følgen af målte værdier, hvorimod måleusikkerhederne, som vist ovenfor, så at sige er indbygget i definitionen. I matematiklitteratur bruges derfor overalt Cauchy følger i stedet for måleserier. Således definerer man de reelle tal som ækvivalensklasser af Cauchy følger, hvor to sådanne er ækvivalente (dvs. repræsenterer samme tal), hvis de tilsvarende måleserier er ækvivalente, hvilket blot betyder, at man ved sammenfletning af de to følger igen får en Cauchy følge: To Cauchy følger $(x_1.x_2, \dots)$ og $(x'_1.x'_2, \dots)$ er ækvivalente, hvis $(x_1, x'_1, x_2, x'_2, \dots)$ er en Cauchy følge. I kraft af bemærkningerne ovenfor er denne definition af $\mathbb{R}$ helt ækvivalent med den tidligere givne ved brug af måleserier. Endvidere følger som nævnt, at $\mathbb{R}$ herved er et fuldstændigt rum i henhold til Definition 2.1:

Lad nemlig $(x_1, x_2, \dots)$ være en Cauchy følge i $\mathbb{R}$. Da er denne, som tidligere bemærket, ækvivalent med en Cauchyfølge i $\mathbb{Q}_0$ og repræsenterer altså et reelt tal

x, som den faktisk konvergerer imod. Dette sidste skulle være intuitivt klart ud fra konstruktionen af reelle tal som (ækvivalensklasser af) måleserier, der tilnærmer det tilsvarende reelle tal, men lad os for fuldstændigheds skyld udmønte argumentet i Cauchy-følge-sprog: Ifølge definitionen af Cauchy følge findes for givet $\varepsilon > 0$ et $N \in \mathbb{N}$ så $|x_n - x_m| \leq \varepsilon$ for $n, m \geq N$. Men for et fast n repræsenteres afstanden $|x - x_n|$ af følgen $(|x_1 - x_n|, |x_2 - x_n|, \dots)$, som er ækvivalent med delfølgen $(|x_N - x_n|, |x_{N+1} - x_n|, \dots)$ (overvej dette!). Men da alle denne følges elementer er $\leq \varepsilon$ for $n \geq N$, følger det, at $|x - x_n| \leq \varepsilon$ for $n \geq N$. Dette viser, at $x_n \to x$ for $n \to \infty$.

Hermed afsluttes vores diskussion af de reelle tals fuldstændighed. Udstyret med Cauchy følge begrebet kan vi imidlertid anvende den ovenfor beskrevne fremgangsmåde ved konstruktionen af de reelle tal på et vilkårligt indre produkt rum V i stedet for $\mathbb{Q}_0$ eller $\mathbb{Q}$. Dvs. vi kan udvide V til et rum H ved at definere H som mængden af Cauchy følger i V, hvor to sådanne identificeres, hvis de er ækvivalente. Alle operationer på V, såsom addition, skalarmultiplikation og indre produkt, kan da overføres til H, som derved bliver et Hilbert rum. Ethvert indre produkt rum har således en *fuldstændiggørelse* H, som er et Hilbert rum. Således er specielt fuldstændiggørelsen af $C([a,b])$ lig med $L_2([a,b])$. Men analogt med tilfældet $\mathbb{R}$ er denne opfattelse af elementerne i $L_2([a,b])$ som ækvivalensklasser af Cauchy følger af kontinuerte funktioner ikke særlig praktisk anvendelig. I stedet bør man tænke på elementerne i $L_2([a,b])$ som grænseværdien af de repræsenterende Cauchy følger, og det er netop, hvad vi har gjort i Afsnit 2.1. Problemet er blot, som vi så, at disse grænseværdier ikke er entydigt bestemte, idet to funktioner, der er ens næsten overalt i $[a,b]$ repræsenterer samme element i $L_2([a,b])$. Denne flertydighed er at sammenligne med den tilsvarende, ganske vist sparsomme, flertydighed af decimalfremstillinger af reelle tal.

Litteraturhenvisninger

1. R. Courant og D. Hilbert: *Methods of Mathematical Physics I, II.* Interscience Publishers, New York, 1953 og 1965.

2. P.A.M. Dirac: *The Principles of Quantum Mechanics.* Oxford University Press 1962.

3. B. Durhuus: *Hilbert rum med anvendelser* HCØ-tryk 1997.

4. J. von Neumann: *Mathematische Grundlagen der Quantenmechanik.* Springer Verlag, Berlin, 1932.

5. M. Reed og B. Simon: *Methods of Modern Mathematical Physics I, II.* Academic Press, New York, 1980 og 1975.

6. H.F. Weinberger: *A First Course in Partial Differential Equations.* Xerox Publishing Company, Lexington, Massachusetts, 1965.

Indeks