Kapitel 1

Udsagnslogik

1.1 Prolog: Et logisk problem om etiketter og krukker

Matematik handler om at løse problemer, der involverer objekter som mængder, funktioner, tal, afledte, integraler og mange flere. Målet med dette kapitel er at træne og forbedre dine problemløsningsevner generelt ved at forklare dig nogle værktøjer fra matematisk logik. For at identificere og motivere disse værktøjer betragter vi som et eksempel følgende problemstilling:

Eksempel 1.1.1

Opgave: Der er givet tre krukker. Du kan ikke se, hvad der er i krukkerne, men de er mærket med etiketterne "Æbler", "Begge" og "Pærer". Etiketten "Begge" betyder blot, at krukken indeholder både æbler og pærer. Problemet er nu, at nogen har byttet rundt på etiketterne på en sådan måde, at ingen etiket sidder på den rigtige krukke længere. Med andre ord: Vi ved, at der for hver krukke gælder, at dens etiket er "Æbler" eller "Begge" eller "Pærer". Vi ved også, at alle etiketterne i øjeblikket er forkerte, hvilket betyder, at den venstre krukke har den sande etiket "Begge" eller "Pærer", den midterste krukke har den sande etiket "Æbler" eller "Pærer", mens den højre krukke har den sande etiket "Æbler" eller "Begge".

For at finde ud af, hvor etiketterne hører til, kan du trække stykker af frugt fra hver krukke. Hvor mange gange skulle du trække fra krukkerne for at finde ud af, hvor etiketterne oprindeligt var?

Vi vil løse denne gåde senere, men du er velkommen til at tænke over det allerede nu!

1.2 Kom i gang med udsagnslogik

Pointen i at stille denne gåde om krukkerne og etiketterne er, at vi ved at tænke over den identificerer adskillige nøgleingredienser, der er nyttige generelt, når man arbejder på et

Figur 1.1: En gåde om etiketter og krukker

matematisk problem. Man bruger ord som "og", "eller", "ikke" og "hvis ... så", når man arbejder på opgaver af denne art. Lad os derfor introducere noget notation fra det, der kendes som *udsagnslogik*. Dette er et emne, der omhandler kombinationer af korte påstande, der kan være enten sande eller falske. Et eksempel på en sådan kort påstand er: etiketten på krukke nummer et er "Æbler". Vi vil kalde en sådan påstand for et logisk *udsagn*. Her er tre eksempler mere på sådanne udsagn: $x = 10$, $1 < y$, $a \neq p$.

Vi benytter typisk variabler som P, Q og tilsvarende til at betegne logiske udsagn. Et logisk udsagn P kan være sandt eller falsk, hvilket mere formelt formuleres således: P kan tage værdien T (for det engelske 'true') eller værdien F (for 'falsk' eller det engelske 'false'). Det er også meget normalt at anvende tallet 1 i stedet for T samt 0 i stedet for F, men i denne tekst vil vi holde os til T og F.

Af og til kan et udsagn opdeles i mindre, enklere udsagn. For eksempel består udsagnet

$$x = 10 \quad \text{og} \quad 1 < y,$$

af de to enklere udsagn $x = 10$, $1 < y$ kombineret med ordet 'og'. I udsagnslogikken skriver man

$$x = 10 \quad \land \quad 1 < y.$$

For at forbedre læsbarheden kan man placere parenteser omkring dele af udtrykket og for eksempel skrive:

$$(x = 10) \land (1 < y).$$

For at være helt præcise i betydningen af $\land$ vil vi gøre det utvetydeligt, hvornår et udsagn af formen $P \land Q$ er sandt. Vi vil gøre dette i følgende definition:

Definition 1.2.1

Lad P og Q være to logiske udsagn. Da er $P \land Q$, som udtales "P og Q", sandt, netop når både P er sandt, og Q er sandt. På tabelform:

P	Q	$P \wedge Q$
T	T	T
F	T	F
T	F	F
F	F	F

Tabellen i denne definition kaldes en *sandhedstabel* for det logiske udsagn $P \wedge Q$. Lad os forklare i flere detaljer, hvordan en sådan sandhedstabel fungerer. De to variabler P og Q kan begge være sande eller falske uafhængigt af hinanden. Med andre ord: P og Q kan antage værdien T eller værdien F uafhængigt. Derfor er der i alt fire tilfælde at overveje for:

1) P og Q antager begge værdien T,

2) P antager værdien F, og Q antager værdien T,

3) P antager værdien T, og Q antager værdien F,

4) P og Q antager begge værdien F.

For hver af disse fire muligheder angiver sandhedstabellen for $P \wedge Q$, hvilken værdi $P \wedge Q$ antager. Hvis P for eksempel antager værdien T, og Q antager værdien F, kan vi aflæse i den tredje række i sandhedstabellen, at $P \wedge Q$ antager værdien F. Dette er grunden til, at sandhedstabellen for $P \wedge Q$ har fire rækker. Hver række angiver, hvilken værdi $P \wedge Q$ antager, hvis P og Q hver især antager specifikke værdier.

Mere komplicerede logiske udsagn har også en sandhedstabel. Her er et eksempel:

Eksempel 1.2.1

Lad P, Q, R være tre logiske udsagn. Betragt nu det logiske udsagn $P \wedge (Q \wedge R)$. Vi har sat parenteser omkring $Q \wedge R$ for at tydeliggøre, at vi betragter P kombineret med $Q \wedge R$ ved hjælp af $\wedge$. Det logiske udsagn $(P \wedge Q) \wedge R$ ser umiddelbart lignende ud men er strengt taget ikke det samme som $P \wedge (Q \wedge R)$!

For at bestemme, hvornår $P \wedge (Q \wedge R)$ er sandt, og hvornår det er falsk, benytter vi Definition 1.2.1 og bestemmer dets sandhedstabel. Da vi nu har tre variabler, vil sandhedstabellen indeholde otte rækker: en række for hver mulig kombination af værdier antaget af P, Q og R. Derfor starter tabellen således:

P	Q	R
T	T	T
F	T	T
T	F	T
F	F	T
T	T	F
F	T	F
T	F	F
F	F	F

Da $P \wedge (Q \wedge R)$ består af P og $Q \wedge R$, er det praktisk først at tilføje en kolonne til $Q \wedge R$. For at udfylde de værdier, $Q \wedge R$ antager i hver af de otte rækker, benytter vi Definition 1.2.1. At de logiske udsagn i Definition 1.2.1 blev kaldt P og Q, er uden betydning, og definitionen kan også anvendes på de logiske udsagn Q og R. For eksempel antager både Q og R værdien T i de første to rækker, hvilket ifølge Definition 1.2.1 betyder, at også $Q \wedge R$ antager værdien T dér. I den tredje og fjerde række antager Q værdien F og R værdien T. Derfor antager $Q \wedge R$ værdien F i disse rækker. Vi fortsætter på denne måde og får:

P	Q	R	$Q \wedge R$
T	T	T	T
F	T	T	T
T	F	T	F
F	F	T	F
T	T	F	F
F	T	F	F
T	F	F	F
F	F	F	F

Herefter tilføjer vi en kolonne til $P \wedge (Q \wedge R)$ og bestemmer sandhedsværdierne, som udsagnet antager for hver af de otte rækker. Antag for eksempel, at (P, Q, R) antager værdierne $(\mathrm{F}, \mathrm{T}, \mathrm{T})$, svarende til værdierne i sandhedstabellens anden række. Så ser vi i den kolonne, vi netop har bestemt, at $Q \wedge R$ antager værdien T. Anvendes Definition 1.2.1 på de logiske udsagn P og $Q \wedge R$, ser vi nu, at $P \wedge (Q \wedge R)$ antager værdien F. Vi fortsætter på denne måde og bestemmer hele kolonnen for $P \wedge (Q \wedge R)$, hvilket fuldfører sandhedstabellen:

P	Q	R	$Q \wedge R$	$P \wedge (Q \wedge R)$
T	T	T	T	T
F	T	T	T	F
T	F	T	F	F
F	F	T	F	F
T	T	F	F	F
F	T	F	F	F
T	F	F	F	F
F	F	F	F	F

Vi kan opfatte $\wedge$ som en logisk operator: givet to logiske udsagn P og Q producerer den et nyt logisk udsagn $P \wedge Q$, uanset hvor komplicerede P og Q måtte være. I dette lys kaldes $\wedge$

typisk *konjunktion*, og $P \wedge Q$ kaldes konjunktionen af P og Q.

Lad os introducere nogle flere logiske operatorer. I Eksempel 1.1.1 ved vi, at alle etiketter er forkerte til at starte med. Derfor har den første krukke fra venstre ikke etiketten "Æbler". Den må derfor have den sande etiket "Begge" eller "Pærer". Dette formaliseres i den næste definition.

Definition 1.2.2

Lad P og Q være to logiske udsagn. Da er $P \vee Q$, som udtales "P eller Q", defineret ved følgende sandhedstabel:

P	Q	$P \vee Q$
T	T	T
F	T	T
T	F	T
F	F	F

Operatoren $\vee$ kaldes *disjunktion* og $P \vee Q$ disjunktionen af P og Q. En tredje logisk operator er negationen af et logisk udsagn. Vi har allerede berørt dette i Eksempel 1.1.1, hvor vi fik givet, at etiketterne er forkert placeret. For eksempel er den sande etiket på den midterste krukke derfor ikke "Begge". Et udsagn som $x \neq 0$ er simpelthen negationen af udsagnet $x = 0$. Lad os definere negationsoperatoren formelt.

Definition 1.2.3

Lad P være et logisk udsagn. Da er $\neg P$, som udtales "ikke P", defineret ved følgende sandhedstabel:

P	$\neg P$
T	F
F	T

Som operator kaldes $\neg$ for *negation*, og $\neg P$ kaldes derfor også negationen af P. Vi har nu rigeligt med ingredienser til at kunne opstille adskillige logiske udsagn. Lad os overveje et eksempel.

Eksempel 1.2.2

Betragt det logiske udsagn $P \vee (Q \wedge \neg P)$. Vi vil bestemme dets sandhedstabel. Da vi kun har to variable P og Q, vil denne sandhedstabel indeholde fire rækker. I $P \vee (Q \wedge \neg P)$ indgår det simplere logiske udsagn $Q \wedge \neg P$, som igen indeholder det logiske udsagn $\neg P$. For at kunne bestemme sandhedstabellen for $P \vee (Q \wedge \neg P)$ giver det derfor mening at tilføje en kolonne til $\neg P$ og en til $Q \wedge \neg P$, hvorved vi gradvist arbejder os hen imod at have bestemt sandhedsværdierne for hele det logiske udsagn $P \vee (Q \wedge \neg P)$. Således bliver resultatet følgende:

P	Q	$\neg P$	$Q \wedge \neg P$	$P \vee (Q \wedge \neg P)$
T	T	F	F	T
F	T	T	T	T
T	F	F	F	T
F	F	T	F	F

Lad os sammenligne vores fundne sandhedstabel med sandhedstabellen for $P \vee Q$ i Definition 1.2.2. For givne sandhedsværdier af P og Q er sandhedsværdierne af $P \vee (Q \wedge \neg P)$ og $P \vee Q$ altid de samme! Med andre ord: hvis vi kun betragter de tre kolonner i vores fundne sandhedstabel, der svarer til P, Q og $P \vee (Q \wedge \neg P)$, får vi præcis samme tabel som sandhedstabellen i Definition 1.2.2. To logiske udsagn, der ser forskellige ud, kan åbenbart have samme sandhedstabel.

1.3 Logisk konsekvens og ækvivalens

De logiske operatorer, vi har introduceret indtil videre, $\neg$, $\wedge$ og $\vee$, gør det muligt at skrive adskillige logiske udsagn på en præcis måde. Pointen med logik er dog at gøre argumenter og ræsonnementer mere præcise. Vi vil gerne være i stand til at sige noget i retning af, hvis P er sandt, kan vi konkludere, at også Q er sandt. For eksempel, hvis $x > 0$, er også $x > -1$. Dette kan formaliseres ved brug af det logiske symbol $\Rightarrow$, kaldet en *implikation*, med skrivemåden $P \Rightarrow Q$ ved følgende definition:

Definition 1.3.1

Det logiske udsagn $P \Rightarrow Q$ defineres ved følgende sandhedstabel:

P	Q	$P \Rightarrow Q$
T	T	T
F	T	T
T	F	F
F	F	T

I almindeligt sprog udtaler man typisk $P \Rightarrow Q$ som "P medfører Q" eller "hvis P, så Q". Det kan blive nødvendigt at skrive det logiske udsagn $P \Rightarrow Q$ som $Q \Leftarrow P$.

Vi vil betegne to særlige typer logiske udsagn ved **T** og **F**. Det logiske udsagn **T** repræsenterer en påstand, der altid er sand, som for eksempel påstanden $5 = 5$. Et sådant logisk udsagn kaldes en *tautologi*. I modsætning hertil repræsenterer det logiske udsagn **F** en påstand, der altid er falsk, som for eksempel $5 \neq 5$. Dette kaldes en *modstrid*. Tænker vi i implikationer, betyder påstanden, at $P \Rightarrow Q$ altid er sandt for givne logiske udsagn P og Q, faktisk, at $P \Rightarrow Q$ er en tautologi. Hvis $P \Rightarrow Q$ er en tautologi, viser sandhedstabellen for implikationen i Definition 1.3.1, at hvis P er sandt, er også Q sandt.

Hvis $P \Rightarrow Q$ er en tautologi, siges Q at være en *logisk konsekvens* af P, alternativt at Q medføres af P, på engelsk 'Q is implied by P', hvilket forklarer brugen af begrebet implikation for symbolet $\Rightarrow$. Lad os se på et eksempel på logisk konsekvens.

Eksempel 1.3.1

Lad P og Q være logiske udsagn. Vi påstår nu, at det logiske udsagn $P \vee Q$ er en logisk konsekvens af P. Vi kan vise påstanden ved at vise, at det logiske udsagn $P \Rightarrow (P \vee Q)$ altid er sandt, med andre ord at $P \Rightarrow (P \vee Q)$ er en tautologi. Vi bestemmer sandhedstabellen for $P \Rightarrow (P \vee Q)$ ved på sædvanlig vis først at opskrive alle kombinationer af sandhedsværdier for P og Q:

P	Q
T	T
F	T
T	F
F	F

Dernæst tilføjer vi en kolonne til $P \vee Q$ for at gøre det nemt for os selv, da det forekommer i det mere komplicerede udsagn $P \Rightarrow (P \vee Q)$, som vi skal arbejde os hen imod. Ved brug af Definition 1.2.2 skriver vi:

P	Q	$P \vee Q$
T	T	T
F	T	T
T	F	T
F	F	F

Nu tilføjes en kolonne til $P \Rightarrow (P \vee Q)$, og ved brug af Definition 1.3.1 bestemmes dens sandhedsværdier ud fra værdierne af P og $P \vee Q$. Resultatet er den sandhedstabel, vi ønsker at bestemme:

P	Q	$P \vee Q$	$P \Rightarrow (P \vee Q)$
T	T	T	T
F	T	T	T
T	F	T	T
F	F	F	T

Da kolonnen tilhørende udtrykket $P \Rightarrow (P \vee Q)$ kun indeholder T'er, kan vi konkludere, at $P \Rightarrow (P \vee Q)$ er en tautologi. Vi har hermed vist, at det ganske rigtigt gælder, at P altid medfører $P \vee Q$, med andre ord at $P \vee Q$ er en logisk konsekvens af P.

Stærkere end en implikation er det, der er kendt som en *biimplikation*, betegnet ved $\Leftrightarrow$ og defineret som følger:

Definition 1.3.2

Det logiske udsagn $P \Leftrightarrow Q$, som udtales "P hvis og kun hvis Q", er defineret ved følgende sandhedstabel:

P	Q	$P \Leftrightarrow Q$
T	T	T
F	T	F
T	F	F
F	F	T

Sætningen "P hvis og kun hvis Q" for det logiske udsagn $P \Leftrightarrow Q$ kan opdeles i to dele, nemlig "P hvis Q" og "P kun hvis Q". Den første del, "P hvis Q", fortæller blot, at $P \Leftarrow Q$, mens "P kun hvis Q" kan koges ned til udsagnet $P \Rightarrow Q$. Dette forklarer begrebet biimplikation for symbolet $\Leftrightarrow$: to implikationer er kombineret i ét symbol. Senere i Sætning 1.3.4, Ligning (1.22) kommer vi til at se en mere formel måde at udtrykke en biimplikation på som to implikationer.

Eksempel 1.3.2

I Eksempel 1.2.2 så vi, at sandhedstabellen for $P \vee Q$ er identisk med den for $P \vee (Q \wedge \neg P)$. Hvad betyder dette for sandhedstabellen for det logiske udsagn $(P \vee Q) \Leftrightarrow (P \vee (Q \wedge \neg P))$? Lad os udvide tabellen i Definition 1.2.2 med resultatet fra Eksempel 1.2.2:

P	Q	$P \vee Q$	$P \vee (Q \wedge \neg P)$
T	T	T	T
F	T	T	T
T	F	T	T
F	F	F	F

En kolonne til det logiske udsagn $(P \vee Q) \Leftrightarrow (P \vee (Q \wedge \neg P))$ tilføjes tabellen, og ved brug af Definition 1.3.2 udfyldes kolonnen til følgende resultat:

P	Q	$P \vee Q$	$P \vee (Q \wedge \neg P)$	$(P \vee Q) \Leftrightarrow (P \vee (Q \wedge \neg P))$
T	T	T	T	T
F	T	T	T	T
T	F	T	T	T
F	F	F	F	T

Da den yderste kolonne kun indeholder T'er, kan vi konkludere, at $(P \vee Q) \Leftrightarrow (P \vee (Q \wedge \neg P))$ er en tautologi.

Pointen er nu, at hvis $R \Leftrightarrow S$ er en tautologi for nogle, muligvis komplicerede, logiske udsagn R og S, så har R og S samme sandhedstabel. Med andre ord: hvis R er sandt, så er også S sandt, og omvendt hvis S er sandt, så er også R sandt. Hvis $R \Leftrightarrow S$ er en tautologi, siges de logiske udsagn R og S derfor at være *logisk ækvivalente*. I Eksempel 1.3.2 kan vi konkludere, at de logiske udsagn $P \vee Q$ og $P \vee (Q \wedge \neg P)$ er logisk ækvivalente. Pointen med eksempelet var at vise, at man nogle gange kan omskrive en logisk påstand til en enklere form. Der findes adskillige ganske praktiske tautologier, der kan benyttes til at omskrive logiske udsagn til en enklere form. Vi oplister herunder som en start nogle, der omhandler konjunktion, disjunktion og negation.

Sætning 1.3.1

Lad P, Q og R være logiske udsagn. Da er alle følgende udtryk tautologier.

$$P \wedge P \Leftrightarrow P \tag{1.1}$$

$$P \vee P \Leftrightarrow P \tag{1.2}$$

$$P \wedge Q \Leftrightarrow Q \wedge P \tag{1.3}$$

$$P \vee Q \Leftrightarrow Q \vee P \tag{1.4}$$

$$P \wedge (Q \wedge R) \Leftrightarrow (P \wedge Q) \wedge R \tag{1.5}$$
$$P \vee (Q \vee R) \Leftrightarrow (P \vee Q) \vee R \tag{1.6}$$
$$P \wedge (Q \vee R) \Leftrightarrow (P \wedge Q) \vee (P \wedge R) \tag{1.7}$$
$$P \vee (Q \wedge R) \Leftrightarrow (P \vee Q) \wedge (P \vee R) \tag{1.8}$$

Bevis. Vi kan bevise, at de nævnte logiske udsagn er tautologier, ved at bestemme deres sandhedstabeller. Det ville fylde temmelig mange sider at gøre for dem alle, så lad os blot bevise én af dem her, nemlig Ligning (1.5). Vi ønsker dermed at vise, at $P \wedge (Q \wedge R) \Leftrightarrow (P \wedge Q) \wedge R$ er en tautologi. I Eksempel 1.2.1 bestemte vi sandhedstabellen for $P \wedge (Q \wedge R)$, så vi undlader at gentage det her. Vi skal her blot bestemme sandhedstabellen for $(P \wedge Q) \wedge R$, hvilket kan gøres ligesom for $P \wedge (Q \wedge R)$ i Eksempel 1.2.1. Derefter vil vi kunne bestemme sandhedstabellen for $P \wedge (Q \wedge R) \Leftrightarrow (P \wedge Q) \wedge R$ ved hjælp af Definition 1.3.2. Resultatet bliver følgende:

P	Q	R	$P \wedge (Q \wedge R)$	$P \wedge Q$	$(P \wedge Q) \wedge R$	$P \wedge (Q \wedge R) \Leftrightarrow (P \wedge Q) \wedge R$
T	T	T	T	T	T	T
F	T	T	F	F	F	T
T	F	T	F	F	F	T
F	F	T	F	F	F	T
T	T	F	F	T	F	T
F	T	F	F	F	F	T
T	F	F	F	F	F	T
F	F	F	F	F	F	T

Vi ser, at det logiske udsagn $P \wedge (Q \wedge R) \Leftrightarrow (P \wedge Q) \wedge R$ kun antager sandhedsværdien T, uanset hvilke værdier P, Q og R antager. Derfor kan vi konkludere, at $P \wedge (Q \wedge R) \Leftrightarrow (P \wedge Q) \wedge R$ er en tautologi.

Alle de andre påstande i sætningen kan vises på samme vis, hvilket vi undlader at gøre her. Læseren opfordres til selv at bevise mindst én påstand mere. □

I ord siger Ligning (1.6), at når man bestemmer disjunktionen af tre logiske udsagn, betyder placeringen af parenteserne ikke noget. Derfor vil man typisk helt udelade parenteserne og skrive $P \vee Q \vee R$. Tilsvarende siger Ligning (1.5), at også for konjunktionen af tre logiske udsagn kan parenteserne placeres frit, hvorfor man også typisk skriver $P \wedge Q \wedge R$ uden nogen risiko for tvetydighed. Men dette ændrer sig, hvis både konjunktion og disjunktion forekommer i det samme udtryk. Så betyder parenteserne noget. Lad os se på et eksempel.

Eksempel 1.3.3

Betragt de logiske udsagn $(P \wedge Q) \vee R$ og $P \wedge (Q \vee R)$. Vi hævder, at disse ikke er logisk ækvivalente. Dette kan vises ved bestemmelse af deres sandhedstabeller. Men for at vise, at to logiske udsagn ikke er logisk ækvivalente, kan vi faktisk nøjes med blot at finde nogle eksempelværdier for P, Q og R, for hvilke $(P \wedge Q) \vee R$ og $P \wedge (Q \vee R)$ ikke begge er sande og ikke begge er falske. Lad os for eksempel undersøge, hvornår $(P \wedge Q) \vee R$ er falsk. Dette sker, netop når $P \wedge Q$ er falsk, og R er falsk. Derfor er $(P \wedge Q) \vee R$ falsk, netop hvis P og Q ikke begge er sande, og R er falsk. Samtidig er $P \wedge (Q \vee R)$ falsk, når P er falsk. Hvis (P, Q, R) antager værdierne (F, T, T), så er $(P \wedge Q) \vee R$ derfor sandt, men $P \wedge (Q \vee R)$ er falsk. Dette betyder, at der i sandhedstabellen for de to udtryk er en række, der ser således ud:

P	Q	R	$\ldots$	$(P \wedge Q) \vee R$	$P \wedge (Q \vee R)$
$\vdots$					
F	T	T	$\ldots$	T	F
$\vdots$					

Dette er nok til at konkludere, at de logiske udsagn $(P \wedge Q) \vee R$ og $P \wedge (Q \vee R)$ ikke er logisk ækvivalente. Hvis de var det, ville det logiske udsagn $(P \wedge Q) \vee R \Leftrightarrow P \wedge (Q \vee R)$ være en tautologi og derfor kun tage værdien T, men vi har nu set, at dens sandhedstabel vil indeholde følgende række:

P	Q	R	$\ldots$	$(P \wedge Q) \vee R$	$P \wedge (Q \vee R)$	$(P \wedge Q) \vee R \Leftrightarrow P \wedge (Q \vee R)$
$\vdots$						
F	T	T	$\ldots$	T	F	F
$\vdots$						

Vi har derfor, at det logiske udsagn $(P \wedge Q) \vee R \Leftrightarrow P \wedge (Q \vee R)$ ikke er en tautologi, og derfor, at de logiske udsagn $(P \wedge Q) \vee R$ og $P \wedge (Q \vee R)$ ikke er logisk ækvivalente.

Der findes flere nyttige tautologier til arbejdet med logiske udsagn. Ud over konjunktionen $\wedge$ og disjunktionen $\vee$ involverer de følgende også negationen $\neg$. Vi overlader beviserne til læseren.

Sætning 1.3.2

Lad P, Q og R være logiske udsagn. Da er alle følgende udtryk tautologier.

$$\begin{aligned} P \vee \neg P &\Leftrightarrow \mathbf{T} && (1.9) \\ P \wedge \neg P &\Leftrightarrow \mathbf{F} && (1.10) \\ P &\Leftrightarrow \neg(\neg P) && (1.11) \\ \neg(P \vee Q) &\Leftrightarrow \neg P \wedge \neg Q && (1.12) \\ \neg(P \wedge Q) &\Leftrightarrow \neg P \vee \neg Q && (1.13) \\ \neg \mathbf{T} &\Leftrightarrow \mathbf{F} && (1.14) \\ \neg \mathbf{F} &\Leftrightarrow \mathbf{T} && (1.15) \end{aligned}$$

Identiteterne (1.12) og (1.13) kaldes *De Morgans love*. Der findes også et par tautologier, der beskriver, hvordan $\wedge$ og $\vee$ opfører sig i kombinationer med tautologier og modstridigheder. Igen overlader vi beviserne for disse til læseren.

Sætning 1.3.3

Lad P, Q og R være logiske udsagn. Da er alle følgende udtryk tautologier.

$$P \vee \mathbf{F} \quad \Leftrightarrow \quad P \tag{1.16}$$
$$P \wedge \mathbf{T} \quad \Leftrightarrow \quad P \tag{1.17}$$
$$P \wedge \mathbf{F} \quad \Leftrightarrow \quad \mathbf{F} \tag{1.18}$$
$$P \vee \mathbf{T} \quad \Leftrightarrow \quad \mathbf{T} \tag{1.19}$$

Listerne over tautologier i Sætningerne 1.3.1, 1.3.2 og 1.3.3 kan ofte hjælpe med omskrivning af logiske udsagn til en logisk ækvivalent form. Lad os se på et eksempel.

Eksempel 1.3.4

Betragt som i Eksemplerne 1.2.2 og 1.3.2 det logiske udsagn $P \vee (Q \wedge \neg P)$. Vi har allerede set, at det er logisk ækvivalent med $P \vee Q$. Lad os bevise det ved hjælp af Sætning 1.3.1 som et alternativ til at bestemme sandhedstabeller. Først og fremmest ser vi ved hjælp af Ligning (1.8), at

$$P \vee (Q \wedge \neg P) \Leftrightarrow (P \vee Q) \wedge (P \vee \neg P).$$

Ved hjælp af Ligning (1.9) konkluderer vi, at

$$P \vee (Q \wedge \neg P) \Leftrightarrow (P \vee Q) \wedge \mathbf{T},$$

hvilket ved brug af Ligning (1.17) kan reduceres til

$$P \vee (Q \wedge \neg P) \Leftrightarrow P \vee Q.$$

Med andre ord, ved hjælp af Sætning 1.3.1 kan man bevise logiske ækvivalenser uden at skulle bestemme sandhedstabeller. Beviset af sætningen selv kræver selvfølgelig, at der bestemmes adskillige sandhedstabeller, men det behøver kun at blive gjort én gang. I matematikken er pointen med en sætning generelt, at den indeholder et eller flere nyttige resultater med et bevis. Når beviset er givet, kan man benytte resultatet i sætningen hvor nødvendigt uden at skulle bevise sætningen igen.

Tautologierne i Sætning 1.3.1 involverer kun negation, konjunktion og disjunktion. Her er yderligere tre, der er særdeles nyttige, som involverer implikation og biimplikation.

Sætning 1.3.4

Lad P og Q være logiske udsagn. Da er alle følgende udtryk tautologier.

$$(P \Rightarrow Q) \quad \Leftrightarrow \quad (\neg P \vee Q) \tag{1.20}$$
$$(P \Rightarrow Q) \quad \Leftrightarrow \quad (\neg Q \Rightarrow \neg P) \tag{1.21}$$
$$(P \Leftrightarrow Q) \quad \Leftrightarrow \quad (P \Rightarrow Q) \wedge (Q \Rightarrow P) \tag{1.22}$$
$$P \quad \Leftrightarrow \quad (\neg P \Rightarrow \mathbf{F}) \tag{1.23}$$

Bevis. Ligesom med Sætning 1.3.1 kan disse påstande bevises ved bestemmelse af sandhedstabeller for hver af dem. Dette gør vi her for den andennævnte påstand, og beviser for de resterende overlades til læseren:

P	Q	$P \Rightarrow Q$	$\neg P$	$\neg Q$	$\neg Q \Rightarrow \neg P$	$(P \Rightarrow Q) \Leftrightarrow (\neg Q \Rightarrow \neg P)$
T	T	T	F	F	T	T
F	T	T	T	F	T	T
T	F	F	F	T	F	T
F	F	T	T	T	T	T

Da den højre kolonne kun indeholder T'er, konkluderer vi, at $(P \Rightarrow Q) \Leftrightarrow (\neg Q \Rightarrow \neg P)$ er en tautologi. $\square$

Ligning (1.20) betyder, at en implikation i princippet kan udtrykkes ved hjælp af negation og disjunktion. Ligning (1.21) kaldes *kontraposition*. Dette betyder, at hvis man vil bevise, at Q er en logisk konsekvens af P, er det nok at vise, at $\neg P$ er en logisk konsekvens af $\neg Q$. Lad os se et lille eksempel på kontraposition.

Eksempel 1.3.5

Overvej påstanden, at der for vilkårlige reelle tal x og y gælder, at

$$(x \cdot y = 0) \Rightarrow ((x = 0) \vee (y = 0)).$$

Denne påstand er sand, og for en gangs skyld er målet ikke at bevise den. Vi ønsker blot at undersøge, hvad kontrapositionen af en sådan påstand er.

Først og fremmest er den givne påstand et logisk udsagn af formen $P \Rightarrow Q$, hvor P er ligningen $x \cdot y = 0$ og Q udsagnet $(x = 0) \vee (y = 0)$. Med ord kan implikationen $P \Rightarrow Q$ formuleres således: hvis ligningen $x \cdot y = 0$ gælder for nogle reelle tal x og y, så er $x = 0$ eller $y = 0$.

Hvad er kontrapositionen af dette? Ligning (1.21) fortæller, at kontrapositionen er $\neg Q \Rightarrow \neg P$. Oversat til vores eksempel bliver den søgte kontraposition dermed

$$\neg((x = 0) \vee (y = 0)) \Rightarrow \neg(x \cdot y = 0).$$

Dette kan dog reduceres en smule. Først og fremmest kan $\neg(x \cdot y = 0)$ omskrives til $x \cdot y \neq 0$. Desuden kan vi ved hjælp af Ligning (1.12), en af De Morgans love, omskrive $\neg((x = 0) \vee (y = 0))$ til $\neg(x = 0) \wedge \neg(y = 0)$, hvilket igen kan skrives som $(x \neq 0) \wedge (y \neq 0)$. Derfor kan kontrapositionen af

$$(x \cdot y = 0) \Rightarrow (x = 0) \vee (y = 0)$$

skrives som

$$((x \neq 0) \wedge (y \neq 0)) \Rightarrow (x \cdot y \neq 0).$$

Sagt i ord: kontrapositionen af udsagnet "hvis $x \cdot y = 0$, så er $x = 0$ eller $y = 0$" er simpelthen "hvis $x \neq 0$ og $y \neq 0$, så er $x \cdot y \neq 0$".

Denne sidstnævnte påstand er en sand påstand, da den er logisk ækvivalent med den sande påstand, vi startede med i dette eksempel.

Ligning (1.22) siger, at to logiske udsagn er logisk ækvivalente, netop hvis de er logiske konsekvenser af hinanden. Ganske ofte er det lettere at vise, at $P \Rightarrow Q$ og $Q \Rightarrow P$ er sande hver for sig, end at vise direkte, at $P \Leftrightarrow Q$ er sand. Også Ligning 1.23 benyttes af og til til at bevise logiske udsagn: i stedet for at vise, at P er sandt, antager man, at P er falsk, og prøver derefter at opnå en modstrid. Hvis man opnår en modstrid, kan man konkludere, at $\neg P \Rightarrow \mathbf{F}$ er sand. Ligning (1.23) viser så, at P dermed også er sandt. Denne metode kaldes et modstridsbevis.

I senere kapitler vil vi regelmæssigt benytte Ligningerne (1.21), (1.22) og (1.23), når forskellige matematiske udsagn undersøges. I næste afsnit vil vi desuden vise anvendelser af logik i matematikken.

1.4 Anvendelsen af logik i matematikken

Logik kan hjælpe, når matematiske problemstillinger skal løses, og når det matematiske ræsonnement skal tydeliggøres. I dette afsnit giver vi en række eksempler på dette.

Eksempel 1.4.1

Spørgsmål: Bestem alle reelle tal x, således at $-x \leq 0 \leq x - 1$.

Svar: $-x \leq 0 \leq x - 1$ er faktisk en forenklet udgave af det logiske udsagn

$$-x \leq 0 \quad \wedge \quad 0 \leq x - 1.$$

Den første ulighed er logisk ækvivalent med uligheden $x \geq 0$, mens den anden er ækvivalent med $x \geq 1$. Derfor er et reelt tal x en løsning, hvis og kun hvis

$$x \geq 0 \quad \wedge \quad x \geq 1.$$

Svaret er derfor alle reelle tal x, således at $x \geq 1$.

Eksempel 1.4.2

Spørgsmål: Bestem alle reelle tal x, således at $2|x| = 2x + 1$. Her betegner $|x|$ absolutværdien af x.

Svar: Hvis $x < 0$, så er $|x| = -x$, mens hvis $x \geq 0$, så er $|x| = x$. Derfor er det praktisk at overveje tilfældene $x < 0$ og $x \geq 0$ hver for sig. Mere formelt har vi følgende sekvens af logisk ækvivalente udsagn:

$$\begin{array}{lll}
2|x| = 2x+1 & & \\
\Leftrightarrow & & \\
2|x| = 2x+1 & \wedge & (x<0 \quad \vee \quad x \geq 0) \\
\Leftrightarrow & & \\
(2|x| = 2x+1 \quad \wedge \quad x<0) & \vee & (2|x| = 2x+1 \quad \wedge \quad x \geq 0) \\
\Leftrightarrow & & \\
(-2x = 2x+1 \quad \wedge \quad x<0) & \vee & (2x = 2x+1 \quad \wedge \quad x \geq 0) \\
\Leftrightarrow & & \\
(-4x = 1 \quad \wedge \quad x<0) & \vee & (0 = 1 \quad \wedge \quad x \geq 0) \\
\Leftrightarrow & & \\
(x = -1/4 \quad \wedge \quad x<0) & \vee & (\mathbf{F} \quad \wedge \quad x \geq 0) \\
\Leftrightarrow & & \\
x = -1/4 \quad \vee \quad \mathbf{F} & & \\
\Leftrightarrow & & \\
x = -1/4 & &
\end{array}$$

Den eneste løsning til ligningen $2|x| = 2x+1$ er derfor $x = -1/4$.

Eksempel 1.4.3

Spørgsmål: Bestem alle ikke-negative reelle tal, således at $\sqrt{x} = -x$.

Observation: Det er fristende at opløfte begge sider i anden potens, hvilket ville resultere i $x = x^2$, og derefter konkludere, at $x = 0$ og $x = 1$ er løsningerne til ligningen $\sqrt{x} = -x$. $x = 0$ er ganske rigtigt en løsning, men $x = 1$ er faktisk ikke, da $\sqrt{1} \neq -1$. Hvad gik galt?

Svar: Ræsonnementet ovenfor viser, at hvis x opfylder ligningen $\sqrt{x} = -x$, så er $x = x^2$, hvilket igen medfører, at $x = 0$ eller $x = 1$. Derfor er følgende udsagn helt korrekt:

$$(\sqrt{x} = -x) \Rightarrow (x = 0 \vee x = 1).$$

I den forstand gik intet galt, og enhver løsning til ligningen $\sqrt{x} = -x$ skal ganske rigtigt være enten $x = 0$ eller $x = 1$. Hvad der kan forvirre, er, at dette slet ikke betyder, at $x = 0$ og $x = 1$ begge er løsninger til ligningen $\sqrt{x} = -x$. Dette ville nemlig svare til udsagnet

$$(x = 0 \vee x = 1) \Rightarrow (\sqrt{x} = -x),$$

hvilket ikke er, hvad vi har vist, og faktisk er det ikke sandt. For at løse opgaven, skal vi blot eftertjekke, om de fundne potentielle løsninger $x = 0$ og $x = 1$ virkelig er løsninger. Vi får da, at $x = 0$ er den eneste løsning.

1.5 Epilog: det logiske problem om etiketter og krukker

Lad os vende tilbage til problemet med krukker og etiketter fra første afsnit.

Eksempel 1.5.1

Lad os ved $P_1(A)$ betegne udsagnet, at den venstre krukke har den sande etiket "Æbler". På samme måde kan vi ved $P_1(B)$, henholdsvis $P_1(P)$, betegne udsagnet, at den venstre krukke har den sande etiket "Begge", henholdsvis "Pærer". Vi ved så, at $P_1(B) \vee P_1(P)$ altid er sandt, da den venstre krukke ikke kan have etiketten "Æbler". Ligeledes kan vi for den midterste krukke introducere $P_2(A)$, $P_2(B)$ og $P_2(P)$ som betegnelser for udsagnene, at den midterste krukke har den sande etiket "Æbler", "Begge" eller "Pærer", og konkludere, at $P_2(A) \vee P_2(P)$ er et sandt udsagn. Igen for den højre krukke får vi, at $P_3(A) \vee P_3(B)$ er et sandt udsagn. Vi konkluderer, at

$$(P_1(B) \vee P_1(P)) \wedge (P_2(A) \vee P_2(P)) \wedge (P_3(A) \vee P_3(B)) \tag{1.24}$$

altid er sandt. Ved anvendelse af Ligning (1.7) gentagne gange kan vi omskrive dette til det logisk ækvivalente udsagn

$$\begin{array}{lcll}
(P_1(B) \wedge P_2(A) \wedge P_3(A)) & \vee & (P_1(B) \wedge P_2(A) \wedge P_3(B)) & \vee \\
(P_1(B) \wedge P_2(P) \wedge P_3(A)) & \vee & (P_1(B) \wedge P_2(P) \wedge P_3(B)) & \vee \\
(P_1(P) \wedge P_2(A) \wedge P_3(A)) & \vee & (P_1(P) \wedge P_2(A) \wedge P_3(B)) & \vee \\
(P_1(P) \wedge P_2(P) \wedge P_3(A)) & \vee & (P_1(P) \wedge P_2(P) \wedge P_3(B)). &
\end{array}$$

Dette udsagn er stadig sandt, da det er logisk ækvivalent med udsagnet fra Ligning 1.24. Da vi ved, at hver etiket skal anvendes netop én gang, kan et udsagn som $P_1(B) \wedge P_2(A) \wedge P_3(A)$, hvor den samme etiket optræder to gange, ikke være korrekt, hvilket vil sige, at det er en modstrid. Da disjunktion absorberer modstridigheder, hvilket Ligning (1.16) viser, konkluderer vi derfor, at

$$(P_1(B) \wedge P_2(P) \wedge P_3(A)) \quad \vee \quad (P_1(P) \wedge P_2(A) \wedge P_3(B)) \tag{1.25}$$

nødvendigvis altid er sandt.

Dette viser, at der kun er to mulige korrekte måder at mærke krukkerne på. Dette er ganske nyttig viden, da vi ikke har trukket frugt endnu! Lad os nu undersøge, hvad effekten af at trække fra en krukke er. Hvis vi trækker fra den venstre krukke, lærer vi ikke meget om etiketten på den krukke. Da den sande etiket er "Begge" eller "Pærer", ved vi, hvis vi trækker et æble fra den, at den sande etiket ikke kan være "Pærer", men hvis vi trækker en pære fra den, kunne den sande etiket stadig være "Begge" eller "Pærer". Ligeledes kan et træk fra den højre krukke heller ikke med sikkerhed fastslå krukkens sande etiket. Situationen er anderledes for den midterste krukke. Da den sande etiket på den midterste krukke er "Æbler" eller "Pærer", ved vi, at hvis vi trækker et æble fra den, kan dens sande etiket ikke være "Pærer". I det tilfælde vil etiketten skulle være "Æbler". Ligeledes, hvis vi trækker en pære fra den midterste krukke, er dens sande etiket "Pærer". Vi er dermed kommet frem til følgende løsningsmetode til problemet:

Løsning:

Trin 1: Træk fra den midterste krukke. Da vi ved, at alle etiketter er forkerte, indeholder den midterste krukke, der har etiketten "Begge", enten kun æbler eller kun pærer. Hvis vi trækker et æble fra den midterste krukke, kan vi konkludere, at den korrekte etiket skulle have været "Æbler", mens vi, hvis vi trækker en pære fra den midterste krukke, kan konkludere, at den korrekte etiket skulle have været "Pærer".

Trin 2: Vi ved, at det logiske udsagn i Ligning 1.25 altid er sandt. Dette medfører, at hvis vi i trin 1 fandt den korrekte etiket for den midterste krukke til at være "Æbler", så er $P_1(P) \wedge P_2(A) \wedge P_3(B)$ sandt, mens hvis den korrekte etiket på den midterste krukke blev identificeret som "Pærer" i trin 1, så er $P_1(B) \wedge P_2(P) \wedge P_3(A)$ sandt.

Konklusion: Vi behøver kun at trække én gang! Derefter kan vi placere alle tre etiketter korrekt. Desuden har vi fundet frem til en simpel, trinvis procedure til at bestemme den korrekte mærkning. Dette er et eksempel på, hvad man kalder en algoritme. Vi kan opskrive proceduren som en computer-algoritme som følger:

Algoritme 1 Etiketidentifikator

1: Træk fra krukken mærket "Begge", og betegn resultatet ved R.
2: **if** $R =$ æble **then**
3: Identificér etiketterne på krukkerne som "Pærer","Æbler","Begge",
4: **else**
5: Identificér etiketterne på krukkerne som "Begge","Pærer","Æbler".

Bemærk, at vi i pseudo-kode som her fastholder klassiske engelske betegnelser for kodespecifikke udtryk så som if-then-else-betingelser.

Der findes adskillige gåder af denne type. Her er endnu en. Du er velkommen til at prøve at løse den selv, før du læser løsningen.

Eksempel 1.5.2

En politibetjent undersøger et indbrud og har kunnet indsnævre antallet af mistænkte til tre. Betjenten er helt sikker på, at en af disse tre har begået forbrydelsen, og at gerningsmanden arbejdede alene. Under afhøringerne kommer de tre mistænkte hver især med følgende udtalelser:

Mistænkt1:	"Mistænkt2 gjorde det";
	"Jeg var der ikke";
	"Jeg er uskyldig"
Mistænkt2:	"Mistænkt3 er uskyldig";
	"Alt, hvad Mistænkt1 sagde, er løgn";
	"Jeg gjorde det ikke"
Mistænkt3:	"Jeg gjorde det ikke";
	"Mistænkt1 lyver, hvis han sagde, at han ikke var der";
	"Mistænkt2 lyver, hvis han sagde, at alt, hvad Mistænkt1 sagde, er løgn"

Forvirret går politibetjenten til sin chef, politikommissæren. Politikommissæren siger: "Jeg kender disse mistænkte ret godt, og hver eneste af dem lyver altid mindst én gang i deres udtalelser." Kan du hjælpe politibetjenten med at finde ud af, hvilken mistænkt der er skyldig i indbruddet?

Løsning: Lad os introducere nogle logiske udsagn for at kunne analysere situationen. Først og fremmest er P_1 udsagnet "Mistænkt1 gjorde det", og på samme måde står P_2 for "Mistænkt2 gjorde det" og P_3 for "Mistænkt3 gjorde det". Med denne notation på plads har vi, at

$$P_1 \vee P_2 \vee P_3$$

er sandt, da politibetjenten er helt sikker på, at en af de tre mistænkte begik indbruddet.

Lad os nu analysere udtalelserne fra de mistænkte:

Udtalelse fra Mistænkt1:	
"Mistænkt2 gjorde det";	dette er bare P_2
"Jeg var der ikke";	vi kalder dette R_1
"Jeg er uskyldig";	dette svarer til $\neg P_1$

Lad os tage politikommissærens input med i overvejelserne: alle tre mistænkte har løjet mindst én gang i deres udtalelser. Altså lyver Mistænkt1 om mindst én af sine tre påstande, hvilket betyder, at $\neg P_2 \vee \neg R_1 \vee \neg(\neg P_1)$ er et sandt udsagn. Ved hjælp af Ligning (1.11), konkluderer vi, at også

$$\neg P_2 \vee \neg R_1 \vee P_1$$

er et sandt udsagn.

Udtalelse fra Mistænkt2:	
"Mistænkt3 er uskyldig";	dette er $\neg P_3$
"Alt, hvad Mistænkt1 sagde, er løgn";	dette svarer til $\neg P_2 \wedge \neg R_1 \wedge P_1$
"Jeg gjorde det ikke";	dette er $\neg P_2$

Lad os igen inddrage politikommissærens input. For Mistænkt2 får vi, at $P_3 \vee \neg(\neg P_2 \wedge \neg R_1 \wedge P_1) \vee P_2$ er et sandt udsagn. Man kan reducere dette udtryk ved hjælp af Sætning 1.3.1. Først og fremmest er udsagnet $\neg(\neg P_2 \wedge \neg R_1 \wedge P_1)$ logisk ækvivalent med $\neg(\neg P_2) \vee \neg(\neg R_1) \vee \neg P_1)$, som igen er logisk ækvivalent med $P_2 \vee R_1 \vee \neg P_1$ ifølge Ligning (1.11). Erstattes dette med det oprindelige udsagn, ser vi, at $P_3 \vee (P_2 \vee R_1 \vee \neg P_1) \vee P_2$ er et sandt udsagn. Ved at reducere $P_2 \vee P_2$ til P_2 ved brug af Ligning (1.2) får vi, at også

$$P_3 \vee P_2 \vee R_1 \vee \neg P_1$$

er et sandt udsagn.

Udtalelsen fra Mistænkt3 er lidt indviklet, så før vi samler hans udsagn i en tabel, tager vi et kig på hans to sidstnævnte udsagn. Andennævnte udsagn fra Mistænkt3 er, at "Mistænkt1 lyver, hvis han sagde, at han ikke var der". Med andre ord: "Mistænkt1 var der ikke" $\Rightarrow$ "Mistænkt1 lyver". Men politikommissæren har allerede fortalt os, at udsagnet "Mistænkt1 lyver" altid er sandt. Dette betyder, at implikationen "Mistænkt1 var der ikke " $\Rightarrow$ "Mistænkt1 lyver" er et sandt udsagn. Ligeledes er det tredje udsagn fra Mistænkt3, at "Mistænkt2 lyver, hvis han sagde, at alt, hvad Mistænkt1 sagde, er løgn", sandt. Derfor giver det andet og tredje udsagn fra Mistænkt3 os ikke nogen information, som vi ikke allerede vidste.

Udtalelse fra Mistænkt3:

"Jeg gjorde det ikke"; dette er $\neg P_3$
"Mistænkt1 lyver, hvis han sagde, at han ikke var der";
"Mistænkt2 lyver, hvis han sagde, at alt, hvad Mistænkt1 sagde, er løgn;

Lad os for tredje gang overveje politikommissærens input. Først og fremmest medfører dette input, at andet og tredje udsagn fra Mistænkt3 er sande, da vi ved, at Mistænkt1 og Mistænkt2 lyver. Da Mistænkt3 ifølge samme input fra politikommissæren løj, konkluderer vi, at $\neg P_3$ må være en løgn. Med andre ord, P_3 må være sandt.

Samler vi det hele, har vi nu bestemt, at følgende alle er sande: $P_1 \vee P_2 \vee P_3$, $\neg P_2 \vee \neg R_1 \vee P_1$, $P_3 \vee P_2 \vee R_1 \vee \neg P_1$, P_3. Det faktum, at P_3 er sandt, medfører straks, at den eneste mulighed er, at Mistænkt3 har begået indbruddet, og at Mistænkt1 og Mistænkt2 som følge heraf er uskyldige. Men vi bør stadig kontrollere, om alle de andre udsagn i dette tilfælde rent faktisk er sande. Hvis ikke, ville dette betyde, at der ikke findes nogen løsning, og at politibetjenten eller politikommissæren tager fejl. Først og fremmest, hvis P_3 antager værdien T, så vil $P_1 \vee P_2 \vee P_3$ og $P_3 \vee P_2 \vee R_1 \vee \neg P_1$ være sande jævnfør definitionen af disjunktion. Dette efterlader $\neg P_2 \vee \neg R_1 \vee P_1$. Da Mistænkt2 er uskyldig, antager P_2 værdien F, og som en konsekvens heraf antager $\neg P_2$ værdien T. Derfor er $\neg P_2 \vee \neg R_1 \vee P_1$ et sandt udsagn. Dette betyder, at der ikke er nogen modstrid. Politiet bør anholde Mistænkt3!

Kapitel 2

Mængder og funktioner

2.1 Mængder

Begrebet *mængde* er grundlæggende i matematikken, og derfor vil vi behandle noget terminologi og notation vedrørende mængder i dette afsnit.

Grundlæggende er en mængde A en måde at samle elementer i et "bundt" på, som udgør ét objekt. Hvis vi som et eksempel vil opskrive en mængde bestående af tallene 0 og 1, skriver vi $\{0, 1\}$. Dette er et eksempel på en mængde med to elementer. Elementerne behøver ikke at være tal men kunne i princippet være hvad som helst. Gentagelse af elementer gør ikke en mængde større, forstået sådan at hvis et element optræder to eller flere gange i en mængde, kan alle dets dubletter fjernes. For eksempel har man $\{0, 0, 1\} = \{0, 1\}$ og $\{1, 1, 1, 1\} = \{1\}$. Også den rækkefølge, hvori elementerne er opskrevet i en mængde, er uden betydning. Derfor er for eksempel $\{0, 1\} = \{1, 0\}$.

Nogle mængder af tal bruges så ofte, at der findes en standardnotation for dem:

$\mathbb{N} = \{1, 2, \dots\}$	mængden af *naturlige tal*,
$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$	mængden af *heltal*,
	og
$\mathbb{R}$	mængden af alle *reelle tal*.

At sige, at a tilhører eller er et element i mængden A, udtrykkes som: $a \in A$. Nogle forfattere foretrækker at skrive mængden først og derefter elementet, altså $A \ni a$ i stedet for $a \in A$. Hvis et element a ikke tilhører mængden A, kan man benytte negationen fra udsagnslogikken og skrive $\neg(a \in A)$. Dette kan dog også skrives som $a \notin A$. Hvis to elementer tilhører samme mængde, altså for eksempel $a_1 \in A$ og $a_2 \in A$, vil man typisk skrive $a_1, a_2 \in A$.

Eksempel 2.1.1

Vi har, at $1 \in \mathbb{N}$ og at $-1 \in \mathbb{Z}$, mens $-1 \notin \mathbb{N}$. Derudover har vi, at $\pi \in \mathbb{R}$, mens $\pi \notin \mathbb{Z}$, da $\pi \approx 3.1415$ ikke er et heltal.

En mængde bestemmes af dets elementer, hvilket betyder, at to mængder A og B er ens, $A = B$, hvis og kun hvis de indeholder de samme elementer. Med andre ord er $A = B$, hvis og kun hvis der for alle elementer a gælder, at $a \in A \Leftrightarrow a \in B$. Hvis A og B er mængder, kaldes B en *delmængde* af A, hvis ethvert element i B også er et element i A. Sædvanlig notation for dette er $B \subseteq A$. Med andre ord er udsagnet $B \subseteq A$ pr. definition sandt, hvis og kun hvis udsagnet $a \in B \Rightarrow a \in A$ er sandt for alle elementer a. Naturligvis er $A \subseteq A$, da implikationen $a \in A \Rightarrow a \in A$ er sand for alle a. I stedet for at skrive $B \subseteq A$ kan man også skrive $A \supseteq B$.

Den *tomme mængde* er mængden, der ikke indeholder nogen elementer overhovedet. Den betegnes sædvanligvis ved $\emptyset$, inspireret af bogstavet Ø fra det danske og norske alfabet. Nogle forfattere bruger $\{\}$ for den tomme mængde, men vi vil altid benytte notationen $\emptyset$ for den. Den tomme mængde $\emptyset$ er en delmængde af enhver anden mængde A.

Hvis man ønsker at understrege, at en mængde B er en delmængde af A men ikke lig med hele A, skriver man $B \subsetneq A$, alternativt $A \supsetneq B$. Endelig, hvis man vil udtrykke i en formel, at B ikke er en delmængde af A, kan det logiske negationssymbol $\neg$ benyttes, ved at der skrives $\neg(B \subseteq A)$. Det er dog mere almindeligt at skrive $B \not\subseteq A$, alternativt $A \not\supseteq B$.

Eksempel 2.1.2

Da ethvert naturligt tal er et heltal, har vi $\mathbb{N} \subseteq \mathbb{Z}$. Ethvert heltal $n \in \mathbb{Z}$ er også et reelt tal. Derfor er $\mathbb{Z} \subseteq \mathbb{R}$. Vi har endda $\mathbb{N} \subsetneq \mathbb{Z}$ og $\mathbb{Z} \subsetneq \mathbb{R}$. For at vise $\mathbb{N} \subsetneq \mathbb{Z}$ skal vi blot sikre os, at $\mathbb{N} \subseteq \mathbb{Z}$ (hvilket vi allerede har observeret), samt at $\mathbb{N} \neq \mathbb{Z}$. Da $-1 \in \mathbb{Z}$, mens $-1 \notin \mathbb{N}$, kan vi konkludere, at $\mathbb{N} \neq \mathbb{Z}$. Ligeledes er $\mathbb{Z} \subsetneq \mathbb{R}$, da $\pi \in \mathbb{R}$, mens $\pi \notin \mathbb{Z}$.

En sædvanlig måde at konstruere delmængder af en mængde A på er ved at vælge elementer fra det, for hvilke et vist logisk udtryk er sandt. For notationens skyld vil vi lade $P(a)$ betegne et sådant logisk udtryk. $\{a \in A \mid P(a)\}$ betegner dermed netop den delmængde af A, der består af de elementer $a \in A$, for hvilke det logiske udtryk $P(a)$ er sandt.

Eksempel 2.1.3

Lad $\mathbb{Z}$ som før være mængden af heltal. Så er $\{a \in \mathbb{Z} \mid a \geq 1\}$ mængden $\{1, 2, 3, 4, \ldots\}$, og $\{a \in \mathbb{Z} \mid a \leq 3\} = \{\ldots, -1, 0, 1, 2, 3\}$. Ligeledes er $\{a \in \mathbb{Z} \mid 1 \leq a \leq 3\} = \{1, 2, 3\}$.

Eksempel 2.1.4

Blandt standardnotationerne, hvoraf vi allerede har introduceret $\mathbb{N}$, $\mathbb{Z}$ og $\mathbb{R}$, er endnu et eksempel mængden $\mathbb{Q}$: mængden af alle *rationelle tal*, det vil sige mængden af brøker af heltal. Mere præcist har vi

$$\mathbb{Q} = \left\{ \frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0 \right\}.$$

Dette betyder, at et element i $\mathbb{Q}$ kan skrives på formen a/b, hvor både a og b er heltal, mens b ikke er nul. Bemærk, at brøker som $1/2$ og $2/4$ er ens, da $2/4$ kan reduceres til $1/2$ ved, at både tæller og nævner divideres med 2. Mere generelt er to brøker a/b og c/d ens, hvis og kun hvis $ad = bc$.

Da ethvert heltal $n \in \mathbb{Z}$ kan skrives som $n/1$, har vi, at $\mathbb{Z} \subseteq \mathbb{Q}$. Omvendt ser vi dog, at mens $1/2 \in \mathbb{Q}$, så er $1/2 \notin \mathbb{Z}$, hvorfor vi har $\mathbb{Z} \subsetneq \mathbb{Q}$. Yderligere er enhver brøk af heltal et reelt tal, så $\mathbb{Q} \subseteq \mathbb{R}$. Det viser sig, at $\mathbb{Q} \subsetneq \mathbb{R}$. En måde at indse dette på er at finde et reelt tal, der ikke kan skrives som en brøk af heltal. Et eksempel på et sådant reelt tal er $\sqrt{2}$, men vi vil ikke vise her, hvorfor $\sqrt{2} \notin \mathbb{Q}$.

Hvis to reelle tal a og b er givet, således at $a < b$, kan man definere en type af delmængder i $\mathbb{R}$, der kaldes *intervaller*. Disse er:

$$[a,b] = \{x \in \mathbb{R} \mid a \leq x \leq b\},$$

$$[a,b[= \{x \in \mathbb{R} \mid a \leq x < b\},$$

$$]a,b] = \{x \in \mathbb{R} \mid a < x \leq b\}$$

og

$$]a,b[= \{x \in \mathbb{R} \mid a < x < b\}.$$

Intervaller af formen $[a,b]$ kaldes *lukkede* , mens intervaller af formen $]a,b[$ kaldes *åbne*.

Man vil også typisk definere

$$\mathbb{R}_{\geq a} = \{x \in \mathbb{R} \mid x \geq a\},$$

$$\mathbb{R}_{>a} = \{x \in \mathbb{R} \mid x > a\},$$

$$\mathbb{R}_{\leq a} = \{x \in \mathbb{R} \mid x \leq a\}$$

og

$$\mathbb{R}_{<a} = \{x \in \mathbb{R} \mid x < a\}.$$

Eksempel 2.1.5

Intervallet $]0,1]$ består af alle reelle tal x, der opfylder $0 < x \leq 1$. Dette interval er hverken lukket eller åbent. Mængden $\mathbb{R}_{\geq 0}$ er mængden af alle ikke-negative, reelle tal, mens $\mathbb{R}_{>0}$ er mængden af alle positive, reelle tal. Notationen $\mathbb{R}_+$ benyttes også ofte til at betegne mængden af alle positive, reelle tal.

Det føles intuitivt, at to mængder er ens, hvis og kun hvis de er delmængder af hinanden. Lad os mere præcist forklare, hvorfor dette er sandt, og opskrive det som en lemma.

Lemma 2.1.1

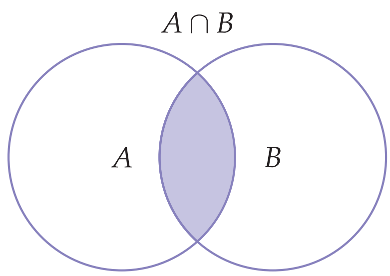
Lad A og B være to mængder. Da er $A = B$, hvis og kun hvis $A \subseteq B$ og $A \supseteq B$.

Bevis. Udsagnet $A = B$ for to mængder A og B er logisk ækvivalent med udsagnet $a \in A \Leftrightarrow a \in B$ for alle a. Ved hjælp af Ligning (1.22) kan vi opdele biimplikationen i to implikationer. Derefter får vi det logisk ækvivalente udsagn $(a \in A \Rightarrow a \in B) \quad \wedge \quad (a \in A \Leftarrow a \in B)$ for alle a. Dette er ækvivalent med udsagnet $A \subseteq B \quad \wedge \quad A \supseteq B$. $\square$

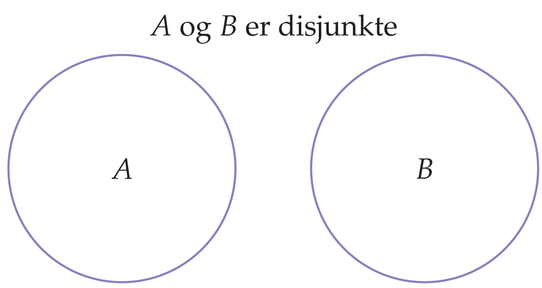
I stedet for $\subseteq$ og $\supseteq$ foretrækker nogle forfattere symbolerne $\subset$ og $\supset$. Men endnu andre forfattere bruger symbolerne $\subset$ og $\supset$ med den ovenfor beskrevne betydning af $\subsetneq$ og $\supsetneq$, inspireret af brugen af $<$ og $>$ til angivelse af skarpe uligheder. For at undgå forvirring vil vi ikke benytte symbolerne $\subset$ eller $\supset$.

Lad os nu definere nogle grundlæggende mængdeoperationer, som vi får brug for senere hen. Disse eksemplificeres i Eksempel 2.1.6. Som den første definerer vi for to mængder A og B *fællesmængden* af A og B, betegnet $A \cap B$, som mængden bestående af alle elementer, der er både i A og i B. Med andre ord:

$$A \cap B = \{a \mid a \in A \wedge a \in B\}. \tag{2.1}$$



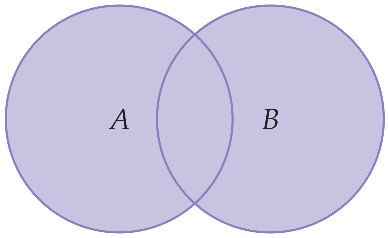
To mængder A og B kaldes *disjunkte*, hvis $A \cap B = \emptyset$.



Foreningsmængden af A og B er defineret som:

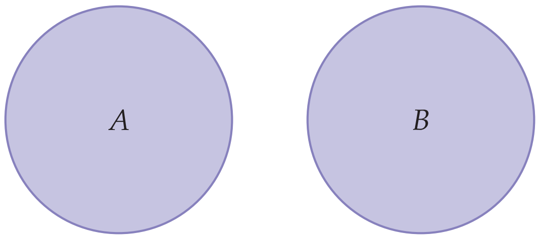
$$A \cup B = \{a \mid a \in A \vee a \in B\}. \tag{2.2}$$

$A \cup B$



Foreningsmængden $A \cup B$ kaldes en *disjunkt forening* af A og B, hvis $A \cap B = \emptyset$.

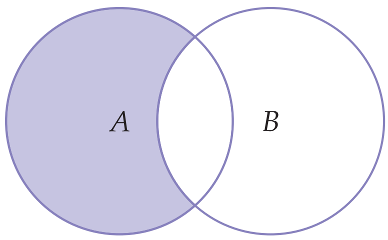
$A \cup B$ er en disjunkt forening af A og B



Mængdedifferencen af A og B, som kan udtales A minus B, defineres ved:

$$A \setminus B = \{a \mid a \in A \wedge a \notin B\}.$$

$A \setminus B$



Afslutningsvis definerer vi *det kartesiske produkt* af A og B som mængden:

$$A \times B = \{(a, b) \mid a \in A \wedge b \in B\}.$$

Det kartesiske produkt af to mængder A og B er simpelthen mængden af alle par (a, b), hvis første koordinat kommer fra A, og hvis anden koordinat kommer fra B. Det kartesiske produkt af en mængde A med sig selv betegnes typisk A^2. Med andre ord: $A^2 = A \times A$.

Vi kommer hovedsageligt til at benytte det kartesiske produkt af to mængder, men definitionen af det kartesiske produkt udvides nemt til at involvere flere mængder end blot to. Man inkluderer blot flere koordinater, én fra hver mængde, i det kartesiske produkt. Som et eksempel er $A \times B \times C = \{(a,b,c) \mid a \in A, b \in B \text{ og } c \in C\}$. Mere generelt, hvis n er et positivt heltal, og $A_1, \ldots, A_n$ er mængder, så er

$$A_1 \times \cdots \times A_n = \{(a_1, \ldots, a_n) \mid a_1 \in A_1, \ldots, a_n \in A_n\}.$$

Hvis alle mængder er ens, for eksempel $A_1 = A, \ldots, A_n = A$, skrives typisk A^n for deres kartesiske produkt. Med andre ord:

$$A^n = \{(a_1, \ldots, a_n) \mid a_1 \in A, \ldots, a_n \in A\}. \tag{2.3}$$

Lad os illustrere de introducerede mængdebegreber i et eksempel.

Eksempel 2.1.6

Lad $1, 2, 3$ og 4 være de første fire positive heltal. Så haves følgende:

(a) $\{1,2\} \subseteq \{1,2,3\}$ og faktisk $\{1,2\} \subsetneq \{1,2,3\}$,

(b) $\{1,2\} \supseteq \{2\}$ og faktisk $\{1,2\} \supsetneq \{2\}$,

(c) $\{1,4\} \not\subseteq \{1,2,3\}$,

(d) $\{1,2,3\} \cap \{2,3,4\} = \{2,3\}$,

(e) $\{1,2\}$ og $\{3\}$ er disjunkte mængder,

(f) $\{1,2,3\} \cup \{2,3,4\} = \{1,2,3,4\}$,

(g) $\{1,2,3,4\}$ er den disjunkte forening af $\{1,2\}$ og $\{3,4\}$,

(h) $\{1,2,3\} \setminus \{2,3,4\} = \{1\}$,

(i) $\{2,3,4\} \setminus \{1,2,3,4\} = \emptyset$,

(j) $\{1,2\} \times \{3,4\} = \{(1,3), (1,4), (2,3), (2,4)\}$,

(k) $\{1,2\}^2 = \{(1,1), (1,2), (2,1), (2,2)\}$.

I Ligningerne (2.1) og (2.2) var de logiske operatorer $\wedge$ og $\vee$ nyttige. I Sætning 1.3.1 oplistede vi forskellige egenskaber for disse to logiske operatorer. Vi kan her benytte dem til at vise tilsvarende egenskaber for fællesmængder og foreningsmængder:

Sætning 2.1.2

Lad A, B og C være mængder. Da haves følgende:

$$A \cap A = A \tag{2.4}$$

$$A \cup A = A \tag{2.5}$$

$$A \cup B = B \cup A \tag{2.6}$$

$$
\begin{aligned}
A \cap B &= B \cap A && (2.7)\\
A \cup (B \cup C) &= (A \cup B) \cup C && (2.8)\\
A \cap (B \cap C) &= (A \cap B) \cap C && (2.9)\\
A \cap (B \cup C) &= (A \cap B) \cup (A \cap C) && (2.10)\\
A \cup (B \cap C) &= (A \cup B) \cap (A \cup C) && (2.11)
\end{aligned}
$$

Bevis. Lad os bevise sidstnævnte, altså Ligning (2.11). Beviset af de resterende overlades til læseren. Ifølge Ligning (2.1) har vi

$$B \cap C = \{a \mid a \in B \wedge a \in C\}.$$

Samtidig har vi ved anvendelse af Ligning (2.2) på mængderne A og $B \cap C$, at

$$A \cup (B \cap C) = \{a \mid a \in A \vee a \in B \cap C\}.$$

Kombineres disse to ligninger og benyttes Ligning 1.8, får vi følgende:

$$
\begin{aligned}
A \cup (B \cap C) &= \{a \mid a \in A \vee (a \in B \wedge a \in C)\}\\
&= \{a \mid (a \in A \vee a \in B) \wedge (a \in A \vee a \in C)\}\\
&= \{a \mid (a \in A \cup B) \wedge (a \in A \cup C)\}\\
&= (A \cup B) \cap (A \cup C).
\end{aligned}
$$

□

Sætning 2.1.2 viser, hvordan udsagnslogikken kan benyttes til omskrivning af fællesmængder og foreningsmængder. Herunder følger et eksempel omhandlende forskellen mellem mængder. Her vil Sætningerne 1.3.2 og 1.3.3 komme til nytte.

Eksempel 2.1.7

Lad A, B og C være tre mængder. I dette eksempel ønsker vi at vise, at $A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C)$. Først og fremmest har vi

$$
\begin{aligned}
A \cap (B \setminus C) &= \{a \mid a \in A \wedge a \in B \setminus C\}\\
&= \{a \mid a \in A \wedge (a \in B \wedge \neg(a \in C))\}.
\end{aligned}
$$

Vi har dog også, at

$$
\begin{aligned}
(A \cap B) \setminus (A \cap C) &= \{a \mid a \in A \cap B \wedge \neg(a \in A \cap C)\} \\
&= \{a \mid (a \in A \wedge a \in B) \wedge \neg(a \in A \wedge a \in C)\} \\
&= \{a \mid (a \in A \wedge a \in B) \wedge (\neg(a \in A) \vee \neg(a \in C))\} \\
&= \{a \mid (a \in A \wedge a \in B) \wedge \neg(a \in A) \quad \vee \quad (a \in A \wedge a \in B) \wedge \neg(a \in C)\} \\
&= \{a \mid \mathbf{F} \quad \vee \quad (a \in A \wedge a \in B) \wedge \neg(a \in C)\} \\
&= \{a \mid (a \in A \wedge a \in B) \wedge \neg(a \in C)\} \\
&= \{a \mid a \in A \wedge (a \in B \wedge \neg(a \in C))\}.
\end{aligned}
$$

Vi kan konkludere, at $A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C)$ er sandt.

2.2 Funktioner

Et særdeles vigtigt begreb i matematikken er *funktionen*. En funktion f fra A til B, hvor A og B er to givne mængder, tildeler til ethvert $a \in A$ et element $b \in B$. I stedet for formuleringen "f tildeler til a et element b" vil man oftest blot sige "f afbilder a til b". Derfor kaldes en funktion ofte også for en *afbildning*. Et alternativ til "f afbilder a til b" er at sige "f evalueret i a er lig med b".

Mængden A kaldes funktionens *definitionsmængde*, mens mængden B kaldes dens *dispositionsmængde*. Bemærk, at begreberne på engelsk er noget anderledeslydende, nemlig *domain*, henholdsvis *co-domain*. Man benytter typisk den kompakte notationsform $f : A \to B$. Værdien af en funktion f i et specifikt element a betegnes $f(a)$. $f(a)$ kaldes billedet af a ved f, alternativt evalueringen af f i a. At f afbilder værdien a fra A til $f(a)$ kan skrives kort med notationen $a \mapsto f(a)$. Al denne information om en funktion f kan kompakt skrives som følger:

$$
\begin{aligned}
f : A &\to B \\
a &\mapsto f(a)
\end{aligned}
$$

For eksempel kan funktionen, der opløfter et reelt tal til anden potens, opskrives som:

$$
\begin{aligned}
f : \mathbb{R} &\to \mathbb{R} \\
x &\mapsto x^2
\end{aligned}
$$

En funktion som ovenstående gives ofte som $f : \mathbb{R} \to \mathbb{R}$, hvor $f(x) = x^2$. Ofte siger man dog blot, at funktionen er defineret som $f(x) = x^2$. I sådanne tilfælde overlades det til læseren at finde ud af, hvad definitionsmængden og dispositionsmængden for funktionen er. Så vidt det er muligt, vil vi altid tydeligt angive definitionsmængde og dispositionsmængde for funktioner. Hvis definitionsmængden og dispositionsmængden vælges til at være samme mængde A, kan man definere *identitetsfunktionen* id_A på A. Dette er funktionen $\mathrm{id}_A : A \to A$, således at $a \mapsto a$.

Billedet, også kaldet *værdimængden*, af en funktion $f : A \to B$ er en vigtig betegnelse, der defineres som mængden $\{f(a) \mid a \in A\}$. Billedet af en funktion $f : A \to B$ er en delmængde af dens dispositionsmængde B, og som Eksempel 2.2.1 vil illustrere, er billede og dispositionsmængde ikke nødvendigvis det samme. Typisk notation for billedet af en funktion $f : A \to B$ er $f(A)$ eller $\mathrm{image}(f)$. Lad os se på nogle eksempler.

Eksempel 2.2.1

Lad os igen arbejde på funktionen

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^2 \end{aligned}$$

Denne funktion har definitionsmængde $\mathbb{R}$ og dispositionsmængde $\mathbb{R}$. Vi påstår nu, at $f(\mathbb{R}) = \{r \in \mathbb{R} \mid r \geq 0\}$. Med andre ord påstår vi, at $f(\mathbb{R}) = \mathbb{R}_{\geq 0}$. På grund af Lemma 2.1.1 er det tilstrækkeligt at vise, at $f(\mathbb{R}) \subseteq \mathbb{R}_{\geq 0}$ og $\mathbb{R}_{\geq 0} \subseteq f(\mathbb{R})$.

Bemærk først og fremmest, at $f(\mathbb{R}) \subseteq \mathbb{R}_{\geq 0}$ er sandt, da et reelt tal opløftet i anden potens ikke kan være negativt. Omvendt, hvis $r \in \mathbb{R}_{\geq 0}$, så er $\sqrt{r}$ defineret, og $r = (\sqrt{r})^2 = f(\sqrt{r})$. Dette viser, at ethvert ikke-negativt, reelt tal r tilhører billedet af f. Med andre ord har vi vist, at $\mathbb{R}_{\geq 0} \subseteq f(\mathbb{R})$. Ved hjælp af Lemma 2.1.1 kan vi konkludere, at $f(\mathbb{R}) = \mathbb{R}_{\geq 0}$.

Dette eksempel viser, at billedet af en funktion ikke nødvendigvis er lig med dens dispositionsmængde.

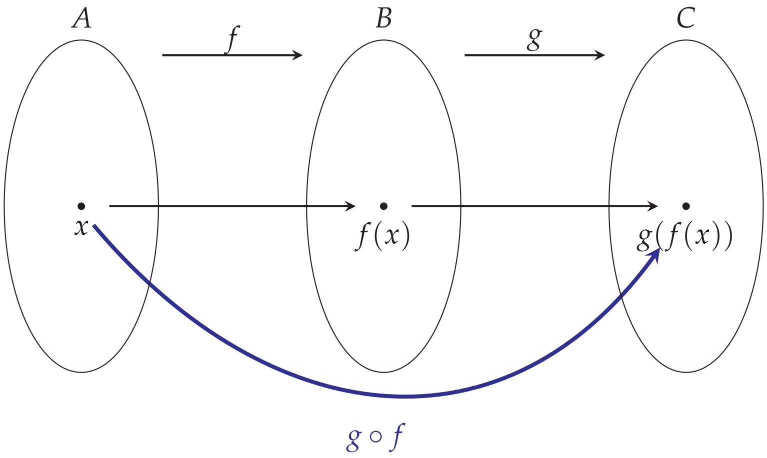
Vi kunne selvfølgelig fra start af have defineret kvadreringsfunktionen, som vi netop har arbejdet på, som $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ med $x \mapsto x^2$. Her er den eneste forskel, at dispositionsmængden er ændret fra $\mathbb{R}$ til $\mathbb{R}_{\geq 0}$. For denne modificerede funktion er billedet det samme som dispositionsmængden, så hvorfor skelner vi mellem billedet og dispositionsmængden for en funktion i den generelle teori? En grund er, at det er praktisk ikke at skulle holde styr på billedet af en funktion hele tiden. Hvis vi ved, at en funktion afbilder reelle tal til reelle tal, kan vi simpelthen sætte dispositionsmængden lig med $\mathbb{R}$ uden at bekymre os yderligere. For komplicerede funktioner kan det endda være meget svært at beregne billedet.

To funktioner $f : A \to B$ og $g : C \to D$ er ens, netop hvis de har samme definitionsmængde, har samme dispositionsmængde, og tildeler de samme værdier til hvert element i deres definitionsmængder. Opskrevet på formel:

$$f = g \iff A = C \quad \wedge \quad B = D \quad \wedge \quad f(a) = g(a) \text{ for alle } a \in A.$$

Eksempel 2.2.2

Betragt funktionerne



Figur 2.1: Sammensætning af funktionerne $f : A \to B$ og $g : B \to C$

$$\begin{aligned} f : \{0,1\} &\to \{0,1\} \\ a &\mapsto a \end{aligned}$$

og

$$\begin{aligned} g : \{0,1\} &\to \{0,1\} \\ a &\mapsto a^2 \end{aligned}$$

Funktionerne f og g har samme definitionsmængde og dispositionsmængde. Desuden er $f(0) = 0$, $f(1) = 1$, mens $g(0) = 0^2 = 0$ og $g(1) = 1^2 = 1$. Derfor er $f = g$.

Dette eksempel viser, at to funktioner kan være ens, selvom de har forskellige forskrifter.

Hvis to funktioner $f : A \to B$ og $g : B \to C$ er givet, er det muligt at definere funktionen

$$\begin{aligned} h : A &\to C \\ a &\mapsto g(f(a)) \end{aligned}$$

Dispositionsmængden for funktionen f og definitionsmængden for funktionen g skal være de samme i denne definition for at garantere, at $g(f(a))$ altid er defineret: for ethvert $a \in A$ ved vi, at $f(a) \in B$, så det er muligt at benytte elementerne $f(a)$ som input i funktionen g, da definitionsmængden for g er defineret til at være B.

Funktionen $h : A \to C$, der opnås på denne måde, betegnes typisk $g \circ f$ (som udtales g *efter* f) og kaldes sammensætningen af g og f. Derfor har vi $(g \circ f)(a) = g(f(a))$.

Eksempel 2.2.3

Lad $\mathbb{R}_{>0}$ betegne mængden af alle positive, reelle tal. Antag, at $f : \mathbb{R} \to \mathbb{R}_{>0}$ er defineret ved

$f(x) = x^2 + 1$, og at $g : \mathbb{R}_{>0} \to \mathbb{R}$ er defineret ved $g(x) = \log_{10}(x)$, hvor $\log_{10}$ betegner logaritmen med grundtal 10. Så er $g \circ f : \mathbb{R} \to \mathbb{R}$ funktionen, der sender $x \in \mathbb{R}$ til $\log_{10}(x^2 + 1)$. Med andre ord:

$$\begin{aligned} g \circ f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \log_{10}(x^2+1) \end{aligned}$$

For eksempel er $(g \circ f)(3) = \log_{10}(3^2 + 1) = \log_{10}(10) = 1$.

Lemma 2.2.1

Lad A, B, C og D være mængder, og antag, at vi er givet funktionerne $h : A \to B$, $g : B \to C$ og $f : C \to D$. Da har vi $(f \circ g) \circ h = f \circ (g \circ h)$.

Bevis. Bemærk som det første, at både $(f \circ g) \circ h$ og $f \circ (g \circ h)$ er funktioner fra A til D, så de har samme definitionsmængde og dispositionsmængde. For at bevise lemmaet er det derfor nok at vise, at vi for alle $a \in A$ har $((f \circ g) \circ h)(a) = (f \circ (g \circ h))(a)$. Fra definitionen af sammensætningen $\circ$, har vi

$$(f \circ (g \circ h))(a) = f((g \circ h)(a)) = f(g(h(a))),$$

mens

$$((f \circ g) \circ h)(a) = (f \circ g)(h(a)) = f(g(h(a))).$$

Vi konkluderer, at der for ethvert $a \in A$ gælder, at $(f \circ (g \circ h))(a) = ((f \circ g) \circ h)(a)$, hvilket er det, der skulle vises. $\square$

Resultatet af dette lemma udtrykkes normalt således: sammensætningen af funktioner er en *associativ* operation. Grundet Lemma 2.2.1 vil man typisk forenkle formler, der involverer sammensætningen af flere funktioner, ved at udelade parenteserne. For eksempel skriver man blot $f \circ g \circ h$, når man bestemmer sammensætningen af tre funktioner.

For en given funktion $f : A \to B$ siges funktionen f at være *injektiv*, netop når to forskellige elementer fra A afbildes til forskellige elementer i B. Skrevet med logiske udtryk betyder dette, at:

$f : A \to B$ er injektiv, hvis og kun hvis for alle $a_1, a_2 \in A$, $(a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2))$.

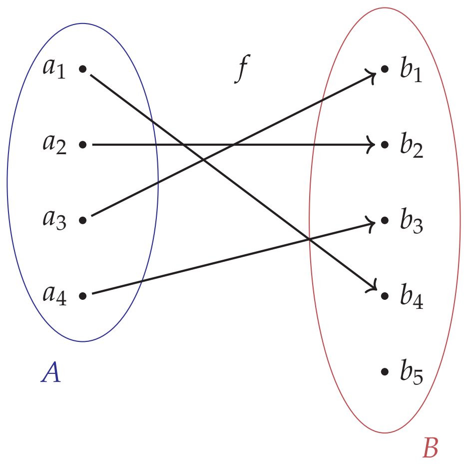
Ligning (1.21) viser, at det er logisk ækvivalent at skrive:

$f : A \to B$ er injektiv, hvis og kun hvis for alle $a_1, a_2 \in A$, $(f(a_1) = f(a_2) \Rightarrow a_1 = a_2)$.

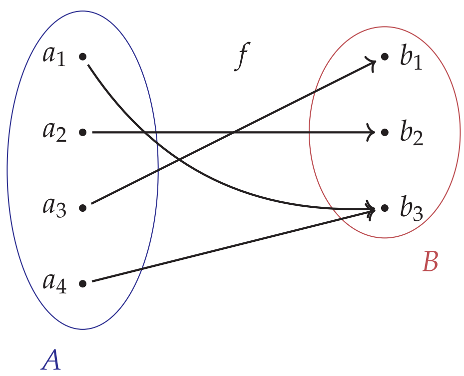
Denne omskrivning kan være praktisk.

En funktion $f : A \to B$ kaldes *surjektiv* , netop når ethvert element i B tilhører billedet af f, det vil sige:

$f : A \to B$ er surjektiv, hvis og kun hvis der for alle $b \in B$ findes et $a \in A$, således at $b = f(a)$.



Figur 2.2: Injektiv funktion $f : A \to B$



Figur 2.3: Surjektiv funktion $f : A \to B$

Ved at benytte notationen $f(A)$ for billedet af f, kan dette kompakt omformuleres til: en funktion $f : A \to B$ kaldes surjektiv, netop når $f(A) = B$.

Eksempel 2.2.4

Et eksempel på en funktion, der er injektiv men ikke surjektiv, er $f : \mathbb{R}\backslash\{0\} \to \mathbb{R}$ givet ved $f(x) = 1/x$. Denne funktion er ikke surjektiv, da dens billede er $\mathbb{R}\backslash\{0\}$, mens dens dispositionsmængde er $\mathbb{R}$. Den er dog injektiv, da vi ser, at hvis $f(a) = f(b)$, altså hvis $1/a = 1/b$, så er $a = b$.

Et eksempel på en funktion, der er surjektiv men ikke injektiv, er $g : \mathbb{R} \to [-1, 1]$ givet ved $g(x) = \sin(x)$. Denne funktion er ikke injektiv, da for eksempel 0 og π indsat i sinusfunktionen begge afbildes i 0.

En funktion $f : A \to B$ kaldes *bijektiv*, hvis den er både injektiv og surjektiv. En bijektiv funktion kaldes også en *bijektion*. Ved at kombinere definitionerne af injektiv og surjektiv ser vi, at funktionen $f : A \to B$ er bijektiv, netop når der for ethvert $b \in B$ findes et unikt $a \in A$, således at $f(a) = b$. I næste afsnit vil vi se flere eksempler på funktioner, men lad os give et eksempel her med det samme.

Eksempel 2.2.5

Betragt funktionen $h : \{0, 1, 2\} \to \{3, 4, 5\}$ givet ved $h(x) = 5 - x$. Vi har, at $h(0) = 5$, $h(1) = 4$, og $h(2) = 3$. Derfor findes der for ethvert $b \in \{3, 4, 5\}$ et unikt $a \in \{0, 1, 2\}$, således at $h(a) = b$. Vi konkluderer, at h er en bijektiv funktion.

Der findes en ganske praktisk forbindelse mellem bijektive funktioner og det, der kaldes inverse funktioner. Lad os for fuldstændighedens skyld først definere, hvad den inverse af en funktion er.

Definition 2.2.1

Lad $f : A \to B$ være en funktion. En funktion $g : B \to A$ kaldes den *inverse funktion* af f, hvis $f \circ g = \mathrm{id}_B$ (identitetsfunktionen på B) og $g \circ f = \mathrm{id}_A$ (identitetsfunktionen på A). Den inverse af f betegnes f^{-1}.

Det viser sig, at en funktion har en invers, netop når den er en bijektiv funktion.

Lemma 2.2.2

Antag, at A og B er mængder, og lad $f : A \to B$ være en funktion. Da er f bijektiv, hvis og kun hvis f har en invers funktion.

Bevis. Antag, at $f : A \to B$ er en bijektion. Som vi har set, er en funktion $f : A \to B$ bijektiv, netop hvis der for ethvert $b \in B$ findes et unikt $a \in A$, således at $f(a) = b$. At a skal være unikt medfører, at vi kan definere en funktion $g : B \to A$, hvor $b \mapsto a$. Vi vil vise, at g er den inverse funktion af f. Hvis $b = f(a)$, har vi

$$(f \circ g)(b) = f(g(b)) = f(a) = b \text{ og } (g \circ f)(a) = g(f(a)) = g(b) = a.$$

Men dette viser, at $f \circ g = \mathrm{id}_B$ og $g \circ f = \mathrm{id}_A$, hvilket ifølge Definition 2.2.1 betyder, at $g = f^{-1}$.

Omvendt, hvis f har en invers funktion, så medfører ligningen $f(a) = b$, at $f^{-1}(f(a)) = f^{-1}(b)$. Da $a = (f^{-1} \circ f)(a) = f^{-1}(f(a))$, ser vi, at $a = f^{-1}(b)$. Derfor findes der for ethvert $b \in B$ et unikt element $a \in A$, således at $f(a) = b$ (nemlig $a = f^{-1}(b)$). Dette viser, at f er bijektiv. □

Eksempel 2.2.6

Lad os igen betragte funktionen $h : \{0, 1, 2\} \to \{3, 4, 5\}$ givet ved $h(x) = 5 - x$ fra Eksempel 2.2.5. Vi så, at funktionen h er bijektiv. Derfor har den ifølge Lemma 2.2.2 en invers $h^{-1} : \{3, 4, 5\} \to \{0, 1, 2\}$. Vi så også, at $h(0) = 5$, $h(1) = 4$, og $h(2) = 3$. Den inverse af h sender billederne tilbage til de oprindelige værdier: $h^{-1}(5) = 0$, $h^{-1}(4) = 1$ og $h^{-1}(3) = 2$.

Bemærk, at de tidligere beregninger faktisk viser, at $h^{-1}(x) = 5 - x$ for alle $x \in \{3, 4, 5\}$. Derfor er $h^{-1} : \{3, 4, 5\} \to \{0, 1, 2\}$ givet ved $h^{-1}(x) = 5 - x$. En lille advarsel: den inverse af en funktion behøver ikke at ligne funktionen selv. Senere vil vi se eksempler på inverse funktioner, hvor dette ikke er tilfældet.

Beregningstekniske aspekter af funktioner

Med vores måde at opfatte en funktion $f : A \to B$ på har vi fuldstændigt ignoreret mere praktiske aspekter så som: hvis du er givet en værdi $a \in A$, hvordan beregner du så rent faktisk $f(a)$? I generel matematisk funktionsteori er dette ikke et problem, og de "indre mekanismer" af funktionen f behandles blot som en sort boks. Men i anvendelser af teorien er det være naturligvis vigtigt at vide, hvordan man beregner funktionsværdier.

Heldigvis kan mange nyttige funktionsberegninger udføres ved hjælp af en *algoritme*. Vi vil ikke gå ind i de præcise detaljer om, hvordan man definerer, hvad en algoritme er, men blot tage udgangspunkt i et intuitivt synspunkt. Grundlæggende er en algoritme et sæt af instruktioner, som man nemt kunne omdanne til et computerprogram, hvis man ville. Disse enkle instruktioner involverer "simple" operationer som multiplikation og addition. Desuden kan mellemliggende resultater gemmes i hukommelsen og benyttes senere i algoritmen, hvis det er nødvendigt. Mere filosofisk åbner algoritmen, som beskriver en funktion f, for den sorte boks og viser dens "indre mekanismer". Lad os se på eksemplet med funktionen $f : \mathbb{R} \to \mathbb{R}$ defineret ved $f(x) = x^3$. Et første forsøg på at beskrive en algoritme, der givet et x beregner $f(x)$, kunne være:

Trin 1. Beregn $x \cdot x$, og husk resultatet af denne beregning.

Trin 2. Tag resultatet fra Trin 1, og multiplicér det med x.

Trin 3. Returnér værdien fra Trin 2.

En smule mere formelt kan vi omskrive dette som:

Trin 0. Betegn det givne input ved x.

Trin 1. Beregn $x \cdot x$, og gem resultatet under navnet y.

Trin 2. Beregn $x \cdot y$, og gem resultatet under navnet z.

Trin 3. Returnér z.

Denne beskrivelse kommer til at ligne en egentlig computeralgoritme, når vi opskriver den som såkaldt *pseudo-kode*. Den primære forskel fra ovenstående beskrivelse er, at en sætning som "Beregn $x \cdot x$, og gem resultatet under navnet y" på kompakt form kan skrives som "$y \leftarrow x \cdot x$". Den algoritmiske pseudo-kodebeskrivelse af funktionen $f : \mathbb{R} \to \mathbb{R}$ defineret ved $f(x) = x^3$ bliver da:

Algoritme 2 for $f : \mathbb{R} \to \mathbb{R}$ defineret ved $f(x) = x^3$

```
Input: x ∈ ℝ
y ← x · x
z ← x · y
return z
```

Lad os se på endnu et eksempel.

Eksempel 2.2.7

Lad $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ være defineret ved $x \mapsto |x|$. Her betegner $|x|$ absolutværdien af x. Som vi så i Eksempel 1.4.2, har vi, at hvis $x < 0$, så er $|x| = -x$, mens hvis $x \geq 0$, så er $|x| = x$. Af denne grund defineres absolutværdien ofte på følgende måde:

$$|x| = \begin{cases} -x & \text{hvis } x < 0, \\ x & \text{ellers.} \end{cases}$$

Når man definerer en funktion stykkevist som her, er det vigtigt at sikre sig: 1) at alle elementer i funktionens definitionsmængde optræder i et af stykkerne, og 2) at et element i funktionens definitionsmængde ikke optræder i mere end ét af stykkerne. Her er funktionens definitionsmængde $\mathbb{R}$, hvilket også er foreningmængden af $\mathbb{R}_{<0}$ og $\mathbb{R}_{\geq 0}$, så 1) er opfyldt. Desuden er $\mathbb{R}_{<0}$ og $\mathbb{R}_{\geq 0}$ disjunkte mængder, så 2) er opfyldt. Med andre ord: 1) og 2) er opfyldt, fordi funktionens definitionsmængde $\mathbb{R}$ er den disjunkte forening af $\mathbb{R}_{<0}$ og $\mathbb{R}_{\geq 0}$. Den givne beskrivelse af funktionen, der udregner absolutværdien, kan let omskrives som en algoritme i pseudo-kode:

Algoritme 3 til udregning af $|x|$ for $x \in \mathbb{R}$

```
Input: x ∈ ℝ
if x < 0 then
    return −x
else
    return x
```

2.3 Eksempler på funktioner

Lad os eksemplificere teorien om funktioner som udviklet ovenfor ved at tage et nærmere kig på nogle elementære funktioner $f : A \to B$, hvor A og B er delmængder af $\mathbb{R}$. For at gøre det nemmere at vise injektivitet af sådanne funktioner benytter vi os af følgende lemma:

Lemma 2.3.1

Lad $f : A \to B$ være en funktion, og antag, at A og B er delmængder af $\mathbb{R}$. Antag enten, at

$$\text{det for alle } a_1, a_2 \in A \text{ gælder, at: } a_1 < a_2 \Rightarrow f(a_1) < f(a_2) \tag{2.12}$$

eller, at

$$\text{det for alle } a_1, a_2 \in A \text{ gælder, at: } a_1 < a_2 \Rightarrow f(a_1) > f(a_2). \tag{2.13}$$

Så er f en injektiv funktion.

Bevis. Antag at funktionen f opfylder Ligning (2.12). Lad a_1 og a_2 være forskellige elementer i A. Da $a_1 \neq a_2$, ved vi, at vi har enten $a_1 < a_2$ eller $a_2 < a_1$. Hvis $a_1 < a_2$, medfører Ligning (2.12), at $f(a_1) < f(a_2)$. Hvis $a_2 < a_1$, medfører Ligning (2.12), at $f(a_2) < f(a_1)$. I begge tilfælde kan vi konkludere, at $f(a_1) \neq f(a_2)$. Derfor er f injektiv. Hvis funktionen f opfylder Ligning (2.13), viser et lignende ræsonnement, at f er injektiv. □

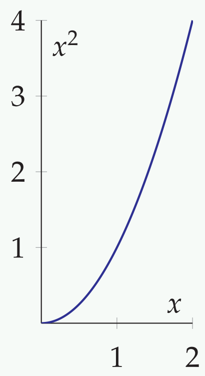
En funktion f, der opfylder Ligning (2.12) eller Ligning (2.13), kaldes *strengt monoton*. Mere præcist kaldes en funktion f, der opfylder Ligning (2.12), *strengt voksende*, mens en funktion f, der opfylder Ligning (2.13), kaldes *strengt aftagende*. Derfor kan Lemma 2.3.1 opsummeres som: en strengt monoton funktion er injektiv.

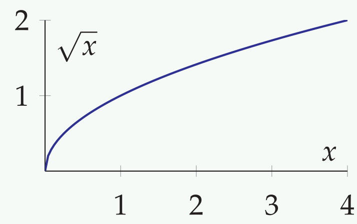
Eksempel 2.3.1

Overvej funktionen $f : \mathbb{R} \to \mathbb{R}$, hvor $f(x) = x^2$. Vi har allerede set i Eksempel 2.2.1, at billedet af denne funktion er lig med $\mathbb{R}_{\geq 0}$. Med andre ord, $f(\mathbb{R}) = \mathbb{R}_{\geq 0}$. Funktionen f er derfor ikke surjektiv. Faktisk er den heller ikke injektiv, da for eksempel $f(-1) = 1$, og $f(1) = 1$.

Da funktionen f ikke er bijektiv, har den ikke en invers. Vi kan forsøge at ændre definitionsmængden og dispositionsmængden for f, så vi opnår en ny funktion, der er bijektiv. Først og fremmest kan vi opstille en funktion $g : \mathbb{R} \to \mathbb{R}_{\geq 0}$ defineret ved $g(x) = x^2$. Forskellen mellem funktionerne f og g er ikke stor: kun deres dispositionsmængder er forskellige. Selvom det for ethvert reelt tal x er sandt, at $f(x) = g(x)$, betragter vi derfor alligevel funktionerne f og g som to forskellige funktioner. Vi introducerer funktionen g af den grund, at g er surjektiv, da $g(\mathbb{R}) = \mathbb{R}_{\geq 0}$, og $\mathbb{R}_{\geq 0}$ er dispositionsmængden for g. Men g har stadig ikke en invers, da g ikke er injektiv, ligesom at f ikke er injektiv af denne grund. Vi har stadig for eksempel, at $g(1) = 1$ og $g(-1) = 1$. Vi vil nu introducere endnu en funktion $h : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ defineret ved $h(x) = x^2$. Funktionen h har samme dispositionsmængde som funktionen g, men bemærk, at definitionsmængden for funktionen h er en delmængde af definitionsmængden for g. Faktisk har h definitionsmængden $\mathbb{R}_{\geq 0}$, hvilket er en skarp delmængde af $\mathbb{R}$, som er definitionsmængden for g. Nu kan man vise, at funktionen h er strengt monoton og derfor jævnfør Lemma 2.3.1 injektiv. Vi har allerede set, at h er surjektiv, så vi kan konkludere, at den er bijektiv. Jævnfør Lemma 2.2.2 har funktionen h derfor en invers. Da der for ethvert $x \in \mathbb{R}_{\geq 0}$ gælder, at $\sqrt{x^2} = x$ og $(\sqrt{x})^2 = x$, er den inverse af h funktionen $h^{-1} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ defineret ved $h^{-1}(x) = \sqrt{x}$.

For at illustrere situationen har vi plottet (dele af) graferne for funktionen h samt dens inverse h^{-1}. Bemærk, at grafen for h^{-1} er spejlbilledet af grafen for h henover linjen $y = x$. Grafen for h illustrerer desuden, at der er tale om en strengt voksende funktion.

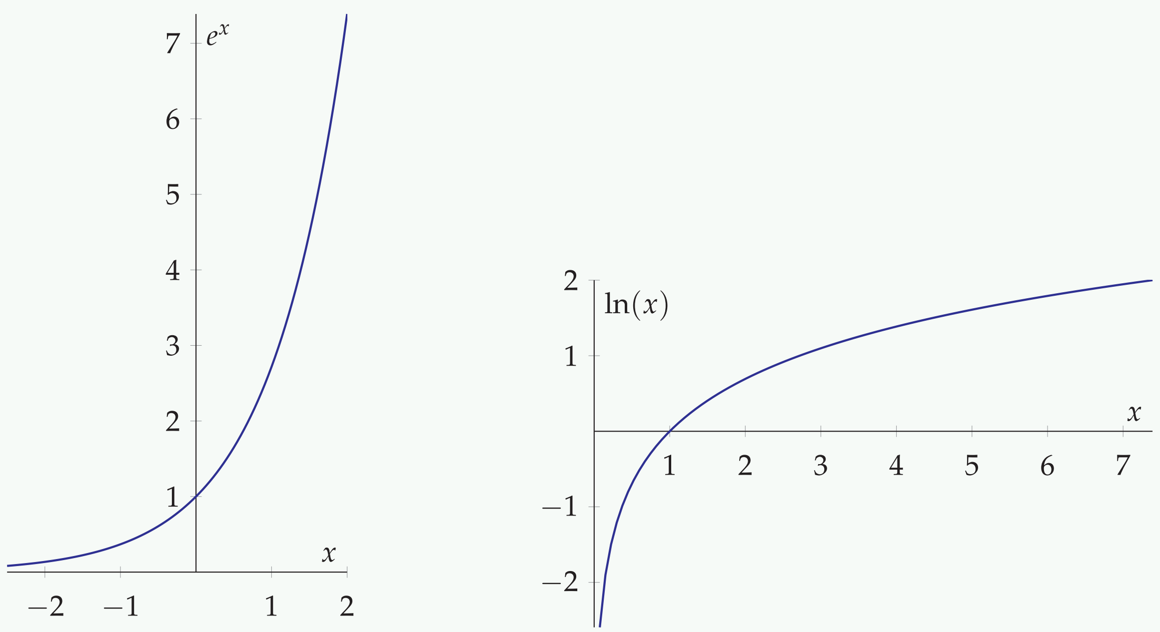




Eksempel 2.3.2

Lad e betegne grundtallet for den naturlige logaritme. Konstanten e kaldes ofte Eulers tal og er omtrent lig med 2.71828. Eksponentialfunktionen $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ er defineret ved $x \mapsto e^x$. Denne er en strengt voksende funktion og er derfor injektiv. Desuden er billedet af eksponentialfunktionen $\mathbb{R}_{>0}$, hvilket medfører, at den er surjektiv. Ved at kombinere disse observationer konkluderer vi, at exp er en bijektiv funktion. Dens inverse betegnes $\ln : \mathbb{R}_{>0} \to \mathbb{R}$. Særligt har vi, at $\ln(e^x) = x$ for alle $x \in \mathbb{R}$, samt $e^{\ln(x)} = x$ for alle $x \in \mathbb{R}_{>0}$.

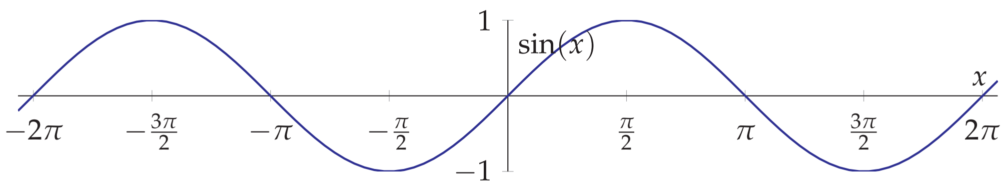
Vi plotter graferne for funktionerne exp og ln for at illustrere situationen.



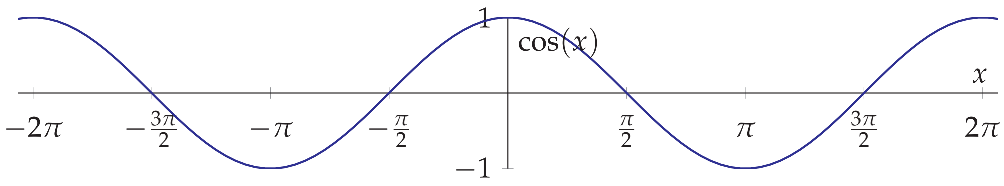
De trigonometriske funktioner sin, cos og tan.

De *trigonometriske funktioner* sinus, cosinus og tangens er ekstremt nyttige eksempler på funktioner og vil dukke op i mange forskellige sammenhænge senere hen. Derfor genbesøger vi dem kort i dette afsnit.

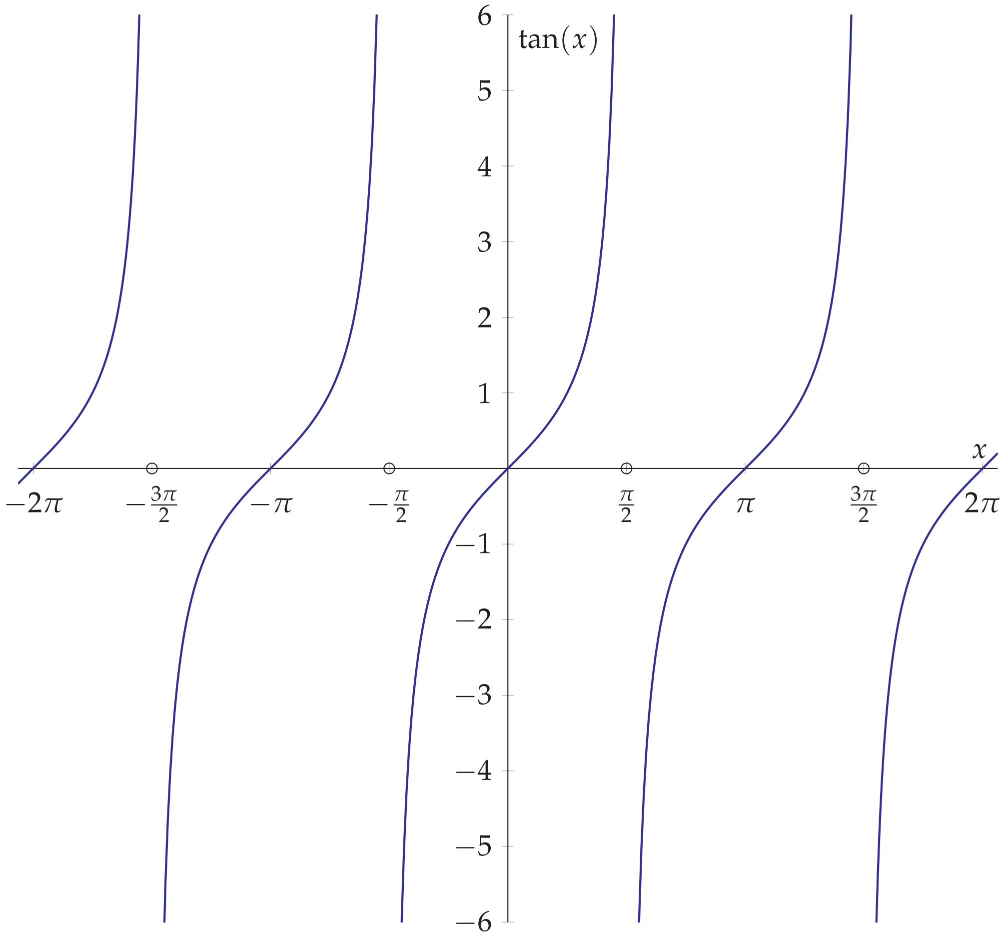
Sinusfunktionen betegnes normalt sin, og lad os i lyset af vores tilgang til funktionsdefinitioner i dette kapitel udspecificere dens definitionsmængde og dispositionsmængde. Vi definerer sinusfunktionen $\sin : \mathbb{R} \to [-1, 1]$ som den funktion, der opfylder, at $x \mapsto \sin(x)$. Billedet af sin er $[-1, 1]$, hvilket betyder, at sin er en surjektiv funktion. Den er ikke en injektiv funktion, da forskellige reelle tal kan have samme funktionsværdi. For eksempel har man $\sin(0) = \sin(\pi) = 0$. Grafen for sinusfunktionen er som følger:



Ligeledes definerer vi $\cos : \mathbb{R} \to [-1, 1]$. Igen er dispositionsmængden valgt som det lukkede interval $[-1, 1]$, hvilket betyder, at funktionen cos er surjektiv. Den er dog ikke injektiv, da for eksempel $\cos(-\pi/2) = \cos(\pi/2) = 0$. Grafen for cosinusfunktionen er:



En tredje ofte benyttet trigonometrisk funktion er tangensfunktionen. Løst sagt har vi sammenhængen $\tan(x) = \sin(x)/\cos(x)$, men denne formel giver kun mening for $x \in \mathbb{R}$, der opfylder $\cos(x) \neq 0$. Derfor kan vi definere $\tan : \{x \in \mathbb{R} \mid \cos(x) \neq 0\} \to \mathbb{R}$, hvor $\tan(x) = \sin(x)/\cos(x)$. Da $\{x \in \mathbb{R} \mid \cos(x) \neq 0\} = \mathbb{R} \setminus \{x \in \mathbb{R} \mid \cos(x) = 0\}$ og $\{x \in \mathbb{R} \mid \cos(x) = 0\} = \{\dots, -3\pi/2, -\pi/2, \pi/2, 3\pi/2, \dots\}$, er definitionsmængden for tangensfunktionen mængden $\mathbb{R} \setminus \{\dots, -3\pi/2, -\pi/2, \pi/2, 3\pi/2, \dots\}$. Grafen for tangensfunktionen er som følger:

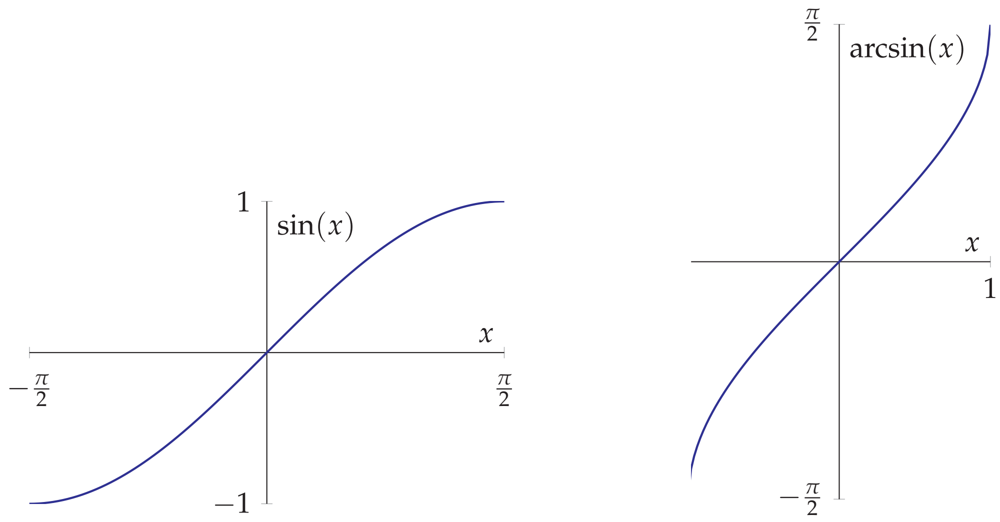


De små cirkler på x-aksen indikerer de værdier af x, for hvilke tangensfunktionen ikke er defineret. Tangensfunktionen er surjektiv, da dens billede er $\mathbb{R}$. Ligesom sinus- og cosinusfunktionerne er den ikke injektiv. Vi har for eksempel $\tan(0) = 0$ og også $\tan(\pi) = 0$.

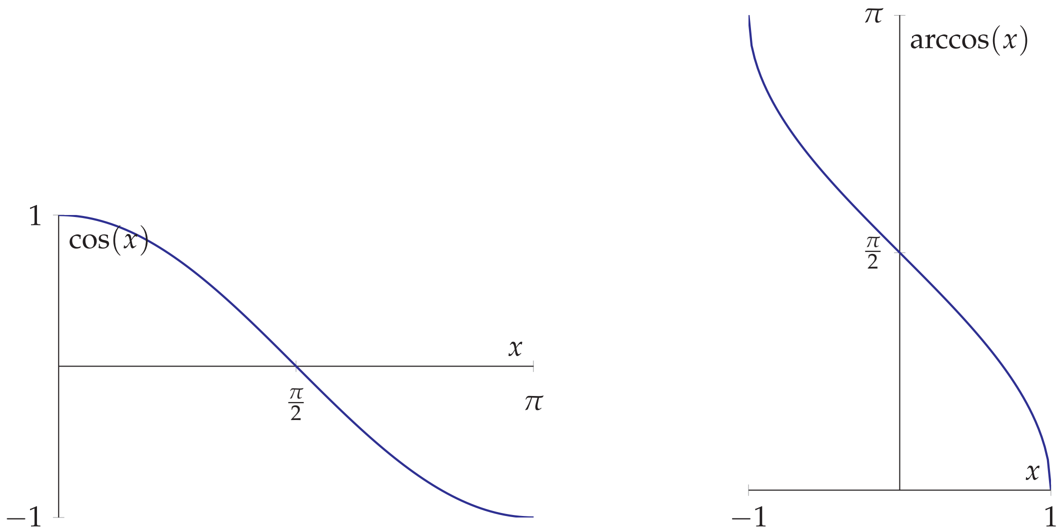
De inverse trigonometriske funktioner

Da ingen af de trigonometriske funktioner sin, cos og tan, der blev diskuteret i foregående afsnit, er bijektioner, kan vi ikke finde inverse til disse funktioner. Men som i Eksempel 2.3.1 kan vi justere definitionsmængden for disse funktioner og derved opnå tilsvarende funktioner, der har en invers. Disse inverse kendes som de *inverse trigonometriske funktioner* (nogle gange også kaldet arcusfunktionerne). I dette afsnit ser vi nærmere på detaljerne bag, hvordan disse er defineret.

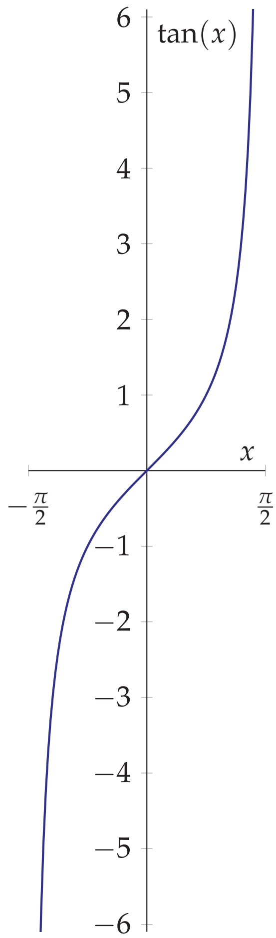
Hvis definitionsmængden for funktionen $\sin : \mathbb{R} \to [-1, 1]$ afgrænses til det lukkede interval $[-\pi/2, \pi/2]$, opnås funktionen $f : [-\pi/2, \pi/2] \to [-1, 1]$ defineret ved $f(x) = \sin(x)$. Funktionen f er en bijektiv funktion, da grafen for sinusfunktionen er strengt voksende på intervallet $[-\pi/2, \pi/2]$ med værdier fra -1 til 1. Den inverse af denne funktion kaldes *arcsinus* og betegnes typisk arcsin i matematiske formler. Derfor er $\arcsin : [-1, 1] \to [-\pi/2, \pi/2]$ den inverse af sinusfunktionen, hvis definitionsmængde er begrænset til $[-\pi/2, \pi/2]$. Graferne for disse to funktioner ser således ud:

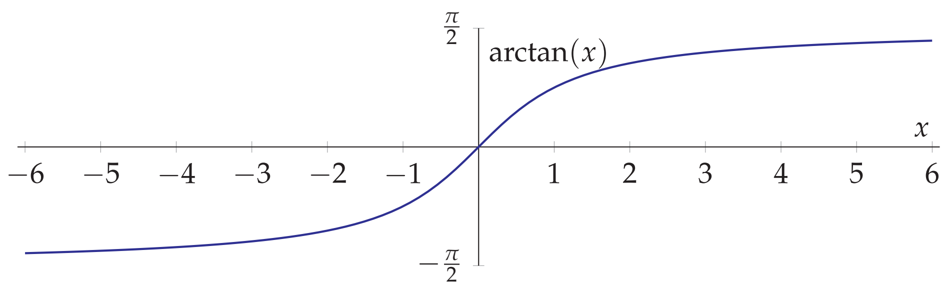


På tilsvarende måde kan vi definere *arccosinus*-funktionen. Først afgrænser vi definitionsmængden for den sædvanlige cosinusfunktion til det lukkede interval $[0, \pi]$. Den resulterende funktion $g : [0, \pi] \to [-1, 1]$, hvor $g(x) = \cos(x)$, er strengt aftagende samt surjektiv og dermed bijektiv. Den inverse af g er arccosinusfunktionen. Den betegnes typisk ved arccos. Derfor er $\arccos : [-1, 1] \to [0, \pi]$ den inverse af cosinusfunktionen, når dens definitionsmængde er begrænset til $[0, \pi]$. Vi illustrerer situationen ved at vise graferne for disse to funktioner.



Slutteligt behandler vi tangensfunktionen i detaljer. Her kan vi overveje funktionen $h :]-\pi/2, \pi/2[\to \mathbb{R}$, hvor $h(x) = \tan(x)$. Med andre ord er funktionen h simpelthen tangensfunktionen med sin definitionsmængde afgrænset til det åbne interval $]-\pi/2, \pi/2[$. Funktionen h er en strengt voksende funktion med billedet $\mathbb{R}$, hvilket medfører, at h er en bijektion. Den inverse af h kaldes *arctangens*-funktionen, typisk betegnet arctan. Mere præcist er $\arctan : \mathbb{R} \to]-\pi/2, \pi/2[$ den inverse af tangensfunktionen, hvis definitionsmængde er begrænset til $]-\pi/2, \pi/2[$. Som før illustrerer vi situationen ved at vise graferne for disse funktioner.





Eksempel 2.3.3

Lad os beregne nogle værdier af de inverse trigonometriske funktioner. Da $\sin(0) = 0$, har vi $\arcsin(0) = 0$. Men selvom $\sin(\pi) = 0$, har vi ikke $\arcsin(0) = \pi$. En funktion må slet ikke kunne opnå to forskellige værdier for den samme inputværdi! Problemet er, at arcsin er den inverse af sinusfunktionen med definitionsmængden begrænset til $[-\pi/2, \pi/2]$. Derfor medfører $\sin(x) = y$ kun $\arcsin(y) = x$, så længe $x \in [-\pi/2, \pi/2]$. For eksempel, da $\sin(\pi/4) = \sqrt{2}/2$, har vi $\arcsin(\sqrt{2}/2) = \pi/4$.

For arccos har vi et lignende fænomen. Man har $\cos(-\pi/4) = \sqrt{2}/2$, men dette medfører ikke $\arccos(\sqrt{2}/2) = -\pi/4$. Denne gang er problemet, at definitionsmængden for cosinusfunktionen blev begrænset til $[0, \pi]$, da arccos-funktionen blev defineret. På intervallet $[0, \pi]$ antager cosinusfunktionen ganske vist værdien $\sqrt{2}/2$, men for $x = \pi/4$. Derfor er $\arccos(\sqrt{2}/2) = \pi/4$.

Som et sidste eksempel har vi $\cos(\pi/3) = 1/2$ og $\sin(\pi/3) = \sqrt{3}/2$. Derfor er $\tan(\pi/3) = \sin(\pi/3)/\cos(\pi/3) = \sqrt{3}$. arctan-funktionen er den inverse af tangensfunktionen med sin definitionsmængde begrænset til $]-\pi/2, \pi/2[$. Da $\pi/3 \in]-\pi/2, \pi/2[$, kan vi derfor konkludere, at $\arctan(\sqrt{3}) = \pi/3$.

Kapitel 3

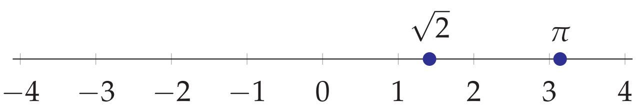
Komplekse tal

3.1 Introduktion til de komplekse tal

I dette kapitel introducerer vi mængden af *komplekse tal*, almindeligvis betegnet $\mathbb{C}$. De komplekse tal viser sig ekstremt nyttige, og ingen forsker eller ingeniør kan undvære dem i den moderne videnskab. Lad os først tage et kort kig på nogle andre talmængder i matematikken. De naturlige tal $\mathbb{N} = \{1, 2, 3, \ldots\}$ har, som deres navn antyder, en meget naturlig fortolkning: de dukker op, når man vil tælle ting. Heltallene $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ kom frem, da forskelle mellem naturlige tal skulle beskrives. I Eksempel 2.1.4 så vi desuden mængden af rationale tal $\mathbb{Q}$ bestående af brøker af heltal.

Man kunne få den tanke, at mængden af rationale tal $\mathbb{Q}$ indeholder alle de tal, man nogensinde ville kunne få brug for, men det er ikke tilfældet. For eksempel viser det sig, at ligningen $z^2 = 2$ ikke har en løsning i $\mathbb{Q}$. I stedet for blot at acceptere, at en sådan ligning ikke har nogen løsninger, udvidede matematikere mængden af rationale tal $\mathbb{Q}$ til mængden af reelle tal $\mathbb{R}$. Inden for $\mathbb{R}$ har ligningen $z^2 = 2$ to løsninger, nemlig $\sqrt{2}$ og $-\sqrt{2}$. Mængden $\mathbb{R}$ er meget stor og indeholder mange interessante tal, såsom e, grundlaget for den naturlige logaritme, og π. Ofte repræsenteres de reelle tal $\mathbb{R}$ som en ret linje, som vi vil kalde *den reelle tallinje*. Ethvert punkt på den reelle tallinje repræsenterer et reelt tal (se Figur 3.1).

Figur 3.1: Den reelle tallinje.



Igen havde man i nogen tid den tanke, at mængden af reelle tal $\mathbb{R}$ må indeholde alle tal, man nogensinde ville kunne få brug for. Men hvad med en ligning som $z^2 = -1$? Indenfor mængden af reelle tal har denne ligning tydeligvis ingen løsninger. Vi er havnet i samme situation som tidligere med ligningen $z^2 = 2$, fra før de reelle tal blev introduceret. Lad os derfor forsøge at finde frem til en mængde af tal, der er endnu større end $\mathbb{R}$, og som indeholder en løsning til ligningen $z^2 = -1$. En umiddelbar tanke vil være at tillade $\sqrt{-1}$ som en løsning til $z^2 = -1$. Oftest benyttes dog i stedet symbolet i. Derfor ønsker vi, at $i^2 = -1$. Med dette in mente definerer vi nu de komplekse tal som følger.

Definition 3.1.1

Mængden $\mathbb{C}$ af komplekse tal defineres som:

$$\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}.$$

Det komplekse tal i opfylder reglen

$$i^2 = -1.$$

Udtrykket $a + bi$ skal blot opfattes som et polynomium i variablen i. Derfor gælder der for eksempel, at $a + bi = a + ib$. Det gør heller ingen forskel at skrive $a + b \cdot i$ fremfor $a + bi$. Dermed har vi for alle $a, b \in \mathbb{R}$:

$$a + bi = a + b \cdot i = a + i \cdot b = a + ib.$$

Slutteligt betegner $a + bi$ nøjagtigt det samme komplekse tal som $bi + a$, præcis som vi ville behandle polynomier.

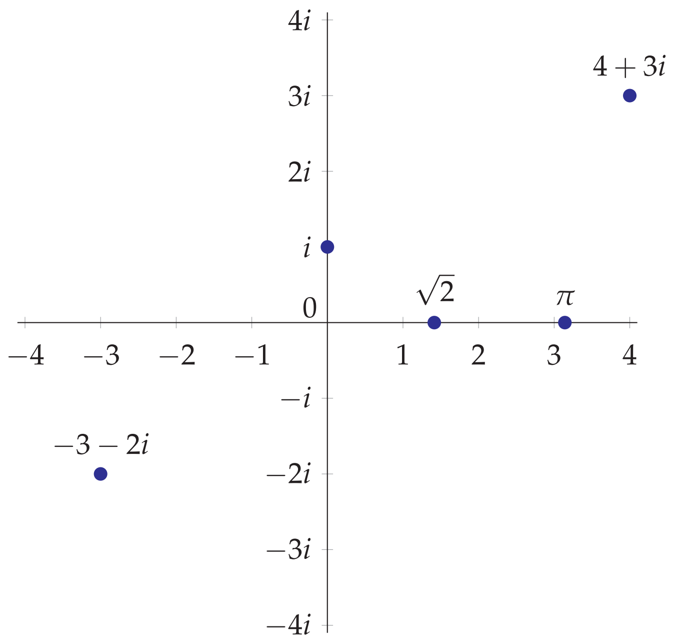
For ethvert $a, b, c, d \in \mathbb{R}$ er de to komplekse tal $a + bi$ og $c + di$ ens, hvis og kun hvis $a = c$, og $b = d$. Hvis $a = 0$, vil man typisk reducere $0 + bi$ til bi, altså $0 + bi = bi$. Ligeledes, hvis $b = 0$, skriver man typisk a i stedet for $a + 0i$. Endelig, hvis $b = 1$, udelades ofte 1 foran i. For eksempel er $5 + 1i = 5 + i$. Med alle disse notationsrettelinjer på plads har vi for eksempel, at $i = 1i = 0 + 1i = 0 + 1 \cdot i$. Mængden af komplekse tal $\mathbb{C}$ indeholder mængden af reelle tal $\mathbb{R}$, da vi for $a \in \mathbb{R}$ har $a = a + 0i$. Med andre ord: $\mathbb{R} \subseteq \mathbb{C}$. Vi har endda $\mathbb{R} \subsetneq \mathbb{C}$, da $i \in \mathbb{C}$, mens $i \notin \mathbb{R}$.

De komplekse tal kan repræsenteres grafisk, men denne gang som et talplan, der kaldes *det komplekse talplan*. Et komplekst tal $a + bi$ repræsenteres af punktet (a, b) i dette talplan. Dette betyder, at tallet i har koordinaterne $(0, 1)$ og derfor vil ligge på andenaksen. Tallet i samt andre eksempler på komplekse tal er indtegnet i det komplekse talplan i Figur 3.2.

Akserne i det komplekse talplan gives særlige navne. Den vandrette akse kaldes *den reelle akse* eller blot *realaksen*, da alle reelle tal ligger på den. Faktisk vil et tal på den reelle akse i det komplekse talplan være på formen $a + 0i$ for ethvert $a \in \mathbb{R}$.

Den lodrette akse kaldes *den imaginære akse* eller blot *imaginæraksen*. Symbolet i er en forkortelse for ordet imaginær. De tal, der ligger på den lodrette akse, kaldes *rent imaginære tal*. Udtrykkene "komplekse tal" og "imaginære tal" er historiske og antyder, at videnskabsfolk har kæmpet med at forstå disse tal. I dag er de komplekse tal helt standard.

Figur 3.2: Det komplekse talplan.



Koordinaterne for et komplekst tal $z \in \mathbb{C}$ i det komplekse talplan tildeles ligeledes særlige navne. Førstekoordinaten kaldes *realdelen* af z (betegnet $\mathrm{Re}(z)$), mens z's andenkoordinat kaldes *imaginærdelen* (betegnet $\mathrm{Im}(z)$). Hvis man kender $\mathrm{Re}(z)$ og $\mathrm{Im}(z)$, kan man opskrive tallet z, da der gælder, at

$$z = \mathrm{Re}(z) + \mathrm{Im}(z)i.$$

Hvis et komplekst tal z er skrevet på formen $\mathrm{Re}(z) + \mathrm{Im}(z)i$, siges tallet z at være på *rektangulær form*. For et givet komplekst tal z kaldes talparret $(\mathrm{Re}(z), \mathrm{Im}(z))$ for z's *rektangulære koordinater*.

Eksempel 3.1.1

Udregn de følgende komplekse tals rektangulære koordinater:

(a) $2 + 3i$

(b) $\sqrt{2}$

(c) i

Svar:

(a) Tallet $2 + 3i$ er på rektangulær form. Derfor kan vi aflæse real- og imaginærdelen direkte. Vi har $\mathrm{Re}(2 + 3i) = 2$ og $\mathrm{Im}(2 + 3i) = 3$. Derfor har det komplekse tal $2 + 3i$ de rektangulære koordinater $(2, 3)$.

(b) Tallet $\sqrt{2}$ er et reelt tal, men vi kan betragte det som et komplekst tal, da $\sqrt{2} = \sqrt{2} + 0i$. Heraf ser vi, at $\mathrm{Re}(\sqrt{2}) = \sqrt{2}$, og $\mathrm{Im}(\sqrt{2}) = 0$. Alle reelle tal har en imaginærdel lig med 0. De rektangulære koordinater for $\sqrt{2}$ er $(\sqrt{2}, 0)$.

(c) Tallet i er et rent imaginært tal, og man kunne også skrive $i = 0 + 1 \cdot i$. Derfor har vi $\mathrm{Re}(i) = 0$ og $\mathrm{Im}(i) = 1$. Alle rent imaginære tal har realdelen 0. De rektangulære koordinater for i er $(0, 1)$.

3.2 Aritmetik med komplekse tal

Nu hvor vi har introduceret de komplekse tal, kan vi begynde vores undersøgelse af deres struktur. Vi er vant til at kunne lægge to tal sammen, trække dem fra hinanden, gange dem sammen og dividere dem med hinanden. Det er endnu ikke klart, om sådanne operationer kan udføres med komplekse tal, men som vi vil se nu, er det naturligvis muligt.

Vi starter med at definere addition og subtraktion.

Definition 3.2.1

Lad $a, b, c, d \in \mathbb{R}$, og lad $a + bi$ og $c + di$ være to komplekse tal i $\mathbb{C}$ skrevet på rektangulær form. Da definerer vi:

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

og

$$(a + bi) - (c + di) = (a - c) + (b - d)i.$$

Addition eller subtraktion af to komplekse tal er meget lig addition eller subtraktion af to polynomier af grad ét (polynomier defineres mere præcist i Definition **??**). Man samler simpelthen de led, der ikke involverer i, og de led, der involverer i. Man kan derfor nemt huske, hvordan addition skal udføres, ved at tilføje et mellemliggende trin som følgende:

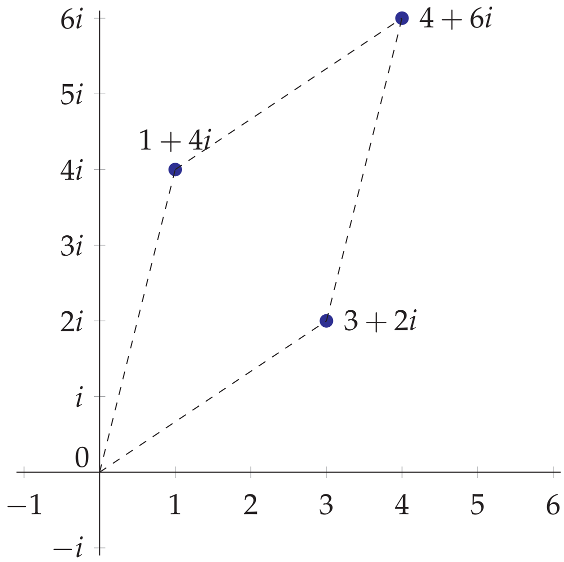
$$\begin{aligned}(a + bi) + (c + di) &= a + bi + c + di \\ &= a + c + bi + di \\ &= (a + c) + (b + d)i.\end{aligned}$$

Subtraktion kan forklares på en tilsvarende måde. Grafisk er addition af komplekse tal meget lig addition af to vektorer i planet, se Figur 3.3. Bemærk, at $(a + bi) + (c + di) = (c + di) + (a + bi)$. Når flere komplekse tal lægges sammen, betyder rækkefølgen, hvori man lægger disse tal sammen, derfor ikke noget.

Eksempel 3.2.1

Reducér følgende udtryk, og opskriv resultaterne på rektangulær form.

Figur 3.3: Addition af komplekse tal. Det er her illustreret grafisk, at $(3+2i)+(1+4i)=4+6i$.



(a) $(3+2i)+(1+4i)$

(b) $(3+2i)-(1+4i)$

(c) $(5-7i)-i$

(d) $(5-7i)-(-10+i)$

Svar:

(a) $(3+2i)+(1+4i)=(3+1)+(2+4)i=4+6i$

(b) $(3+2i)-(1+4i)=(3-1)+(2-4)i=2-2i$

(c) $(5-7i)-i=5+(-7-1)i=5-8i$

(d) $(5-7i)-(-10+i)=(5-(-10))+(-7-1)i=15-8i$

Nu hvor vi har addition og subtraktion af komplekse tal på plads, tager vi et kig på deres multiplikation. Antag for eksempel, at vi vil multiplicere de komplekse tal $a+bi$ og $c+di$, hvorom der som sædvanlig gælder $a,b,c,d \in \mathbb{R}$. Lad os se, hvad der sker, hvis vi bare ganger disse udtryk sammen, mens de opfattes som polynomier i variablen i:

$$(a+bi)\cdot(c+di)=a\cdot(c+di)+bi\cdot(c+di)=a\cdot c+a\cdot d\,i+b\cdot c\,i+b\cdot d\,i^2.$$

Det eneste, vi har gjort indtil nu, er at reducere produktet for at slippe af med parenteserne. For at komme videre skal vi nu huske på, at hele pointen med at introducere i var, at det er en løsning til ligningen $z^2 = -1$. Derfor er $i^2 = -1$. Benyttes dette, fås

$$(a+bi)\cdot(c+di) = a\cdot c + a\cdot di + b\cdot ci + b\cdot d\cdot(-1) = (a\cdot c - b\cdot d) + (a\cdot d + b\cdot c)i.$$

Vi ender igen med et komplekst tal! Alt vi havde brug for, var de sædvanlige regneregler (for at slippe af med parenteserne) samt formlen $i^2 = -1$. Lad os derfor ophøje den formel, vi netop har fundet frem til, som den formelle definition af multiplikation af komplekse tal.

Definition 3.2.2

Lad $a, b, c, d \in \mathbb{R}$, og lad $a + bi$ og $c + di$ være to komplekse tal i $\mathbb{C}$ givet på rektangulær form. Vi definerer:

$$(a+bi)\cdot(c+di) = (a\cdot c - b\cdot d) + (b\cdot c + a\cdot d)i.$$

Der er ingen grund til at huske ovenstående definition udenad. Udregning af et produkt af to komplekse tal på rektangulær form kræver blot, at vi husker, hvordan vi opnåede formlen: vi reducerede produktet ved at gange alle led ud og derefter benytte, at $i^2 = -1$. Bemærk, at $(a + bi)\cdot(c + di) = (c + di)\cdot(a + bi)$, så rækkefølgen af de komplekse tal betyder ikke noget i multiplikation. Man siger, at multiplikation af komplekse tal er *kommutativ*. Vi vil se i Afsnit 3.3, at multiplikationen af to komplekse tal også kan beskrives geometrisk.

Eksempel 3.2.2

Reducér følgende udtryk, og opskriv resultaterne på rektangulær form.

(a) $(1 + 2i)\cdot(3 + 4i)$

(b) $(4 + i)\cdot(4 - i)$

Svar:

(a)

$$\begin{aligned}(1+2i)(3+4i) &= 1\cdot 3 + 1\cdot 4i + 2i\cdot 3 + 2i\cdot 4i\\ &= 3 + 4i + 6i + 8i^2\\ &= 3 + 10i - 8\\ &= -5 + 10i.\end{aligned}$$

(b)

$$\begin{aligned}(4+i)\cdot(4-i) &= 4\cdot 4 + 4\cdot(-i) + i\cdot 4 - i^2\\ &= 16 - 4i + 4i - (-1)\\ &= 17 + 0i\\ &= 17.\end{aligned}$$

I dette tilfælde er resultatet et reelt tal.

Del to af dette eksempel viser, at produktet af to ikke-reelle tal godt kan blive et reelt tal, og eksemplet er et særtilfælde af følgende lemma:

Lemma 3.2.1
Lad $a, b \in \mathbb{R}$, og lad $z = a + bi$ være et komplekst tal på rektangulær form. Da gælder der, at

$$(a+bi) \cdot (a-bi) = a^2 + b^2.$$

Bevis. Vi har

$$\begin{aligned} (a+bi)\cdot(a-bi) &= a\cdot a + a\cdot(-bi) + (bi)\cdot a - b\cdot bi^2 \\ &= a^2 - abi + abi - b^2\cdot(-1) \\ &= a^2 + b^2. \end{aligned}$$

□

Motiveret af dette lemma introducerer vi følgende:

Definition 3.2.3
Lad $z \in \mathbb{C}$ være et komplekst tal og $z = a + bi$ dets rektangulære form. Da defineres den komplekst konjugerede af z som $\bar{z} = a - bi$. Funktionen fra $\mathbb{C}$ til $\mathbb{C}$ defineret ved $z \mapsto \bar{z}$ kaldes *kompleks konjugering*.

Bemærk, at vi direkte af denne definition ser, at $\mathrm{Re}(\bar{z}) = \mathrm{Re}(z)$ og $\mathrm{Im}(\bar{z}) = -\mathrm{Im}(z)$. Derfor er

$$\bar{z} = \mathrm{Re}(z) - \mathrm{Im}(z)i.$$

Dermed medfører Lemma 3.2.1, at

$$z \cdot \bar{z} = \mathrm{Re}(z)^2 + \mathrm{Im}(z)^2. \tag{3.1}$$

Bemærk, at denne ligning medfører, at produktet $z \cdot \bar{z}$ for ethvert $z \in \mathbb{C}$ er et reelt tal.

Kompleks konjugering viser sig at være nyttig, når division af komplekse tal skal defineres. Vi vil gerne kunne dividere ethvert komplekst tal med ethvert komplekst tal udover nul. Bemærk, at vi allerede er i stand til at dividere et komplekst tal $a + bi \in \mathbb{C}$ med et reelt tal $c \in \mathbb{R} \setminus \{0\}$, så længe det ikke er nul, ved at definere:

$$\frac{a+bi}{c} = \frac{a}{c} + \frac{b}{c}i \quad a, b \in \mathbb{R} \text{ og } c \in \mathbb{R} \setminus \{0\}.$$

For at kunne dividere ethvert komplekst tal $z_1 = a + bi$ med ethvert komplekst tal $z_2 = c + di$, der ikke er nul, kan vi udnytte et trick ved at observere følgende:

$$\frac{z_1}{z_2} = \frac{a+bi}{c+di} = \frac{a+bi}{c+di} \cdot \frac{c-di}{c-di} = \frac{(a+bi)\cdot(c-di)}{c^2+d^2}. \tag{3.2}$$

Tælleren på højresiden af denne ligning er blot et produkt af to komplekse tal, hvilket vi allerede er i stand til at håndtere. Nævneren er et reelt tal, der ikke er nul, nemlig $c^2 + d^2$, og vi ved også nu, hvordan vi dividerer et komplekst tal med et sådant reelt tal. Lad os sikre os, at nævneren $c^2 + d^2$ ganske rigtigt er et reelt tal men ikke nul. For det første er det et reelt tal, fordi c og d er reelle tal. For det andet kan et reelt tal opløftet i anden ikke blive negativt, altså $c^2 \geq 0$, og $d^2 \geq 0$. Den eneste måde, hvorpå $c^2 + d^2 = 0$ kan blive opfyldt, er derfor, hvis både $c^2 = 0$, og $d^2 = 0$. Men så er $c = 0$ og $d = 0$, hvilket medfører, at $c + di = 0$, i modstrid med vores antagelse om, at vi dividerer med et komplekst tal, der ikke er nul.

Fremgangsmåden i division af et komplekst tal med et komplekst tal er altså, at hvis $z_1 \in \mathbb{C}$, og $z_2 \in \mathbb{C}\backslash\{0\}$, så kan z_1/z_2 udføres, hvis både tæller og nævner multipliceres med den komplekst konjugerede af z_2, da nævneren så bliver $z_2 \cdot \overline{z_2}$, hvilket er et reelt tal. Ligning (3.2) giver os derfor mulighed for at dividere med komplekse tal, der ikke er nul. Et særtilfælde af Ligning (3.2) er følgende:

$$\frac{1}{c+di} = \frac{1}{c+di} \cdot \frac{c-di}{c-di} = \frac{c-di}{c^2+d^2} = \frac{c}{c^2+d^2} - \frac{d}{c^2+d^2}i. \tag{3.3}$$

Lad os se på nogle eksempler.

Eksempel 3.2.3

Reducér følgende udtryk, og skriv resultaterne på rektangulær form.

(a) $1/(1+i)$

(b) $\dfrac{1+2i}{3+4i}$

Svar:

(a) Bemærk, at $1/(1+i)$ bare er en anden måde at skrive $\frac{1}{1+i}$ på. Derfor får vi ved brug af Ligning (3.2), eller alternativt Ligning (3.3):

$$1/(1+i) = \frac{1 \cdot (1-i)}{(1+i)\cdot(1-i)} = \frac{1-i}{1^2+1^2} = \frac{1-i}{2} = \frac{1}{2} - \frac{1}{2}i.$$

(b) Ved hjælp af Ligning (3.2) får vi

$$\begin{aligned}\frac{1+2i}{3+4i} &= \frac{(1+2i)(3-4i)}{(3+4i)(3-4i)} = \frac{3-4i+6i-8i^2}{3^2+4^2} \\ &= \frac{3+2i+8}{9+16} = \frac{11+2i}{25} = \frac{11}{25} + \frac{2}{25}i.\end{aligned}$$

Lad os samle nogle forskellige egenskaber for multiplikation og addition i en sætning. Vi vil ikke bevise sætningen; flere af udsagnene er faktisk allerede vist i det foregående.

Sætning 3.2.2

Lad $\mathbb{C}$ være mængden af komplekse tal, og lad $z_1, z_2, z_3 \in \mathbb{C}$ være valgt vilkårligt. Da er følgende egenskaber opfyldt:

(i) Addition og multiplikation er *associative*: $z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$ og $z_1 \cdot (z_2 \cdot z_3) = (z_1 \cdot z_2) \cdot z_3$.

(ii) Addition og multiplikation er *kommutative*: $z_1 + z_2 = z_2 + z_1$ og $z_1 \cdot z_2 = z_2 \cdot z_1$.

(iii) Der gælder *distributivitet* af multiplikation over addition: $z_1 \cdot (z_2 + z_3) = z_1 \cdot z_2 + z_1 \cdot z_3$.

Yderligere har man for komplekse tal, ligesom for de reelle tal, følgende egenskaber:

Sætning 3.2.3

(i) Addition og multiplikation har et neutralt element: elementerne 0 og 1 i $\mathbb{C}$ opfylder $z + 0 = z$ og $z \cdot 1 = z$ for alle $z \in \mathbb{C}$.

(ii) Additivt inverse eksisterer: for ethvert $z \in \mathbb{C}$ findes der et element i $\mathbb{C}$, betegnet $-z$, kaldet den additivt inverse af z, således at $z + (-z) = 0$.

(iii) Multiplikativt inverse eksisterer: for ethvert $z \in \mathbb{C} \setminus \{0\}$ findes der et element i $\mathbb{C}$, betegnet z^{-1} eller $1/z$, kaldet den multiplikativt inverse af z, således at $z \cdot z^{-1} = 1$.

Bemærk, at punkt to og tre i Sætning 3.2.3 garanterer eksistensen af additivt og multiplikativt inverse. Det angives dog ikke, hvordan man bestemmer disse inverse. Men faktisk har vi allerede set, hvordan man kan bestemme dem. Vi opstiller i følgende eksempel regnemetoden algoritmisk ved at nedskrive den præcise metodik til beregning af $-z$ og $1/z$ i pseudo-kode:

Eksempel 3.2.4

En mulig algoritme, der udregner $-z$ for et givet komplekst tal z, kan beskrives som følger: Begynd ved at skrive z på rektangulær form, altså find $a, b \in \mathbb{R}$, således at $z = a + bi$. Derefter er $-z = -a - bi$. I pseudo-kode:

Algoritme 4 til udregning af "den additivt inverse af $z \in \mathbb{C}$".

Input: $z \in \mathbb{C}$
$a \leftarrow \mathrm{Re}(z)$
$b \leftarrow \mathrm{Im}(z)$
return $-a - bi$

For at udregne $1/z$ gør vi brug af Ligning (3.3). Bemærk, at $1/z$ ikke eksisterer, hvis $z = 0$. Derfor tjekker algoritmen først, om $z = 0$.

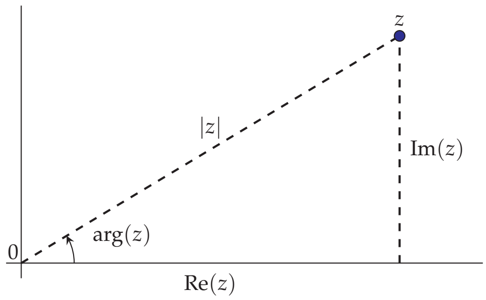
Algoritme 5 til udregning af "den multiplikativt inverse af $z \in \mathbb{C}$".

Input: $z \in \mathbb{C}$
if $z = 0$ **then**
  return "0 har ingen multiplikativt invers!"
else
  $c \leftarrow \mathrm{Re}(z)$,
  $d \leftarrow \mathrm{Im}(z)$,
  $N \leftarrow c^2 + d^2$,
  return $\frac{c}{N} - \frac{d}{N} i$.

3.3 Modulus og argument

Vi har set i Afsnit 3.1, at et komplekst tal z kan bestemmes entydigt ved dets realdel $\mathrm{Re}(z)$ og dets imaginærdel $\mathrm{Im}(z)$, da der for ethvert $z \in \mathbb{C}$ gælder, at $z = \mathrm{Re}(z) + \mathrm{Im}(z)i$. Vi kaldte talparret $(\mathrm{Re}(z), \mathrm{Im}(z))$ for z's rektangulære koordinater. I dette afsnit vil vi introducere en anden måde at beskrive et komplekst tal på. Givet et komplekst tal z kan vi tegne en trekant i det komplekse talplan med hjørner i de komplekse tal 0, $\mathrm{Re}(z)$ og z (se Figur 3.4). Afstanden fra z til 0 kaldes *modulus* eller *absolutværdien* af z og betegnes $|z|$. Vinklen fra den positive del af den reelle akse til vektoren fra 0 til z kaldes *argumentet* for z og betegnes $\arg(z)$.

Vi vil altid angive et argument (faktisk enhver vinkel) i radianer. Da vinklen 2π betegner en fuld omdrejning, kan man altid tilføje et heltalsmultiplum af 2π til en vinkel. For eksempel kan vinklen $-\pi/4$ også angives som $7\pi/4$, da $-\pi/4 + 2\pi = 7\pi/4$. Af denne grund siger man, at argumentet for et komplekst tal kun er bestemt op til et multiplum af 2π. Et udtryk som "$\arg(z) = 5\pi/4$" skal derfor læses som: "$5\pi/4$ er et argument for z". Det er altid muligt at finde et argument for et komplekst tal z i intervallet $]-\pi, \pi]$. Denne værdi kaldes typisk tallets *hovedargument* og betegnes $\mathrm{Arg}(z)$.



Figur 3.4: Modulus og argument af et komplekst tal z.

Fra Figur 3.4 kan vi udlede, at

$$\mathrm{Re}(z) = |z| \cos(\arg(z)) \quad \text{og} \quad \mathrm{Im}(z) = |z| \sin(\arg(z)). \tag{3.4}$$

Givet $|z|$ og $\arg(z)$ kan vi derfor bestemme z's rektangulære koordinater. Dette betyder, at talparret $(|z|, \arg(z))$ fuldstændigt bestemmer det komplekse tal z, da

$$z = |z| \left(\cos(\arg(z)) + \sin(\arg(z)) i \right). \tag{3.5}$$

Talparret $(|z|, \mathrm{Arg}(z))$ kaldes de *polære koordinater* for et komplekst tal $z \in \mathbb{C}$. Hvis et komplekst tal z er skrevet på formen $z = r\left(\cos(\alpha) + i\sin(\alpha)\right)$, hvor r er et positivt, reelt tal, gælder der, at $|z| = r$, og $\arg(z) = \alpha$. Desuden, hvis $\alpha \in]-\pi, \pi]$, så er $\mathrm{Arg}(z) = \alpha$. Vi kan igen fra Figur 3.4 udlede, at

$$|z| = \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2} \quad \text{og} \quad \tan(\arg(z)) = \mathrm{Im}(z)/\mathrm{Re}(z), \text{ hvis } \mathrm{Re}(z) \neq 0. \tag{3.6}$$

Denne ligning er nøglen til at bestemme et tals polære koordinater ud fra dets rektangulære koordinater. Mere præcist, ved brug af den inverse tangensfunktion arctan, som diskuteres i underafsnit 2.3, har vi følgende:

Sætning 3.3.1

Hvis et komplekst tal z, der er forskelligt fra nul, har de polære koordinater (r, α), så er

$$\mathrm{Re}(z) = r\cos(\alpha) \quad \text{og} \quad \mathrm{Im}(z) = r\sin(\alpha).$$

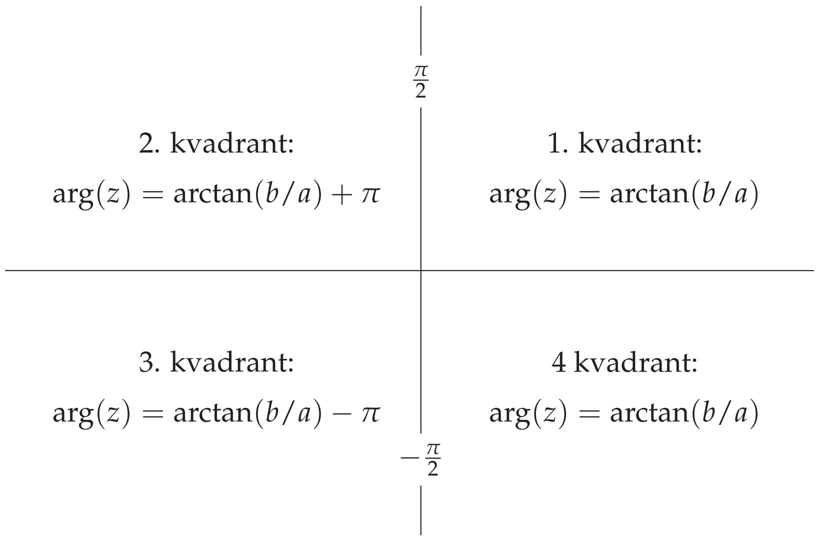
Omvent, hvis et komplekst tal z, der er forskelligt fra nul, har de rektangulære koordinater (a, b), så er:

$$|z| = \sqrt{a^2 + b^2} \quad \text{og} \quad \mathrm{Arg}(z) = \begin{cases} \arctan(b/a) & \text{hvis } a > 0, \\ \pi/2 & \text{hvis } a = 0 \text{ og } b > 0, \\ \arctan(b/a) + \pi & \text{hvis } a < 0 \text{ og } b \geq 0, \\ -\pi/2 & \text{hvis } a = 0 \text{ og } b < 0, \\ \arctan(b/a) - \pi & \text{hvis } a < 0 \text{ og } b < 0. \end{cases}$$

Bevis. Givet de polære koordinater for z kan vi benytte Ligning (3.4) til at bestemme dets rektangulære koordinater. Omvendt, givet de rektangulære koordinater (a, b) for z får vi fra Ligning (3.6), at $|z| = \sqrt{a^2 + b^2}$. Hvis $a = 0$, ligger tallet z på den imaginære akse. I et sådant tilfælde har vi, at $\mathrm{Arg}(z) = \pi/2$, hvis $b > 0$, og $\mathrm{Arg}(z) = -\pi/2$, hvis $b < 0$. Hvis $a \neq 0$, gælder der ifølge Ligning (3.6), at $\tan(\mathrm{Arg}(z)) = b/a$. Derfor gælder der da, at $\mathrm{Arg}(z) = \arctan(b/a) + n\pi$ for et heltal $n \in \mathbb{Z}$. Hvis z ligger i første eller fjerde kvadrant, ligger $\mathrm{Arg}(z)$ i intervallet $]-\pi/2, \pi/2[$. I det tilfælde får vi derfor, at $\mathrm{Arg}(z) = \arctan(b/a)$. Hvis z ligger i anden kvadrant, ligger dets argument i intervallet $]\pi/2, \pi]$. Derfor får vi da, at $\mathrm{Arg}(z) = \arctan(b/a) + \pi$. Ligeledes, hvis z ligger i tredje kvadrant, får vi, at $\mathrm{Arg}(z) = \arctan(b/a) - \pi$. $\square$

Modulus kan forstås som en funktion $f : \mathbb{C} \to \mathbb{R}$, hvor $f(z) = |z|$. Den spiller en lignende rolle for de komplekse tal som absolutværdifunktionen i Eksempel 2.2.7. Hvis $z = a + 0i$ er et reelt tal, gælder der, at $|z| = \sqrt{a^2 + 0^2}$, hvis vi anvender modulusfunktionen. Men vi har også, at $\sqrt{a^2} = |a|$, hvor $|a|$ her betegner absolutværdien af et reelt tal. Derfor er modulus, når den udregnes af et reelt tal a, nøjagtigt det samme som absolutværdien af a. Dette forklarer, hvorfor det giver mening at benytte notationen $|a|$ både for den sædvanlige absolutværdi af et reelt tal samt for modulus af et komplekst tal. Faktisk kaldes $|z|$ ofte også absolutværdien af et komplekst tal. Bemærk slutteligt, at $|z|^2 = \text{Re}(z)^2 + \text{Im}(z)^2 = z \cdot \bar{z}$, hvilket er den sidste lighed, der følger af Ligning (3.1).

Formlen for argumentet af et komplekst tal $a + bi$ afhænger af, i hvilken kvadrant i det komplekse talplan tallet ligger (se Figur 3.5).



Figur 3.5: Formler for argumentet af $z = a + bi$.

Eksempel 3.3.1

Bestem de polære koordinater for følgende komplekse tal.

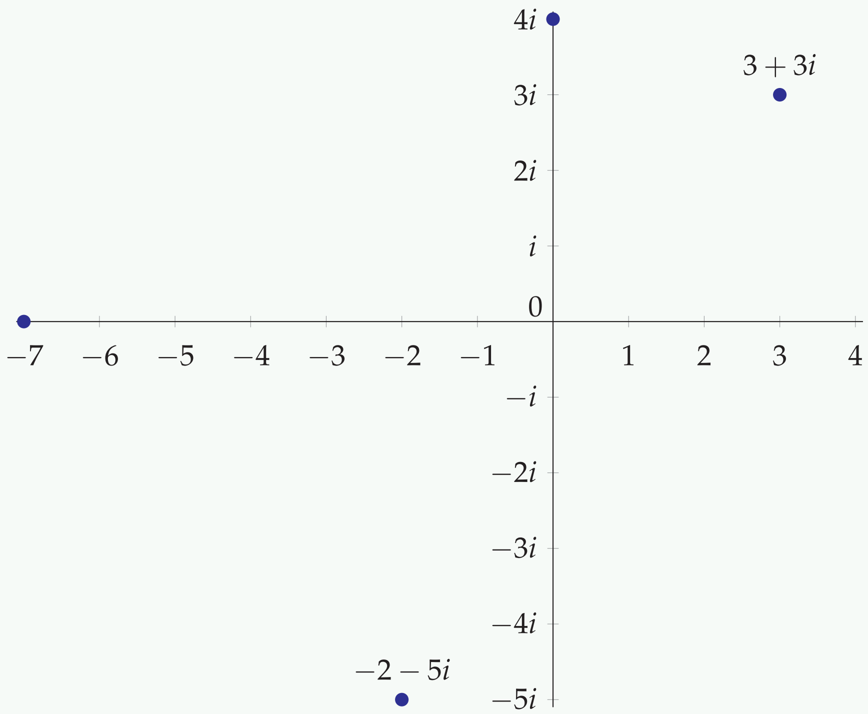
(a) $4i$

(b) -7

(c) $3 + 3i$

(d) $-2 - 5i$

Svar: Vi kan bestemme modulus og argument ved hjælp af Sætning 3.3.1. Figur 3.5 er nyttig, når vi udregner argumentet. Derfor plotter vi først de fire givne komplekse tal i det komplekse talplan.



(a) $|4i| = |0+4i| = \sqrt{0^2+4^2} = 4$ og $\operatorname{Arg}(4i) = \pi/2$. Derfor har $4i$ de polære koordinater $(4, \pi/2)$.

(b) $|-7| = \sqrt{(-7)^2+0^2} = 7$ og $\operatorname{Arg}(-7) = \arctan(0/(-7)) + \pi = \pi$. Derfor har -7 de polære koordinater $(7, \pi)$.

(c) $|3+3i| = \sqrt{3^2+3^2} = 3\sqrt{2}$ og $\operatorname{Arg}(3+3i) = \arctan(3/3) = \pi/4$. Derfor har $3+3i$ de polære koordinater $(3\sqrt{2}, \pi/4)$.

(d) $|-2-5i| = \sqrt{(-2)^2+(-5)^2} = \sqrt{29}$

og

$\operatorname{Arg}(-2-5i) = \arctan((-5)/(-2)) - \pi = \arctan(5/2) - \pi$. Derfor har $-2-5i$ de polære koordinater $(\sqrt{29}, \arctan(5/2) - \pi)$.

Eksempel 3.3.2

De følgende polære koordinater er givet. Bestem de tilsvarende komplekse tal, og skriv dem på rektangulær form.

(a) $(2, \pi/3)$

(b) $(10, \pi)$

(c) $(4, -\pi/4)$

(d) $(2\sqrt{3}, -2\pi/3)$

(e) $(3, 2)$

Svar: Vi benytter Ligning (3.5) til at bestemme de komplekse tal z, der svarer til de givne polære koordinater. Derefter udtrykker vi disse komplekse tal på rektangulær form.

(a) $z = 2 \cdot (\cos(\pi/3) + \sin(\pi/3)i) = 2 \cdot (1/2 + \sqrt{3}/2i) = 1 + \sqrt{3}i$.

(b) $z = 10 \cdot (\cos(\pi) + \sin(\pi)i) = -10 + 0i = -10$.

(c) $z = 4 \cdot (\cos(-\pi/4) + \sin(-\pi/4)i) = 4 \cdot (\sqrt{2}/2 - \sqrt{2}/2i) = 2\sqrt{2} - 2\sqrt{2}i$.

(d) $z = 2\sqrt{3} \cdot (\cos(-2\pi/3) + \sin(-2\pi/3)i) = 2\sqrt{3} \cdot (-1/2 - \sqrt{3}/2i) = -\sqrt{3} - 3i$.

(e) $z = 3 \cdot (\cos(2) + \sin(2)i) = 3\cos(2) + 3\sin(2)i$.

3.4 Den komplekse eksponentialfunktion

Som vi har set, kan mange udregninger, der kan udføres med reelle tal, så som addition, subtraktion, multiplikation og division, også udføres med komplekse tal. Vi kommer i dette afsnit til at se, at også eksponentialfunktionen $\exp : \mathbb{R} \to \mathbb{R}_{>0}$, hvor $\exp(t) = e^t$, kan defineres for komplekse tal. Dette kaldes *den komplekse eksponentialfunktion*.

Definition 3.4.1

Lad $z \in \mathbb{C}$ være et komplekst tal, hvis rektangulære form er givet ved $z = a + bi$ for vilkårlige $a, b \in \mathbb{R}$. Da definerer vi

$$e^z = e^a \cdot (\cos(b) + \sin(b)i).$$

Den komplekse eksponentialfunktion betegnes typisk også exp. Denne gang er funktionens definitionsmængde dog $\mathbb{C}$. Mere præcist er den komplekse eksponentialfunktion funktionen $\exp : \mathbb{C} \to \mathbb{C}$. Bemærk, at hvis z er et reelt tal, det vil sige $z = a + 0i$, så er $e^z = e^a \cdot (\cos(0) + \sin(0)i) = e^a$. Den komplekse eksponentialfunktion giver derfor nøjagtig det samme, som den sædvanlige eksponentialfunktion ville have givet, når den evalueres i et reelt tal. Derfor giver det mening at betegne både eksponentialfunktionen $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ og den komplekse eksponentialfunktion $\exp : \mathbb{C} \to \mathbb{C}$ med samme symbol exp.

Eksempel 3.4.1

Skriv følgende udtryk på rektangulær form:

(a) e^2

(b) e^{1+i}

(c) $e^{\pi i}$

(d) $e^{\ln(2)+i\pi/4}$ (ved ln menes der logaritmen med grundtal e)

(e) $e^{2\pi i}$

Svar: Vi benytter Definition 3.4.1 og reducerer, indtil vi opnår den ønskede rektangulære form.

(a) Da e^2 er et reelt tal, er det allerede på rektangulær form. Hvis vi alligevel benytter Definition 3.4.1, får vi $e^2 = e^{2+0i} = e^2 \cdot (\cos(0) + \sin(0)i) = e^2 \cdot (1 + 0i) = e^2$, hvilket bekræfter, at e^2 allerede var på rektangulær form. Det er også fint at skrive $e^2 = e^2 + 0i$ og derefter returnere $e^2 + 0i$ som svar.

(b) $e^{1+i} = e^1 \cdot (\cos(1) + \sin(1)i) = e\cos(1) + e\sin(1)i$.

(c) $e^{\pi i} = e^{0+\pi i} = e^0 \cdot (\cos(\pi) + \sin(\pi)i) = 1 \cdot (-1 + 0i) = -1$.

(d) $e^{\ln(2)+i\pi/4} = e^{\ln(2)} \cdot (\cos(\pi/4) + \sin(\pi/4)i) = 2(\sqrt{2}/2 + \sqrt{2}/2i) = \sqrt{2} + \sqrt{2}i$.

(e) $e^{2\pi i} = \cos(2\pi) + \sin(2\pi)i = 1 + 0i = 1$. Bemærk, at også $e^0 = 1$. Dette viser, at den komplekse eksponentialfunktion ikke er injektiv.

Direkte fra Definition 3.4.1 ser vi, at vi for ethvert $z \in \mathbb{C}$ har:

$$\mathrm{Re}(e^z) = e^{\mathrm{Re}(z)} \cos(\mathrm{Im}(z)) \quad \text{og} \quad \mathrm{Im}(e^z) = e^{\mathrm{Re}(z)} \sin(\mathrm{Im}(z)).$$

Den komplekse eksponentialfunktion har mange egenskaber til fælles med den sædvanlige reelle eksponentialfunktion. For at vise disse, benytter vi os af følgende lemma.

Lemma 3.4.1

Lad $\alpha_1, \alpha_2 \in \mathbb{R}$. Vi har

$$(\cos(\alpha_1) + \sin(\alpha_1)i) \cdot (\cos(\alpha_2) + \sin(\alpha_2)i) = \cos(\alpha_1 + \alpha_2) + \sin(\alpha_1 + \alpha_2)i.$$

Bevis. Ved at gange parenteserne ud kan vi beregne real- og imaginærdelen af produktet $(\cos(\alpha_1) + \sin(\alpha_1)i) \cdot (\cos(\alpha_2) + \sin(\alpha_2)i)$. Det viser sig, at realdelen er givet ved $\cos(\alpha_1)\cos(\alpha_2) - \sin(\alpha_1)\sin(\alpha_2)$ og imaginærdelen ved $\cos(\alpha_1)\sin(\alpha_2) + \sin(\alpha_1)\cos(\alpha_2)$. Ved at benytte additionsformlerne for cosinus- og sinusfunktionerne følger lemmaet. □

Sætning 3.4.2

Lad z, z_1 og z_2 være komplekse tal og n et heltal. Da gælder der, at

$$e^z \neq 0$$

$$1/e^z = e^{-z}$$

$$e^{z_1} e^{z_2} = e^{z_1+z_2}$$

$$e^{z_1}/e^{z_2} = e^{z_1-z_2}$$

$$(e^z)^n = e^{nz}$$

Bevis. Lad os vise det tredje punkt: $e^{z_1}e^{z_2} = e^{z_1+z_2}$. Først skrives z_1 og z_2 på rektangulær form: $z_1 = a_1 + b_1 i$ og $z_2 = a_2 + b_2 i$. Derefter får vi, at

$$\begin{aligned}
e^{z_1} \cdot e^{z_2} &= e^{a_1} \cdot (\cos(b_1) + \sin(b_1)i) \cdot e^{a_2} \cdot (\cos(b_2) + \sin(b_2)i) \\
&= e^{a_1} \cdot e^{a_2} \cdot (\cos(b_1) + \sin(b_1)i) \cdot (\cos(b_2) + \sin(b_2)i) \\
&= e^{a_1+a_2} \cdot (\cos(b_1) + \sin(b_1)i) \cdot (\cos(b_2) + \sin(b_2)i) \\
&= e^{a_1+a_2} \cdot (\cos(b_1 + b_2) + \sin(b_1 + b_2)i) \text{ (ved at benytte Lemma 3.4.1)} \\
&= e^{a_1+a_2+(b_1+b_2)i} = e^{z_1+z_2}.
\end{aligned}$$

□

3.5 Eulers formel

Den komplekse eksponentialfunktion angiver en forbindelse mellem trigonometri og komplekse tal. Vi vil udforske denne forbindelse i dette afsnit.

Lad t være et reelt tal. Formlen

$$e^{it} = \cos(t) + i\sin(t) \tag{3.7}$$

er kendt som *Eulers formel* og er en konsekvens af Definition 3.4.1. Den medfører, at

$$e^{-it} = \cos(-t) + i\sin(-t) = \cos(t) - i\sin(t). \tag{3.8}$$

Ligningerne (3.7) og (3.8) kan opfattes som ligninger med $\cos(t)$ og $\sin(t)$ som de ukendte. Ved at løse for $\cos(t)$ og $\sin(t)$ får vi:

$$\cos(t) = \frac{e^{it} + e^{-it}}{2} \text{ og } \sin(t) = \frac{e^{it} - e^{-it}}{2i}. \tag{3.9}$$

Ligning (3.9) kan benyttes til omskrivning af produkter af cos- og sin-funktioner til en sum af cos- og sin-funktioner (det vil sige, som en sum af rent harmoniske funktioner). Denne type beregninger forekommer ofte i frekvensanalyse, hvor man forsøger at omskrive vilkårlige funktioner til en sum af rent harmoniske funktioner. Det kan også være nyttigt til udregning af integraler over trigonometriske udtryk, som vi ser i følgende eksempel.

Eksempel 3.5.1

Udregn $\int \sin(3t)\cos(t)dt$.

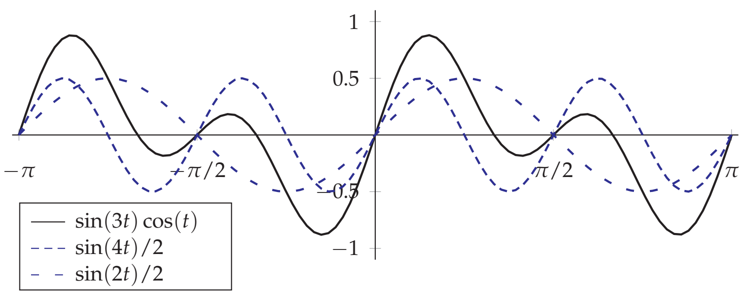
Svar: Vi starter med at benytte identiteterne udledt fra Eulers formel i det foregående til at omskrive udtrykket $\sin(3t)\cos(t)$:

$$\begin{aligned}\sin(3t)\cos(t) &= \frac{e^{i3t}-e^{-i3t}}{2i}\cdot\frac{e^{it}+e^{-it}}{2} = \frac{(e^{i3t}-e^{-i3t})(e^{it}+e^{-it})}{4i}\\ &= \frac{e^{i4t}+e^{i2t}-e^{-i2t}-e^{-i4t}}{4i} = \frac{1}{2}\left(\frac{e^{i4t}-e^{-i4t}}{2i}+\frac{e^{i2t}-e^{-i2t}}{2i}\right)\\ &= \frac{\sin(4t)}{2}+\frac{\sin(2t)}{2}.\end{aligned}$$

Nu får vi

$$\int \sin(3t)\cos(t)dt = \int \frac{\sin(4t)}{2}+\frac{\sin(2t)}{2}dt = -\frac{\cos(4t)}{8}-\frac{\cos(2t)}{4}+c, \quad c\in\mathbb{R}.$$

I Figur 3.6 illustreres identiteten $\sin(3t)\cos(t) = \frac{\sin(4t)}{2}+\frac{\sin(2t)}{2}$ fra eksemplet.



Figur 3.6: Illustration af identiteten $\sin(3t)\cos(t) = \frac{\sin(4t)}{2}+\frac{\sin(2t)}{2}$.

En anden anvendelse af Eulers formel er givet i følgende sætning.

Sætning 3.5.1

Lad $n\in\mathbb{N}$ være et naturligt tal. Da gælder følgende formler:

$$\cos(nt) = \mathrm{Re}((\cos(t)+\sin(t)i)^n)$$

og

$$\sin(nt) = \mathrm{Im}((\cos(t)+\sin(t)i)^n)$$

Bevis. Nøglen er følgende ligning:

$$\cos(nt) + \sin(nt)i = e^{int} = (e^{it})^n = (\cos(t) + \sin(t)i)^n.$$

Sætningen følger ved at udregne real- og imaginærdelene på begge sider af denne lighed. $\square$

Udtrykkene i denne sætning er kendt som *DeMoivres formler*. Lad os se på nogle eksempler.

Eksempel 3.5.2

Udtryk $\cos(2t)$ og $\sin(2t)$ ved $\cos(t)$ og $\sin(t)$.
Svar: Ifølge DeMoivres formler for $n = 2$ har vi $\cos(2t) = \mathrm{Re}((\cos(t) + \sin(t)i)^2)$ og $\sin(2t) = \mathrm{Im}((\cos(t) + \sin(t)i)^2)$. Da

$$\begin{aligned}(\cos(t) + \sin(t)i)^2 &= \cos^2(t) + 2\cos(t)\sin(t)i + \sin^2(t)i^2 \\ &= \cos^2(t) + 2\cos(t)\sin(t)i - \sin^2(t) \\ &= \cos^2(t) - \sin^2(t) + 2\cos(t)\sin(t)i,\end{aligned}$$

får vi, at

$$\cos(2t) = \cos(t)^2 - \sin(t)^2$$

og

$$\sin(2t) = 2\cos(t)\sin(t).$$

Eksempel 3.5.3

Udtryk $\cos(3t)$ og $\sin(3t)$ ved $\cos(t)$ og $\sin(t)$.
Svar: Ifølge DeMoivres formler for $n = 3$ har vi $\cos(3t) = \mathrm{Re}((\cos(t) + i\sin(t))^3)$ og $\sin(3t) = \mathrm{Im}((\cos(t) + i\sin(t))^3)$. Efter nogle omskrivninger opnås $(\cos(t) + i\sin(t))^3 = (\cos(t)^3 - 3\cos(t)\sin(t)^2) + i(3\cos(t)^2\sin(t) - \sin(t)^3)$. Dermed gælder følgende:

$$\cos(3t) = \cos(t)^3 - 3\cos(t)\sin(t)^2$$

og

$$\sin(3t) = 3\cos(t)^2\sin(t) - \sin(t)^3.$$

3.6 Den polære form af et komplekst tal

Lad r være et positivt, reelt tal og α et reelt tal. Så ser vi fra Definition 3.4.1, at $r \cdot e^{\alpha i} = r \cdot (\cos(\alpha) + \sin(\alpha)i)$. Som vi har set i og efter Ligning (3.5), har tallet $r \cdot e^{\alpha i}$ modulus r og et argument lig med α (se Figur 3.7). Vi kan også omskrive Ligning (3.5) som $z = |z|e^{i\arg(z)}$. Denne skrivemåde for et komplekst tal gives et særligt navn:

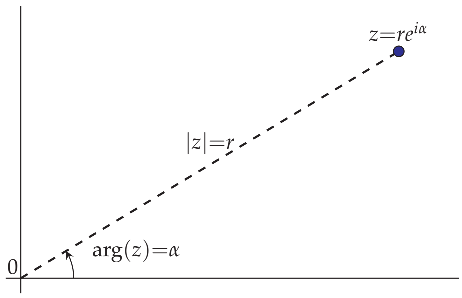
Definition 3.6.1

Lad $z \in \mathbb{C} \setminus \{0\}$ være et komplekst tal forskelligt fra nul. Da kaldes højresiden af ligningen

$$z = |z| \cdot e^{i \arg(z)}$$

for *den polære form* af z.

Hvis $z \neq 0$, kan vi ved brug af z's polære koordinater (r, α) direkte opskrive z på polær form, nemlig som $z = re^{i\alpha}$. Omvendt, givet et udtryk af formen $z = re^{i\alpha}$, hvor $r > 0$ er et positivt, reelt tal, og $\alpha \in]-\pi, \pi]$ er et reelt tal, så kan vi aflæse z's polære koordinater til (r, α). Se Figur 3.7 for en illustration.



Figur 3.7: Den polære form af et komplekst tal z.

Eksempel 3.6.1

Skriv følgende komplekse tal på polær form:

(a) $-1 + i$

(b) $2 + 5i$

(c) e^{7+3i}

(d) $e^{7+3i}/(-1 + i)$

Svar: Man kan for hvert af de givne tal udregne modulus og argument. Når disse er fundet, kan tallet skrives på polær form.

(a) $|-1 + i| = \sqrt{1 + 1} = \sqrt{2}$ og $\arg(-1 + i) = \arctan(1/-1) + \pi = 3\pi/4$. På polær form er tallet dermed givet ved $\sqrt{2}e^{i3\pi/4}$.

(b) $|2 + 5i| = \sqrt{4 + 25} = \sqrt{29}$ og $\arg(2 + 5i) = \arctan(5/2)$. Vi får derfor, at $2 + 5i$ har følgende polære form: $\sqrt{29}e^{i \arctan(5/2)}$.

(c) $e^{7+3i} = e^7 e^{3i}$. Højresiden af denne ligning er allerede tallets polære form, da det er på formen $re^{i\alpha}$ (med $r > 0$ og $\alpha \in \mathbb{R}$). Vi kan aflæse, at modulus af tallet e^{7+3i} er lig med e^7, mens det har et argument lig med 3.

(d) Vi har set i første del af dette eksempel, at $-1+i = \sqrt{2}e^{i3\pi/4}$. Ved at fortsætte derfra får vi, at:
$$\frac{e^{7+3i}}{-1+i} = \frac{e^7 e^{3i}}{\sqrt{2}e^{i3\pi/4}} = \frac{e^7}{\sqrt{2}}\frac{e^{3i}}{e^{i3\pi/4}} = \frac{e^7}{\sqrt{2}}e^{(3-3\pi/4)i}.$$
Det sidste udtryk er den ønskede polære form. Vi kan aflæse, at tallet $e^{7+3i}/(-1+i)$ har modulus $e^7/\sqrt{2}$ og argument $3-3\pi/4$.

I eksemplet så vi, at tallet e^{7+3i} har modulus e^7, mens det har et argument lig 3. Generelt gælder der, at
$$|e^z| = e^{\mathrm{Re}(z)} \quad \text{og} \quad \arg(e^z) = \mathrm{Im}(z). \tag{3.10}$$

I sidste punkt i Eksempel 3.4.1 så vi, at den komplekse eksponentialfunktion ikke er injektiv, da ligningen $e^z = 1$ har flere end én løsning, for eksempel både 0 og $2\pi i$. Lad os benytte det, vi har lært indtil videre, til at undersøge mere generelt, hvordan man løser denne type ligning:

Lemma 3.6.1

Lad $w \in \mathbb{C}$ være et komplekst tal. Hvis $w = 0$, har ligningen $e^z = w$ ingen løsninger. Hvis $w \neq 0$, er løsningerne til ligningen $e^z = w$ netop $z \in \mathbb{C}$ på formen $z = \ln(|w|) + \arg(w)i$, hvor $\arg(w)$ kan være ethvert argument for w.

Bevis. Ligning (3.10) medfører, at $|e^z|$ ikke kan være nul, da $e^{\mathrm{Re}(z)} > 0$ for alle $z \in \mathbb{C}$. Derfor har ligningen $e^z = 0$ ingen løsninger. Antag nu, at $w \neq 0$. Hvis $e^z = w$, medfører Ligning (3.10), at $|w| = |e^z| = e^{\mathrm{Re}(z)}$ og derfor, at $\mathrm{Re}(z) = \ln(|w|)$. Ligeledes ser vi ved at benytte anden del af Ligning (3.10), at $e^z = w$ medfører, at $\arg(w) = \mathrm{Im}(z)$. Bemærk dog, at der er uendeligt mange mulige værdier for $\arg(w)$, da vi altid kan tillægge et heltalsmultiplum af 2π. Indtil videre har vi vist, at hvis $w \neq 0$, så skal enhver løsning til ligningen $e^z = w$ være på formen $z = \ln(|w|) + \arg(w)i$. Omvendt, givet ethvert z, der opfylder $z = \ln(|w|) + \arg(w)i$, hvor $\arg(w)$ er et vilkårligt argument for w, så er $e^z = e^{\ln(|w|)+\arg(w)i} = e^{\ln(|w|)} \cdot e^{i\arg(w)} = |w| \cdot e^{i\arg(w)} = w$, hvor den sidste lighed følger, da $|w|e^{i\arg(w)}$ simpelthen er den polære form af w. $\square$

En direkte konsekvens af dette lemma er, at billedet af den komplekse eksponentialfunktion $\exp : \mathbb{C} \to \mathbb{C}$ med $z \mapsto e^z$ opfylder $\exp(\mathbb{C}) = \mathbb{C} \setminus \{0\}$. Ligningen $e^z = 0$ har ingen løsninger, hvilket medfører, at 0 ikke tilhører billedet, mens lemmaet forklarer for ethvert komplekst tal w forskelligt fra nul, hvordan man finder komplekse tal z, der afbildes i w ved den komplekse eksponentialfunktion.

Vi kan nu genbesøge de polære koordinater og benytte egenskaberne for den komplekse eksponentialfunktion som givet i Sætning 3.4.2 til at bevise følgende sætning.

Sætning 3.6.2
Lad $z_1, z_2 \in \mathbb{C} \setminus \{0\}$ være to komplekse tal, begge forskellige fra nul. Da gælder følgende:

$$|z_1 \cdot z_2| = |z_1| \cdot |z_2|$$

og

$$\arg(z_1 \cdot z_2) = \arg(z_1) + \arg(z_2).$$

Vi har også

$$|z_1/z_2| = |z_1|/|z_2|$$

og

$$\arg(z_1/z_2) = \arg(z_1) - \arg(z_2).$$

Endelig, lad $n \in \mathbb{Z}$ være et heltal og $z \in \mathbb{C} \setminus \{0\}$ et komplekst tal forskelligt fra nul. Da haves

$$|z^n| = |z|^n$$

og

$$\arg(z^n) = n \arg(z).$$

Bevis. Vi viser kun de første to dele af sætningen. Lad os definere $r_1 = |z_1|$, $r_2 = |z_2|$, $\alpha_1 = \arg(z_1)$ og $\alpha_2 = \arg(z_2)$. Ifølge Ligning (3.5) har vi

$$\begin{aligned} z_1 \cdot z_2 &= r_1 \cdot e^{\alpha_1 i} \cdot r_2 \cdot e^{\alpha_2 i} \\ &= r_1 \cdot r_2 \cdot e^{\alpha_1 i} \cdot e^{\alpha_2 i} \\ &= r_1 \cdot r_2 \cdot e^{\alpha_1 i + \alpha_2 i} \\ &= r_1 \cdot r_2 \cdot e^{(\alpha_1 + \alpha_2) i} \end{aligned}$$

Vi benyttede det tredje punkt i Sætning 3.4.2 i den tredje lighed. Vi kan nu konkludere, at

$$|z_1 \cdot z_2| = r_1 \cdot r_2 = |z_1| \cdot |z_2| \quad \text{og} \quad \arg(z_1 \cdot z_2) = \alpha_1 + \alpha_2 = \arg(z_1) + \arg(z_2).$$

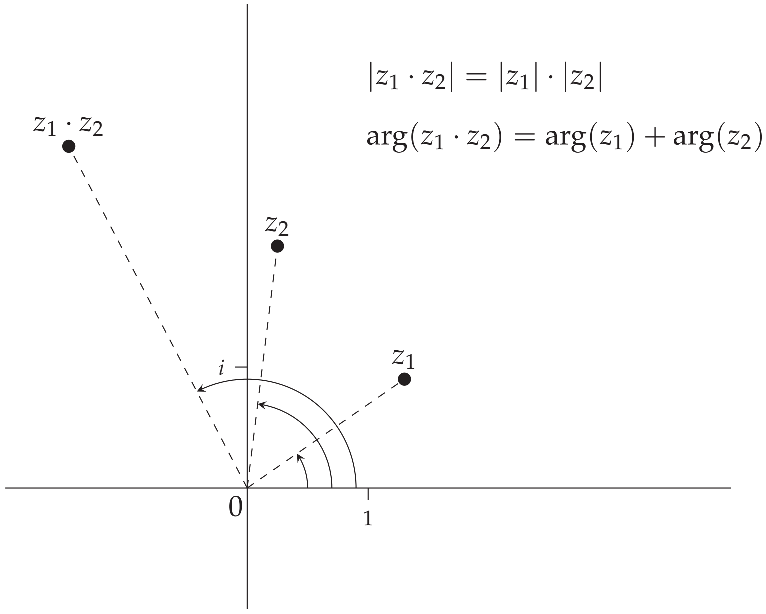
□

Sætning 3.6.2 angiver en geometrisk metode til at beskrive multiplikationen af to komplekse tal: længden af et produkt er produktet af længderne, og argumentet af et produkt er summen af argumenterne (se Figur 3.8).

Den polære form af et komplekst tal kan vise sig særdeles nyttig ved udregninger af heltalspotenser af komplekse tal. Lad os se på et eksempel.

Eksempel 3.6.2
Skriv følgende komplekse tal på rektangulær form. Tip: brug polære former.

(a) $(1+i)^{13}$.



Figur 3.8: Grafisk illustration af Sætning 3.6.2.

(b) $(-1-\sqrt{3}i)^{15}$.

Svar:

(a) Tallet $1+i$ har argument $\pi/4$ og modulus $\sqrt{2}$. Derfor er $1+i = \sqrt{2}\cdot e^{i\pi/4}$. Så

$$\begin{aligned}
(1+i)^{13} &= \left(\sqrt{2}\cdot e^{i\pi/4}\right)^{13} \\
&= \sqrt{2}^{13}\cdot e^{i13\pi/4} \\
&= \sqrt{2}^{13}\cdot(\cos(13\pi/4)+\sin(13\pi/4)i) \\
&= \sqrt{2}^{13}\cdot(\cos(-3\pi/4)+\sin(-3\pi/4)i) \\
&= 64\sqrt{2}\cdot(\cos(-3\pi/4)+\sin(-3\pi/4)i) \\
&= 64\sqrt{2}\cdot\left(-\sqrt{2}/2 - i\sqrt{2}/2\right) \\
&= -64-64i.
\end{aligned}$$

(b) Først udregner vi modulus og argument for $-1-\sqrt{3}i$. Ifølge Sætning 3.3.1 gælder der, at

$$\arg(-1-\sqrt{3}i) = \arctan((-\sqrt{3})/(-1)) - \pi = -2\pi/3$$

og

$$|-1-\sqrt{3}i| = \sqrt{(-1)^2+(-\sqrt{3})^2} = 2.$$

Derfor er $-1-\sqrt{3}i = 2\cdot e^{-i2\pi/3}$. Dermed har vi

$$\begin{aligned}
(-1-\sqrt{3}i)^{15} &= \left(2\cdot e^{-i2\pi/3}\right)^{15}\\
&= 2^{15}\cdot e^{-i30\pi/3}\\
&= 2^{15}\cdot(\cos(-30\pi/3)+\sin(-30\pi/3)i)\\
&= 2^{15}\cdot(\cos(-10\pi)+\sin(-10\pi)i)\\
&= 2^{15}\cdot(\cos(0)+\sin(0)i)\\
&= 2^{15}\cdot(1+0i)\\
&= 2^{15}.
\end{aligned}$$

Kapitel 4

Polynomier

4.1 Definition af polynomier

I dette kapitel vil vi undersøge en bestemt type af funktioner kaldet polynomier. Polynomier vil blive taget op igen senere, når vi skal behandle differentialligninger, eksempler på vektorrum og egenværdier af en matrix. Vi lægger ud med at definere, hvad et polynomium er.

Definition 4.1.1
Et *polynomium* $p(Z)$ i en variabel Z er et udtryk på formen:

$$p(Z) = a_0Z^0 + a_1Z^1 + a_2Z^2 + \cdots + a_nZ^n, \text{ hvor } n \in \mathbb{Z}_{\geq 0} \text{ er et ikke-negativt heltal.}$$

Her betegner symbolerne $a_0, a_1, a_2, \ldots, a_n \in \mathbb{C}$ komplekse tal, der kaldes *koefficienter* i polynomiet $p(Z)$. Udtrykkene $a_0Z^0, a_1Z^1, \ldots, a_nZ^n$ kaldes *led* i polynomiet $p(Z)$. Den største i, for hvilken $a_i \neq 0$, kaldes *graden* af $p(Z)$ og betegnes $\deg(p(Z))$. Den tilsvarende koefficient kaldes *den ledende koefficient*. Mængden af alle polynomier i Z med komplekse koefficienter betegnes $\mathbb{C}[Z]$.

Typisk skrives ikke Z^0, og typisk skrives Z i stedet for Z^1. Et polynomium skrives dermed helt simpelt som $p(Z) = a_0 + a_1Z + a_2Z^2 + \cdots + a_nZ^n$. Et polynomium af grad nul kan fortolkes som en konstant a_0 forskellig fra nul, mens et polynomium af grad ét har formen $a_0 + a_1Z$. Polynomiet, for hvilket alle koefficienter er nul, kaldes *nulpolynomiet* og angives blot med et 0. Det er sædvane at definere graden af nulpolynomiet til at være $-\infty$, minus uendelig.

Per definition bestemmer koefficienterne et polynomium fuldstændigt. Med andre ord: to polynomier $p_1(Z) = a_0 + a_1Z \cdots + a_nZ^n$ af grad n og $p_2(Z) = b_0 + b_1Z \cdots + b_mZ^m$ af grad m er ens, hvis og kun hvis $n = m$, og $a_i = b_i$ for alle i. Rækkefølgen af leddene er ikke vigtig. For eksempel er polynomierne $Z^2 + 2Z + 3$, $Z^2 + 3 + 2Z$ og $3 + 2Z + Z^2$ alle det samme polynomium. Notationen $\mathbb{C}[Z]$ for mængden af alle polynomier med koefficienter fra $\mathbb{C}$ benyttes som standard, men symbolet, der angiver variablen, i vores tilfælde Z, varierer fra bog til bog. Vi har her valgt Z, da vi har brugt z for komplekse tal. Andre mængder af polynomier kan opnås ved at erstatte $\mathbb{C}$ med noget andet. For eksempel vil vi ofte benytte $\mathbb{R}[Z]$, der repræsenterer mængden af alle polynomier med koefficienter fra $\mathbb{R}$. Bemærk, at $\mathbb{R}[Z] \subseteq \mathbb{C}[Z]$, da $\mathbb{R} \subseteq \mathbb{C}$.

Eksempel 4.1.1

Hvilke af følgende udtryk er elementer i mængden $\mathbb{C}[Z]$? Hvis udtrykket er et polynomium, angiv da dets grad og ledende koefficient.

(a) $1 + Z^2$

(b) $Z^{-1} + 1 + Z^3$

(c) i

(d) $\sin(Z) + Z^{12}$

(e) $1 + 2Z + 5Z^{10} + 0Z^{11}$

(f) $1 + Z + Z^{2.5}$

(g) $(1 + Z)^2$

Svar:

(a) $1 + Z^2$ er et polynomium tilhørende $\mathbb{C}[Z]$. Hvis vi vil skrive det på formen $a_0 + a_1 Z + a_2 Z^2 + \cdots + a_n Z^n$ som i Definition 4.1.1, kan vi skrive det som $1 + 0Z + 1Z^2$. Vi ser dermed, at $n = 2$, $a_0 = a_2 = 1$, og $a_1 = 0$. Fordi $a_2 \neq 0$, er polynomiet af grad 2, mens dets ledende koefficient a_2 er lig med 1.

(b) $Z^{-1} + 1 + Z^3$ er ikke et polynomium fra $\mathbb{C}[Z]$ grundet leddet Z^{-1}. Eksponenterne på Z'erne i et polynomiums led må ikke være negative.

(c) Det komplekse tal i kan fortolkes som et polynomium tilhørende $\mathbb{C}[Z]$. Man vælger $n = 0$ og $a_0 = i$ i Definition 4.1.1. Polynomiet i har derfor grad 0 og ledende koefficient i.

(d) $\sin(Z) + Z^{12}$ er ikke et polynomium grundet leddet $\sin(Z)$.

(e) $1 + 2Z + 5Z^{10} + 0Z^{11}$ er et polynomium tilhørende $\mathbb{C}[Z]$. Leddet med eksponenten 11 kan udelades, da koefficienten for Z^{11} er 0. Den højeste potens af Z med en koefficient forskellig fra nul er derfor 10. Dette betyder, at $\deg(1 + 2Z + 5Z^{10} + 0Z^{11}) = 10$, mens dets ledende koefficient er 5.

(f) $1 + Z + Z^{2.5}$ er ikke et polynomium på grund af leddet $Z^{2.5}$. Eksponenterne på Z'erne skal være naturlige tal.

(g) $(2+Z)^2$ er et polynomium tilhørende $\mathbb{C}[Z]$, selvom det ikke umiddelbart er skrevet på formen som i Definition 4.1.1. Det kan nemlig omskrives til denne form, da $(2+Z)^2 = 4+4Z+Z^2 = 4+4Z+1Z^2$. Vi har, at $\deg((2+Z)^2) = 2$. Den ledende koefficient i $(1+Z)^2$ er 1.

Man kan evaluere et givet polynomium $p(Z) \in \mathbb{C}[Z]$ i ethvert komplekst tal $z \in \mathbb{C}$. Mere præcist, hvis $p(Z) = a_0 + a_1 Z + \cdots + a_n Z^n \in \mathbb{C}[Z]$, og $z \in \mathbb{C}$, så kan vi definere $p(z) = a_0 + a_1 \cdot z + \cdots + a_n \cdot z^n \in \mathbb{C}$. På denne måde giver ethvert polynomium $p(Z) \in \mathbb{C}[Z]$ anledning til en funktion $p : \mathbb{C} \to \mathbb{C}$, defineret ved $z \mapsto p(z)$. En funktion $f : \mathbb{C} \to \mathbb{C}$ kaldes en *polynomiefunktion*, hvis der findes et polynomium $p(Z) \in \mathbb{C}[Z]$, således at der for alle $z \in \mathbb{C}$ gælder, at $f(z) = p(z)$. Ligeledes kaldes en funktion $f : \mathbb{R} \to \mathbb{R}$ en polynomiefunktion, hvis der findes et polynomium $p(Z) \in \mathbb{R}[Z]$, således at der for alle $x \in \mathbb{R}$ gælder, at $f(x) = p(x)$.

To polynomier $p_1(Z) = a_0 + a_1 Z \cdots + a_n Z^n$ og $p_2(Z) = b_0 + b_1 Z \cdots + b_m Z^m$ kan ganges sammed ved at leddene $a_i b_j Z^{i+j}$ for $0 \leq i \leq n$ og $0 \leq j \leq m$ lægges sammen. Dette betyder blot, at for at udregne $p_1(Z) \cdot p_2(Z)$ ganges hvert led i $p_1(Z)$ med hvert led i $p_2(Z)$, hvorefter alle de resulterende led lægges sammen. Lad os se på nogle eksempler.

Eksempel 4.1.2

Skriv følgende polynomier på formen angivet i Definition 4.1.1.

(a) $(Z+5) \cdot (Z+6)$.

(b) $(3Z+2) \cdot (3Z-2)$.

(c) $(Z-1) \cdot (Z^2+Z+1)$.

Svar:

(a) $(Z+5) \cdot (Z+6) = Z \cdot (Z+6) + 5 \cdot (Z+6) = Z^2 + 6Z + 5Z + 30 = Z^2 + 11Z + 30$.

(b) $(3Z+2) \cdot (3Z-2) = (3z)^2 - 6Z + 6Z - 2^2 = 9Z^2 - 4$.

(c) I dette eksempel er den eneste forskel fra de to foregående, at der bringes flere led i spil under multiplikationen:

$$\begin{aligned} (Z-1) \cdot (Z^2+Z+1) &= Z \cdot (Z^2+Z+1) - (Z^2+Z+1) \\ &= Z^3 + Z^2 + Z - Z^2 - Z - 1 \\ &= Z^3 - 1. \end{aligned}$$

Bemærk, at hvis et polynomium er et produkt af to andre polynomier, så som $p(Z) = p_1(Z) \cdot p_2(Z)$, så er $\deg p(Z) = \deg p_1(Z) + \deg p_2(Z)$. Med andre ord:

$$p(Z) = p_1(Z) \cdot p_2(Z) \quad \Rightarrow \quad \deg p(Z) = \deg p_1(Z) + \deg p_2(Z). \tag{4.1}$$

Hvis $p(Z) \in \mathbb{C}[Z]$ er et polynomium, så kaldes ligningen $p(z) = 0$ en *polynomiumsligning*. Løsninger til en polynomiumsligning tildeles et særligt navn:

Definition 4.1.2
Lad $p(Z) \in \mathbb{C}[Z]$ være et polynomium. Et komplekst tal $\lambda \in \mathbb{C}$ kaldes en *rod* i $p(Z)$, hvis $p(\lambda) = 0$.

Bemærk, at ifølge definitionen er et komplekst tal en rod i et polynomium $p(Z)$, hvis og kun hvis det er en løsning til polynomiumsligningen $p(z) = 0$.

4.2 Polynomier af anden grad med reelle koefficienter

Vi vil i dette afsnit vise en vigtig anvendelse af de komplekse tal, når vi skal bestemme rødderne i et polynomium $p(Z) \in \mathbb{R}[Z]$ af anden grad. Bemærk, at vi antager, at $p(Z) \in \mathbb{R}[Z]$, så polynomiet $p(Z)$ har reelle koefficienter. Et sådant polynomium $p(Z)$ kan derfor skrives på formen

$$p(Z) = aZ^2 + bZ + c,$$

hvor $a, b, c \in \mathbb{R}$, og $a \neq 0$. For at bestemme dets rødder skal vi løse polynomiumsligningen $az^2 + bz + c = 0$. Nu gælder følgende:

$$\begin{aligned} az^2 + bz + c = 0 &\Leftrightarrow 4a^2z^2 + 4abz + 4ac = 0 \\ &\Leftrightarrow (2az)^2 + 2(2az)b + b^2 = b^2 - 4ac \\ &\Leftrightarrow (2az + b)^2 = b^2 - 4ac. \end{aligned} \tag{4.2}$$

Udtrykket $b^2 - 4ac$ kaldes diskriminanten for polynomiet $aZ^2 + bZ + c$. Vi vil betegne det ved D. Fra Ligning (4.2) følger det, at for at bestemme rødderne i polynomiet $aZ^2 + bZ + c$ skal vi tage kvadratroden af dets diskriminant D. Hvis $D \geq 0$, kan man benytte den sædvanlige kvadratrod, men nu vil vi definere kvadratroden af ethvert reelt tal:

Definition 4.2.1
Lad D være et reelt tal. Da definerer vi

$$\sqrt{D} = \begin{cases} \sqrt{D} & \text{hvis } D \geq 0, \\ i\sqrt{|D|} & \text{hvis } D < 0. \end{cases}$$

Hvis $D \geq 0$, så er $\sqrt{D}$ netop, hvad vi er vant til, og der gælder, at $\sqrt{D}^2 = D$. Hvis $D < 0$, gælder der, at $\sqrt{D}^2 = (i\sqrt{|D|})^2 = i^2\sqrt{|D|}^2 = (-1)|D| = D$. Derfor gælder der for alle reelle tal D, at $\sqrt{D}^2 = D$. Dette er præcis den egenskab, vi gerne vil have, at kvadratrodsymbolet skal have. Desuden kan alle løsninger til ligningen $z^2 = D$ nu angives: de er $z = \sqrt{D}$ og $z = -\sqrt{D}$. Senere, i Sætning 4.4.1, vil vi endda kunne beskrive alle løsningerne til ligninger af formen $z^n = w$ for ethvert $n \in \mathbb{N}$ og $w \in \mathbb{C}$. Vi vender nu tilbage til bestemmelsen af rødderne i polynomiet $p(z) = az^2 + bz + c$. Ved hjælp af den udvidede kvadratrodsdefinition samt Ligning (4.2) får vi, at

$$
\begin{aligned}
az^2+bz+c=0 &\Leftrightarrow (2az+b)^2=b^2-4ac \\
&\Leftrightarrow (2az+b)=\sqrt{b^2-4ac} \quad \vee \quad (2az+b)=-\sqrt{b^2-4ac} \\
&\Leftrightarrow z=\frac{-b+\sqrt{b^2-4ac}}{2a} \quad \vee \quad z=\frac{-b-\sqrt{b^2-4ac}}{2a}.
\end{aligned} \tag{4.3}
$$

Vi opnår den velkendte formel til at løse en ligning af anden grad, men kvadratroden af diskriminanten er nu også defineret, hvis diskriminanten er negativ. Faktisk har vi nu vist følgende sætning.

Sætning 4.2.1

Polynomiet $p(Z)=aZ^2+bZ+c\in\mathbb{R}[Z]$ med $a\neq 0$, har følgende rødder i $\mathbb{C}$:

$$\frac{-b+\sqrt{D}}{2a} \text{ og } \frac{-b-\sqrt{D}}{2a}, \text{ hvor } D=b^2-4ac.$$

For at være mere præcis har polynomiet

(i) to reelle rødder $z=\dfrac{-b\pm\sqrt{D}}{2a}$, hvis $D>0$,

(ii) en reel rod $z=\dfrac{-b}{2a}$, hvis $D=0$,

(iii) to ikke-reelle rødder $z=\dfrac{-b\pm i\sqrt{|D|}}{2a}$, hvis $D<0$.

Beskrivelsen af rødderne i Sætning 4.2.1 er algoritmisk i sin natur. Lad os skrive noget pseudo-kode og opstille algoritmen:

Algoritme 6 for udregning af rødderne i $p(Z)\in\mathbb{R}[Z]$ af anden grad.

Input: $p(Z)\in\mathbb{R}[Z]$ med $\deg(p(Z))=2$

1: $a\leftarrow$ koefficienten af Z^2 i $p(Z)$

2: $b\leftarrow$ koefficienten af Z^1 i $p(Z)$

3: $c\leftarrow$ koefficienten af Z^0 i $p(Z)$

4: $D\leftarrow b^2-4ac$

5: **if** $D\geq 0$ **then**

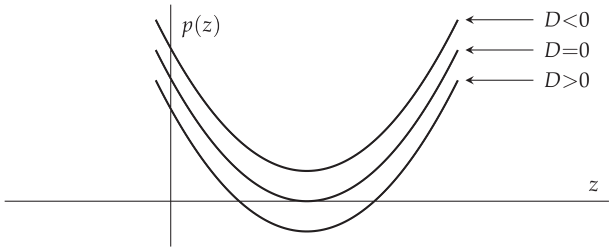
6: **return** $\dfrac{-b+\sqrt{D}}{2a}$ og $\dfrac{-b-\sqrt{D}}{2a}$

7: **else**

8: **return** $\dfrac{-b+i\sqrt{|D|}}{2a}$ og $\dfrac{-b-i\sqrt{|D|}}{2a}$

I Figur 4.1 har vi tegnet graferne for nogle andengradspolynomier. Reelle rødder i et andengradspolynomium svarer til skæringspunkter mellem x-aksen og dets graf. Hvis der ikke er nogen skæringspunkter, har polynomiet ikke reelle rødder men komplekse rødder.

Hvis $D = b^2 - 4ac = 0$, har polynomiumsligningen $az^2 + bz + c = 0$ én løsning, og vi siger i dette tilfælde, at polynomiet har en *dobbeltrod*, eller en rod med multiplicitet to. Hvis $D \neq 0$, siger man, at rødderne hver især har multiplicitet én. Vi ser, at ethvert polynomium af anden grad har to rødder, hvis rødderne tælles med deres multipliciteter. Vi vender tilbage til rødder og multipliciteter i flere detaljer i Afsnit 4.6. Grafen for en polynomiefunktion $f : \mathbb{R} \to \mathbb{R}$, der er dannet af et andengradspolynomium fra $\mathbb{R}[Z]$, skærer den vandrette akse to gange, hvis $D > 0$, én gang, hvis $D = 0$, og slet ikke, hvis $D < 0$. Se Figur 4.1 for en illustration.



Figur 4.1: Et polynomium $p(Z) \in \mathbb{R}[Z]$ af anden grad har to reelle rødder, hvis $D > 0$, en dobbeltrod, hvis $D = 0$, og to komplekse, ikke-reelle rødder, hvis $D < 0$.

Eksempel 4.2.1

Bestem alle komplekse rødder i polynomiet $2Z^2 - 4Z + 10 = 0$.
Svar: Diskriminanten tilhørende polynomiet $2Z^2 - 4Z + 10$ er

$$D = (-4)^2 - 4 \cdot 2 \cdot 10 = -64.$$

Ifølge Definition 4.2.1 får vi så, at

$$\sqrt{D} = \sqrt{-64} = i\sqrt{64} = 8i.$$

Derfor har polynomiumsligningen $2z^2 - 4z + 10 = 0$ to ikke-reelle rødder, nemlig

$$z = \frac{-(-4) + 8i}{2 \cdot 2} = 1 + 2i \quad \vee \quad z = \frac{-(-4) - 8i}{2 \cdot 2} = 1 - 2i.$$

Selvom Sætning 4.2.1 garanterer, at $1 + 2i$ og $1 - 2i$ er rødderne i polynomiet $2Z^2 - 4Z + 10$, så lad os alligevel eftertjekke ved håndregning, at $1 + 2i$ er en rod:

$$\begin{aligned} 2 \cdot (1+2i)^2 - 4 \cdot (1+2i) + 10 &= 2 \cdot (1^2 + 4i + (2i)^2) - 4 \cdot (1+2i) + 10 \\ &= 2 \cdot (1 - 4 + 4i) - 4 \cdot (1+2i) + 10 \\ &= 2 \cdot (-3 + 4i) - 4 \cdot (1+2i) + 10 \\ &= (-6 + 8i) - (4 + 8i) + 10 \\ &= 0. \end{aligned}$$

Vi har nu vist, at $1 + 2i$ ganske rigtigt og præcis som teorien forudsiger, er en rod i $2Z^2 - 4Z + 10$.

4.3 Polynomier med reelle koefficienter

I det foregående afsnit studerede vi polynomier af anden grad med reelle koefficienter. Mange af de polynomier, vi vil støde på senere, vil have reelle koefficienter. I dette afsnit vil vi derfor samle fakta om sådanne polynomier. Kompleks konjugering som introduceret i Definition 3.2.3 vil spille en vigtig rolle. Kompleks konjugering har flere gode egenskaber. Vi oplister nogle af disse i følgende lemma.

Lemma 4.3.1

Lad $z, z_1, z_2 \in \mathbb{C}$ være komplekse tal. Da gælder der, at

(i) $\overline{\overline{z}} = z$,

(ii) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$,

(iii) $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$,

(iv) $\overline{1/z} = 1/\overline{z}$ forudsat, at $z \neq 0$,

(v) $\overline{z^n} = (\overline{z})^n$, hvor $n \in \mathbb{Z}$.

Bevis. Lad os bevise lemmaets andet og tredje punkt. Beviserne af de resterende punkter overlades til læseren. Om en sum af to komplekse tal $z_1 = a + bi$ og $z_2 = c + di$ på rektangulær form gælder der, at

$$\overline{z_1 + z_2} = \overline{(a+c) + (b+d)i} = (a+c) - (b+d)i = (a - bi) + (c - di) = \overline{z_1} + \overline{z_2}.$$

Om et produkt af to komplekse tal $z_1 = a + bi$ og $z_2 = c + di$ på rektangulær form gælder der $z_1 \cdot z_2 = (ac - bd) + (ad + bc)i$. Derfor har vi

$$\overline{z_1 \cdot z_2} = (ac - bd) - (ad + bc)i.$$

Samtidig har vi følgende:

$$\begin{aligned}
\overline{z_1} \cdot \overline{z_2} &= (a - bi) \cdot (c - di) \\
&= ac - adi - bci + (-b) \cdot (-d)i^2 \\
&= ac - (ad + bc)i + bd \cdot (-1) \\
&= ac - bd - (ad + bc)i.
\end{aligned}$$

Dette viser, at $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$. □

Eksempel 4.3.1

Udtryk følgende komplekse tal på rektangulær form.

(a) $\overline{-3 + 6i}$

(b) $\overline{\pi}$

(c) $\overline{-97i}$

Svar:

(a) Jævnfør definitionen af den komplekst konjugerede har vi $\overline{-3+6i} = -3-6i$.

(b) $\overline{\pi} = \overline{\pi + 0i} = \pi - 0i = \pi$. Dette illustrerer det mere generelle faktum, at $\overline{z} = z$, hvis z er et reelt tal.

(c) $\overline{-97i} = -(-97i) = 97i$. Det viser sig mere generelt, at $\overline{z} = -z$ for alle rent imaginære tal.

Kompleks konjugering fungerer også fint sammen med den komplekse eksponentialfunktion.

Lemma 4.3.2
Lad $z \in \mathbb{C}$ være et komplekst tal og $\alpha \in \mathbb{R}$ et reelt tal. Der gælder, at

(i) $\overline{e^z} = e^{\overline{z}}$,

(ii) $\overline{e^{i\alpha}} = e^{-i\alpha}$,

(iii) $\overline{z} = |z|e^{-i\arg(z)}$.

Bevis. Vi beviser her de første to dele af lemmaet. Den tredje del af lemmaet er illustreret i Figur 4.2. Antag, at $z = a + bi$ er z's rektangulære form. Jævnfør definitionen af den komplekse eksponentialfunktion får vi

$$\overline{e^z} = \overline{e^a \cos(b) + e^a \sin(b)i} = e^a \cos(b) - e^a \sin(b)i$$

$$= e^a \cos(-b) + e^a \sin(-b)i = e^{a-bi} = e^{\overline{z}}.$$

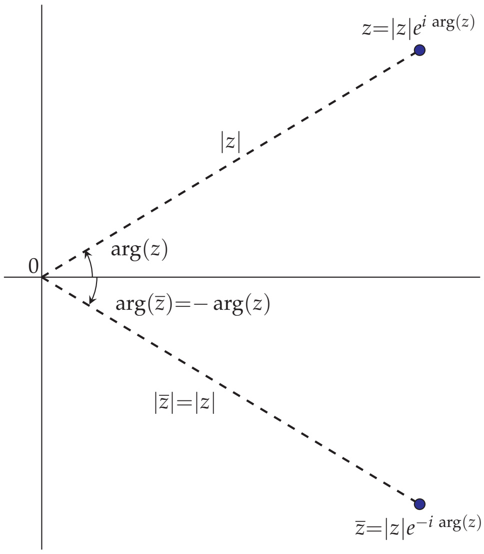
Hvis $z = i\alpha$ (med $\alpha \in \mathbb{R}$), opnår vi særtilfældet

$$\overline{e^{i\alpha}} = e^{\overline{i\alpha}} = e^{-i\alpha}.$$

□

Eksempel 4.3.2
Skriv det komplekse tal $\overline{5e^{i\pi/3}}$ på polær form.
Svar:

$\overline{5e^{i\pi/3}} = \overline{5}\,\overline{e^{i\pi/3}} = 5e^{-i\pi/3}$. Dette illustrerer den tredje del af det foregående lemma, som siger, at $\overline{z} = |z|e^{-i\arg(z)}$.



Figur 4.2: Et komplekst tal z og dets komplekse konjugerede $\overline{z}$ på polær form.

Lad os nu vende tilbage til vores diskussion af polynomier med reelle koefficienter. Grunden til, at vi genopfriskede kompleks konjugering her, er følgende egenskab:

Lemma 4.3.3

Lad $p(Z) \in \mathbb{R}[Z]$ være et polynomium med reelle koefficienter, og lad $\lambda \in \mathbb{C}$ være en rod i $p(Z)$. Da er det komplekse tal $\overline{\lambda} \in \mathbb{C}$ også en rod i $p(Z)$.

Bevis. Lad os opskrive $p(Z) = a_n Z^n + \cdots + a_1 Z + a_0$. Da $p(Z)$ har reelle koefficienter, gælder der, at $a_n, \ldots, a_0 \in \mathbb{R}$. Det er givet, at $\lambda \in \mathbb{C}$ er en rod i $p(Z)$, og derfor gælder der, at

$$0 = a_n \lambda^n + \cdots + a_1 \lambda + a_0.$$

Vi vil nu vise, at $\overline{\lambda}$ også er en rod i $p(Z)$. Først udføres kompleks konjugering af hver side af ligningen, hvilket giver

$$0 = \overline{a_n \lambda^n + \cdots + a_1 \lambda + a_0},$$

hvorefter vi ved anvendelse af egenskaberne givet i Lemma 4.3.1 herpå, får:

$$\begin{aligned}
0 &= \overline{a_n \lambda^n + a_{n-1}\lambda^{n-1} \cdots + a_1 \lambda + a_0} \\
&= \overline{a_n \lambda^n} + \overline{a_{n-1}\lambda^{n-1}} + \cdots + \overline{a_1 \lambda} + \overline{a_0} \\
&= \overline{a_n}\,\overline{\lambda^n} + \overline{a_{n-1}}\,\overline{\lambda^{n-1}} + \cdots + \overline{a_1}\,\overline{\lambda} + \overline{a_0} \\
&= \overline{a_n}(\overline{\lambda})^n + \overline{a_{n-1}}(\overline{\lambda})^{n-1} + \cdots + \overline{a_1}\,\overline{\lambda} + \overline{a_0} \\
&= a_n(\overline{\lambda})^n + a_{n-1}(\overline{\lambda})^{n-1} + \cdots + a_1\,\overline{\lambda} + a_0 \\
&= p(\overline{\lambda})
\end{aligned}$$

I den femte lighed har vi benyttet os af, at koefficienterne i polynomiet $p(Z)$ er reelle tal, så $\overline{a_j} = a_j$ for alle j mellem 0 og n. Vi har nu vist, at $p(\overline{\lambda}) = 0$, og derfor kan vi konkludere, at også $\overline{\lambda}$ er en rod i polynomiet $p(Z)$. $\square$

Lemma 4.3.3 har følgende konsekvens: ikke-reelle rødder i et polynomium med reelle koefficienter forekommer i par. Tag for eksempel polynomiet $2Z^2 - 4Z + 10$. Vi så i Eksempel 4.2.1, at en af dets rødder er $1 + 2i$. Lemma 4.3.3 medfører, at det komplekse tal $1 - 2i$ dermed også er en rod i $2Z^2 - 4Z + 10$. Eksempel 4.2.1 viste, at dette rent faktisk er tilfældet.

4.4 Binomier

I dette afsnit behandler vi polynomier på formen $Z^n - w$ for et naturligt tal $n \in \mathbb{N}$ og et komplekst tal $w \in \mathbb{C}$ forskelligt fra 0. Tallet n er graden af polynomiet $Z^n - w$. Fordi et polynomium på formen $Z^n - w$ kun består af to led, nemlig Z^n og $-w$, kaldes det typisk et *binomium*. Den tilsvarende ligning $z^n = w$ kaldes en *binom ligning*. Vi vil i det følgende give en nøjagtig beskrivelse af alle rødder i et binomium $Z^n - w \in \mathbb{C}[Z]$. Vi vil altså bestemme alle $z \in \mathbb{C}$, der opfylder ligningen $z^n = w$. Det viser sig, at den polære form af det komplekse tal w er til stor hjælp.

Sætning 4.4.1

Lad $w \in \mathbb{C}\backslash\{0\}$. Ligningen $z^n = w$ har præcis n forskellige løsninger, nemlig:

$$z = \sqrt[n]{|w|}e^{i(\frac{\arg(w)}{n}+p\frac{2\pi}{n})}, \; p \in \{0, \dots, n-1\}.$$

Her betegner $\sqrt[n]{|w|}$ det entydige, positive, reelle tal, der opfylder $\left(\sqrt[n]{|w|}\right)^n = |w|$.

Bevis. Ideen i dette bevis er at forsøge at bestemme alle løsninger z til ligningen $z^n = w$ på polær form. Derfor skriver vi $z = |z|e^{iu}$, og vi vil forsøge at bestemme de mulige værdier af $|z|$ og u, således at $z^n = |w|e^{i\alpha}$. Først og fremmest har vi $z^n = (|z|e^{iu})^n = |z|^n e^{inu}$, og dette udtryk skal være lig med $|w|e^{i\alpha}$. Dette gælder, hvis og kun hvis $|w| = |z|^n$, og $e^{inu} = e^{i\alpha}$, eller med andre ord hvis og kun hvis $|w| = |z|^n$, og $e^{i(nu-\alpha)} = 1$. Ligningen $|w| = |z|^n$ har præcis én løsning for $|z| \in \mathbb{R}_{>0}$, nemlig $|z| = \sqrt[n]{|w|}$, mens ligningen $e^{i(nu-\alpha)} = 1$ ifølge Lemma 3.6.1 er opfyldt, hvis og kun hvis $nu - \alpha = \arg(1)$. De mulige argumenter for 1 er ethvert multiplum af 2π, det vil sige $\arg(1) = p2\pi$ for ethvert heltal $p \in \mathbb{Z}$.

Alle løsninger til $z^n = w$ er derfor på formen $z = \sqrt[n]{|w|}e^{i(\frac{\alpha}{n}+p\frac{2\pi}{n})}$, hvor $p \in \mathbb{Z}$. I princippet har vi en løsning for ethvert valg af $p \in \mathbb{Z}$, men når p gennemløber mængden $\{0, \dots, n-1\}$, har vi allerede opnået alle de muligheder, der udgør forskellige løsninger z. $\square$

Når de tegnes i det komplekse talplan, danner løsningerne til ligningen $z^n = w$ hjørnerne af en regulær n-kant med centrum i 0. Lad os illustrere dette i et eksempel.

Eksempel 4.4.1

I dette eksempel vil vi bestemme alle rødder i polynomiet $Z^4 + 8 - i8\sqrt{3}$ og skrive dem på rektangulær form.

Svar: Vi kan benytte Sætning 4.4.1 med $n = 4$ og $w = -(8 - i8\sqrt{3})$. Først opskrives det komplekse tal $-(8 - i8\sqrt{3}) = -8 + i8\sqrt{3}$ på polær form. Vi udregner

$$|-8 + i8\sqrt{3}| = \sqrt{(-8)^2 + (8\sqrt{3})^2} = 16$$

og

$$\arg(-8 + i8\sqrt{3}) = \arctan(8\sqrt{3}/(-8)) + \pi = 2\pi/3,$$

og derfor får vi, at $-8 + i8\sqrt{3} = 16e^{i2\pi/3}$, hvilket er den ønskede polære form. Ifølge Sætning 4.4.1 er alle løsningerne til $z^4 = -8 + i8\sqrt{3}$ givet ved:

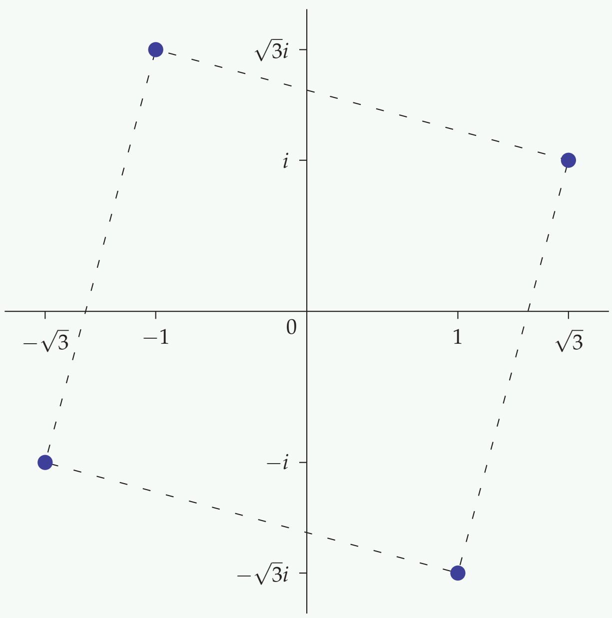
$$z = \sqrt[4]{16}e^{i\left(\frac{2\pi}{3\cdot 4} + p\frac{2\pi}{4}\right)}, \text{hvor } p \text{ kan vælges frit fra mængden } \{0, 1, 2, 3\}, \text{ så}$$

$$z = 2e^{i\frac{\pi}{6}} \quad \vee \quad z = 2e^{i\frac{2\pi}{3}} \quad \vee \quad z = 2e^{i\frac{7\pi}{6}} \quad \vee \quad z = 2e^{i\frac{5\pi}{3}}.$$

Vi skal opskrive disse rødder på rektangulær form. Ved hjælp af formlen $e^{it} = \cos(t) + i\sin(t)$ fås:

$$z = \sqrt{3} + i \quad \vee \quad z = -1 + i\sqrt{3} \quad \vee \quad z = -\sqrt{3} - i \quad \vee \quad z = 1 - i\sqrt{3}.$$

Som nævnt i en bemærkning efter Sætning 4.4.1 udgør disse løsninger hjørnerne af en regulær 4-kant (det vil sige, et kvadrat) med centrum i origo. Dette ses i følgende figur.



Polynomier fra $\mathbb{C}[Z]$ af grad to

I Afsnit 4.2 så vi, hvordan man bestemmer rødderne i polynomier fra $\mathbb{R}[Z]$ af grad to. Vi ved nu også, hvordan man bestemmer rødderne i binome polynomier, og vi er dermed klar til at bestemme rødderne i polynomier af grad to fra $\mathbb{C}[Z]$ uden meget ekstra arbejde. Den vigtigste observation er, at for ethvert polynomium $aZ^2 + bZ + c \in \mathbb{C}[Z]$, hvor $a \neq 0$, er Ligning (4.3) stadig gyldig. Derfor har vi $az^2 + bz + c = 0 \Leftrightarrow (2az + b)^2 = b^2 - 4ac$. Vi ved fra Sætning 4.4.1, at ligningen $s^2 = b^2 - 4ac$ har præcis to løsninger, som vi kan kalde s og $se^{i\pi} = -s$. Vi har dermed $az^2 + bz + c = 0 \Leftrightarrow 2az + b = s \vee 2az + b = -s$. Ved at løse for z får vi derefter følgende resultat:

Sætning 4.4.2

Lad $p(Z) = aZ^2 + bZ + c \in \mathbb{C}[Z]$ være et polynomium af grad to. Yderligere, lad $s \in \mathbb{C}$ være en løsning til den binome ligning $s^2 = b^2 - 4ac$. Da har $p(Z)$ netop følgende rødder:

$$\frac{-b+s}{2a} \text{ og } \frac{-b-s}{2a}.$$

Eksempel 4.4.2

Lad os som et eksempel bestemme rødderne i polynomiet $Z^2 + 2Z + 1 - i$.

Svar: Diskriminanten tilhørende polynomiet $Z^2 + 2Z + 1 - i$ er lig med $2^2 - 4 \cdot 1 \cdot (1 - i) = 4i$. Derfor skal vi først løse den binome ligning $s^2 = 4i$. Vi har $|4i| = 4$ og $\mathrm{Arg}(4i) = \pi/2$. Ved hjælp af Sætning 4.4.1, ser vi, at ligningen $s^2 = 4i$ har løsningerne

$$2 \cdot e^{\pi/4\, i} = 2 \cdot (\cos(\pi/4) + i \sin(\pi/4)) = 2 \cdot (\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} i) = \sqrt{2} + \sqrt{2}\, i$$

og

$$2 \cdot e^{(\pi/4+\pi)\, i} = 2 \cdot (\cos(5\pi/4) + i \sin(5\pi/4)) = 2 \cdot (-\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} i) = -\sqrt{2} - \sqrt{2}\, i.$$

Derfor får vi ved hjælp af Sætning 4.4.2 rødderne i polynomiet $Z^2 + Z + 1 - i$ til at være

$$\frac{-2 + \sqrt{2} + i\sqrt{2}}{2} = -1 + \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} i \quad \text{og} \quad \frac{-2 - \sqrt{2} - i\sqrt{2}}{2} = -1 - \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} i.$$

4.5 Divisionsalgoritmen

I det foregående afsnit så vi, hvordan man bestemmer rødderne i visse typer af polynomier. For at studere røddernes opførsel i mere generelle polynomier begynder vi med følgende observation:

Lemma 4.5.1

Lad $p(Z) \in \mathbb{C}[Z]$ være et polynomium, og antag, at $p(Z) = p_1(Z) \cdot p_2(Z)$ for to polynomier $p_1(Z), p_2(Z) \in \mathbb{C}[Z]$. Yderligere, lad $\lambda \in \mathbb{C}$. Da er λ en rod i $p(Z)$, hvis og kun hvis λ er en rod i $p_1(Z)$ eller i $p_2(Z)$.

Lad os, før vi beviser dette lemma, relatere påstanden i lemmaet til udsagnslogikken fra Kapitel 1 for at tydeliggøre, hvad påstanden siger. En påstand som

"λ er en rod i $p(Z)$, hvis og kun hvis λ er en rod i $p_1(Z)$ eller i $p_2(Z)$"

i en matematisk tekst er blot en måde at formulere et udsagn fra udsagnslogikken på i mere almindeligt sprog. Omformulerer vi til begreber fra udsagnslogikken, opnår vi udsagnet

$$\lambda \text{ er en rod i } p(Z) \quad \Leftrightarrow \quad \lambda \text{ er en rod i } p_1(Z) \vee \lambda \text{ er en rod i } p_2(Z).$$

Vi kan endda gå endnu videre og fjerne alle ord:

$$p(\lambda) = 0 \quad \Leftrightarrow \quad p_1(\lambda) = 0 \vee p_2(\lambda) = 0.$$

Det er en god vane at sikre sig, at man forstår, hvad et matematisk udsagn nøjagtigt betyder, hvis det formuleres i almindeligt sprog. I dette tilfælde er det for eksempel muligt, at λ er en rod i både $p_1(Z)$ og $p_2(Z)$, selvom "eller" ofte benyttes i betydningen "enten den ene eller den anden, men ikke begge". I matematiske tekster har "eller" typisk samme betydning som "$\vee$". Med dette in mente vil vi fortsætte til beviset for lemmaet:

Bevis. Tallet λ er en rod i $p(Z)$, hvis og kun hvis $p(\lambda) = 0$. Da $p(Z) = p_1(Z)p_2(Z)$, er dette ækvivalent med at sige, at $p_1(\lambda)p_2(\lambda) = 0$, og derfor ækvivalent med udsagnet, at $p_1(\lambda) = 0 \vee p_2(\lambda) = 0$. Dette udsagn er logisk ækvivalent med at sige, at λ er en rod i $p_1(Z)$ eller i $p_2(Z)$. □

Ønsker man at bestemme alle rødder i et polynomium, antyder ovenstående lemma, at det altid er en god idé at prøve at skrive polynomiet som et produkt af polynomier af lavere grad. Hvis $p(Z) = p_1(Z) \cdot p_2(Z)$ som i det foregående lemma, siges $p_1(Z)$ og $p_2(Z)$ at være *faktorer* i polynomiet $p(Z)$. Det er derfor nyttigt at have en algoritme, der gør det muligt at afgøre, om et givet polynomium $p_1(Z) \in \mathbb{C}[Z]$ er en faktor i et givet andet polynomium $p(Z) \in \mathbb{C}[Z]$. Ligning (4.1) er her til en vis hjælp, da den medfører, at $p(Z) = p_1(Z) \cdot p_2(Z)$ kun kan være sandt, hvis $\deg p(Z) = \deg p_1(Z) + \deg p_2(Z)$. Specifikt kan $p_1(Z)$ ikke være en faktor i $p(Z)$, hvis $\deg p_1(Z) > \deg p(Z)$. Men dette efterlader stadig tilfældet $\deg p_1(Z) \leq \deg p(Z)$ åbent. Før vi opstiller algoritmen, der løser problemet i sin helhed, vil vi først se på et par eksempler.

Eksempel 4.5.1

(a) Afgør, om polynomiet $Z+3$ er en faktor i polynomiet $2Z^2+3Z-9$.

(b) Afgør, om polynomiet $Z+4$ er en faktor i polynomiet $3Z^3+2Z+1$.

(c) Afgør, om polynomiet $2Z^2+Z+3$ er en faktor i polynomiet $6Z^4+3Z^3+19Z^2+5Z+15$.

Svar:

(a) Vi vil forsøge at bestemme et polynomium $q(Z) \in \mathbb{C}[Z]$, således at $(Z+3)\cdot q(Z) = 2Z^2+3Z-9$. Hvis et sådant $q(Z)$ findes, skal det have graden 1 på grund af Ligning (4.1). Hvis et $q(Z)$ findes, skal det derfor være af formen $q(Z) = b_1Z+b_0$, hvor $b_1, b_0 \in \mathbb{C}$. Vi prøver først at bestemme b_1. Inden vi påbegynder en reducering af produktet $(Z+3)\cdot(b_1Z+b_0)$, kan vi med det samme se, at den højeste eksponent på Z i produktet er 2, og at koefficienten af Z^2 i produktet er b_1. Dette betyder, at $(Z+3)\cdot(b_1Z+b_0) = b_1Z^2 +$ led af grad mindre end 2. Da der allerede gælder, at $(Z+3)\cdot(b_1Z+b_0) = 2Z^2+3Z-9$, har vi dermed, at b_1 skal være 2. Nu hvor vi ved, at $b_1 = 2$, vil vi bestemme b_0. På den ene side skal $(Z+3)\cdot(2Z+b_0) = 2Z^2+3Z-9$ være opfyldt, men på den anden side kan vi skrive $(Z+3)\cdot(2Z+b_0) = (Z+3)\cdot 2Z + (Z+3)\cdot b_0$. Derfor kan vi konkludere, at

$$(Z+3)\cdot b_0 = 2Z^2+3Z-9-(Z+3)\cdot 2Z = -3Z-9. \tag{4.4}$$

Den vigtige observation her er, at vi tidligere har valgt b_1 på en sådan måde, at Z^2-leddet i Ligning (4.4) er væk. Ved at se på koefficienterne af Z, konkluderer vi, at $b_0 = -3$. Vi har vist implikationen $(Z+3)\cdot q(Z) = 2Z^2+3Z-9 \Rightarrow q(Z) = 2Z-3$. En direkte kontrol verificerer, at det faktisk er tilfældet, at $2Z^2+3Z-9 = (Z+3)\cdot(2Z-3)$. Vi kan konkludere, at $Z+3$ ganske rigtigt er en faktor i $2Z^2+3Z-9$. Da -3 er roden i $Z+3$, medfører Lemma 4.5.1, at -3 også er rod i polynomiet $2Z^2+3Z-9$. Og ganske rigtigt har vi $2\cdot(-3)^2+3\cdot(-3)-9 = 0$.

Der findes en mere bekvem måde at opskrive de beregninger, vi netop har udført, på. Det første trin var at udregne b_1 og trække $b_1\cdot(Z+3)$ fra $2Z^2+3Z-9$:

$$\begin{array}{r|l|l} Z+3 & 2Z^2+3Z-9 & 2Z \\ \hline & 2Z^2+6Z & \\ \cline{2-2} & \quad -3Z-9 & \end{array}$$

Den første linje indeholder de polynomier, vi startede med, nemlig $Z+3$ og $2Z^2+3Z-9$, samt alle led i $q(Z)$, som vi fandt frem til i første trin. Den anden linje består af det multiplum af $Z+3$, som vi trak fra $2Z^2+3Z-9$ i Ligning (4.4). Den tredje linje angiver udtrykket $2Z^2+3Z-9-2Z\cdot(Z+3)$ før vores reduceringer. Vi kom også frem til dette i højresiden af Ligning (4.4). Det næste trin var at bestemme b_0. Vi kom frem til, at

$b_0 = -3$, og kan opdatere ovenstående skema som følger:

$$\begin{array}{r|rrr|l}
Z+3 & 2Z^2 & +3Z & -9 & 2Z-3 \\
 & 2Z^2 & +6Z & & \\ \hline
 & & -3Z & -9 & \\
 & & -3Z & -9 & \\ \hline
 & & & 0 &
\end{array}$$

Dette betyder blot, at $2Z^2 + 3Z - 9 - (Z+3) \cdot (2Z-3) = 0$. Dette nul i højresiden kommer fra den sidste linje i ovenstående skema. Konklusionen er derfor, at $Z+3$ er en faktor i polynomiet $2Z^2 + 3Z - 9$. Vi kan endda skrive faktoriseringen ned, da vi har vist, at $2Z^2 + 3Z - 9 = (Z+3) \cdot (2Z-3)$.

(b) Denne gang skal vi undersøge, om polynomiet $Z+4$ er en faktor i polynomiet $3Z^3 + 2Z + 1$. Vi prøver at bestemme et polynomium $q(Z)$, således at $(Z+4) \cdot q(Z) = 3Z^3 + 2Z + 1$. Vi ser, at $q(Z)$ skal have graden 2, det vil sige $q(Z) = b_2 Z^2 + b_1 Z + b_0$, og vi vil bestemme dets tre koefficienter. Fra leddet med den højeste eksponent på Z ser vi, at $b_2 = 3$. Lad os denne gang benytte den skematiske procedure som beskrevet i første del af dette eksempel fra start af. Først får vi:

$$\begin{array}{r|rrrr|l}
Z+4 & 3Z^3 & & +2Z & +1 & 3Z^2 \\
 & 3Z^3 & +12Z^2 & & & \\ \hline
 & & -12Z^2 & +2Z & +1 &
\end{array}$$

Nu kan vi se, at koefficienten af Z i $q(Z)$ skal være -12, og vi får:

$$\begin{array}{r|rrrr|l}
Z+4 & 3Z^3 & & +\ 2Z & +1 & 3Z^2 - 12Z \\
 & 3Z^3 & +12Z^2 & & & \\ \hline
 & & -12Z^2 & +\ 2Z & +1 & \\
 & & -12Z^2 & -48Z & & \\ \hline
 & & & 50Z & +1 &
\end{array}$$

Vi kan nu aflæse, at konstantleddet b_0 i $q(Z)$ skal være 50, og vi får:

$$\begin{array}{r|rrrr|l}
Z+4 & 3Z^3 & & +\ 2Z & +\ 1 & 3Z^2 - 12Z + 50 \\
 & 3Z^3 & +12Z^2 & & & \\ \hline
 & & -12Z^2 & +\ 2Z & +\ 1 & \\
 & & -12Z^2 & -48Z & & \\ \hline
 & & & 50Z & +\ 1 & \\
 & & & 50Z & +200 & \\ \hline
 & & & & -199 &
\end{array}$$

Denne gang får vi ikke et nul i den sidste linje. Hvad ovenstående skema faktisk viser, er, at $3Z^3 + 2Z + 1 - (Z+4) \cdot (3Z^2 - 12Z + 50) = -199$. Dette betyder, at $Z+4$ ikke kan være en faktor i $3Z^3 + 2Z + 1$, da $Z+4$ så også ville være en faktor i $3Z^3 + 2Z + 1 - (Z+4) \cdot (3Z^2 - 12Z + 50) = -199$. Dette er umuligt, da $\deg(Z+4) = 1 > 0 = \deg(-199)$. Bemærk, at -4 ikke er rod i polynomiet $3Z^3 + 2Z + 1$, da $3 \cdot (-4)^3 + 2 \cdot (-4) + 1 = -199$.

(c) Vi angiver kun den fuldendte skematiske procedure denne gang:

$$
\begin{array}{r|r|l}
2Z^2+Z+3 & 6Z^4+3Z^3+19Z^2+5Z+15 & 3Z^2+5 \\
\hline
 & 6Z^4+3Z^3+\ \ 9Z^2 & \\
\cline{2-2}
 & 10Z^2+5Z+15 & \\
 & 10Z^2+5Z+15 & \\
\cline{2-2}
 & 0 &
\end{array}
$$

Konklusionen er, at $6Z^4 + 3Z^3 + 19Z^2 + 5Z + 15 - (2Z^2 + Z + 3) \cdot (3Z^2 + 5) = 0$ og derfor, at $6Z^4 + 3Z^3 + 19Z^2 + 5Z + 15 = (2Z^2 + Z + 3) \cdot (3Z^2 + 5)$. Derfor er $2Z^2 + Z + 3$ en faktor i polynomiet $6Z^4 + 3Z^3 + 19Z^2 + 5Z + 15$.

Algoritmen beskrevet i ovenstående eksempler kaldes *polynomisk division* eller *divisionsalgoritmen* eller af og til også *lang division*. Lad os beskrive den i sin fulde, generelle form.

Der er givet to polynomier $p(Z), d(Z) \in \mathbb{C}[Z]$ som input, hvor $d(Z)$ ikke er nulpolynomiet. Vi ønsker at bestemme to polynomier $q(Z)$ og $r(Z)$ tilhørende $\mathbb{C}[Z]$, således at:

(i) $p(Z) = d(Z)q(Z) + r(Z)$.

(ii) $r(Z) = 0 \quad \vee \quad deg(r(z)) < \deg(d(z))$.

Polynomiet $q(Z)$, som vi forsøger at finde frem til, kaldes *kvotienten* af $p(Z)$ modulo $d(Z)$, mens polynomiet $r(Z)$ kaldes *resten* af $p(Z)$ modulo $d(Z)$. Polynomiet $d(Z)$ er en faktor i $p(Z)$, hvis og kun hvis denne rest er nulpolynomiet. Derfor kan divisionsalgoritmen også benyttes til at afgøre, om et givet polynomium overhovedet kan divideres op i $p(Z)$.

For at bestemme kvotienten og resten begynder vi med følgende skematiske procedure:

$$
\begin{array}{r|c|l}
d(Z) & p(Z) & 0 \\
\hline
\end{array}
$$

Er vi heldige, har vi $\deg p(Z) < \deg d(Z)$. I så fald kan vi stoppe divisionsalgoritmen allerede her og returnere værdierne $q(Z) = 0$ og $r(Z) = p(Z)$. Hvis ikke, så starter vi den lange division og bestemmer et simpelt multiplum af $d(Z)$ med samme grad og ledende koefficient som $p(Z)$. Lad os betegne graden af $d(Z)$ ved m, den ledende koefficient i $d(Z)$ ved d_m og den ledende koefficient i $p(Z)$ ved b. Så har polynomiet $bd_m^{-1}Z^{\deg p(Z)-m} \cdot d(Z)$ nøjagtig samme grad og ledende koefficient som $p(Z)$. Derfor opdaterer vi den skematiske procedure som følger:

$$
\begin{array}{r|l|l}
d(Z) & p(Z) & bd_m{}^{-1}Z^{\deg p(Z)-m} \\
\hline
 & bd_m{}^{-1}Z^{\deg p(Z)-m} \cdot d(Z) & \\
\cline{2-2}
 & p(Z) - bd_m{}^{-1}Z^{\deg p(Z)-m} \cdot d(Z) &
\end{array}
$$

Bemærk, at graden af polynomiet $p(Z) - bd_m{}^{-1}Z^{\deg p(Z)-m} \cdot d(Z)$ er skarpt mindre end $\deg p(Z)$, da de ledende koefficienter i $p(Z)$ og $bd_m{}^{-1}Z^{\deg p(Z)-m} \cdot d(Z)$ er ens og derfor annullerer hinanden, når forskellen mellem de to polynomier udregnes. Skulle det ske, at graden af det resulterende polynomium $p(Z) - bd_m{}^{-1}Z^{\deg p(Z)-m} \cdot d(Z)$ er skarpt mindre end graden af $d(Z)$, er vi færdige og kan returnere som svar polynomierne $p(Z) - bd_m{}^{-1}Z^{\deg p(Z)-m} \cdot d(Z)$ for $r(Z)$ og $bd_m{}^{-1}Z^{\deg p(Z)-m} \cdot d(Z)$ for $q(Z)$. Hvis ikke, så fortsætter vi til næste linje.

Antag nu, at vi har udført proceduren et par gange og er nået frem til følgende:

$$
\begin{array}{r|c|l}
d(Z) & p(Z) & q^*(Z) \\
\hline
 & \ddots & \\
\cline{2-2}
 & r^*(Z) &
\end{array}
$$

Hvis $\deg r^*(Z) < \deg d(Z)$, så er vi færdige og kan returnere $q^*(Z)$ og $r^*(Z)$ som kvotienten og resten, vi leder efter. Ellers udfører vi et trin mere i den lange division og finder et simpelt multiplum af $d(Z)$, der har samme grad og ledende koefficient som $r^*(Z)$. Meget lig det første trin i den lange division, hvor vi nu betegner den ledende koefficient i $r^*(Z)$ ved b, får vi, at polynomiet $bd_m^{-1}Z^{\deg r^*(Z)-m} \cdot d(Z)$ har nøjagtig samme grad og ledende koefficient som $r^*(Z)$. Derfor opdaterer vi den skematiske procedure som følger:

$$
\begin{array}{r|l|l}
d(Z) & p(Z) & q^*(Z) + bd_m{}^{-1}Z^{\deg r^*(Z)-m} \\
\hline
 & \ddots & \\
\cline{2-2}
 & r^*(Z) & \\
 & bd_m{}^{-1}Z^{\deg r^*(Z)-m} \cdot d(Z) & \\
\cline{2-2}
 & r^*(Z) - bd_m{}^{-1}Z^{\deg r^*(Z)-m} \cdot d(Z) &
\end{array}
$$

Da graden af polynomiet i bunden af skemaet falder for hvert trin i iterationen, vil vi efter et endeligt antal trin nå situationen:

$$
\begin{array}{r|c|l}
d(Z) & p(Z) & q(Z) \\
\hline
 & \cdots & \\
\cline{2-2}
 & \ddots & \\
\cline{2-2}
 & r(Z) &
\end{array}
$$

Her er $r(Z)$ enten nulpolynomiet, eller $\deg r(Z) < \deg d(Z)$. Kvotienten og resten er dermed de polynomier $q(Z)$ og $r(Z)$, der ses i skemaet. Lad os opstille denne algoritme i pseudo-kode. For at angive, at algoritmen skal fortsætte, lige så længe $\deg r^*(Z) \geq \deg d(Z)$, benytter vi en såkaldt while-løkke i pseudo-koden.

Algoritme 7 til udførelse af lang division i $\mathbb{C}[Z]$

Input: $p(Z) \in \mathbb{C}[Z]$, $d(Z) \in \mathbb{C}[Z] \setminus \{0\}$.
$m \leftarrow \deg d(Z)$
$d_m \leftarrow$ ledende koefficient i $d(Z)$
$q^*(Z) \leftarrow 0$ og $r^*(Z) \leftarrow p(Z)$
while $\deg r^*(Z) \geq m$ **do**
  $b \leftarrow$ ledende koefficient i $r^*(Z)$
  $q^*(Z) \leftarrow q^*(Z) + bd_m^{-1}Z^{\deg r(Z) - m}$
  $r^*(Z) \leftarrow r^*(Z) - bd_m^{-1}Z^{\deg r^*(Z) - m} \cdot d(Z)$
return $q^*(Z), r^*(Z)$

4.6 Rødder, multipliciteter og faktoriseringer

En overraskende og smuk sætning siger, at ethvert polynomium $p(Z) \in \mathbb{C}[Z]$ af grad mindst 1 har en rod i $\mathbb{C}$. Dette resultat kaldes normalt *algebraens fundamentalsætning*. Lad os her formulere denne sætning, så vi kan referere til den fremover.

Sætning 4.6.1 Algebraens fundamentalsætning

Lad $p(Z) \in \mathbb{C}[Z]$ være et polynomium af grad mindst én. Da har $p(Z)$ en rod $\lambda \in \mathbb{C}$.

Vi vil ikke bevise denne sætning, da beviset er temmelig indviklet. Vi har dog set, at sætningen er sand for polynomier af anden grad i Sætning 4.4.2. Bemærk, at et polynomium ikke nødvendigvis har en reel rod. Som eksempel har polynomiet $Z^2 + 1$ ikke en reel rod men derimod et par af (ikke-reelle) komplekse rødder, nemlig i og $-i$.

Det kan være svært eller direkte umuligt at finde et nyttigt, nøjagtigt udtryk for et givet polynomiums rødder, hvor man ofte må nøjes med en numerisk tilnærmelse af rødderne. Man kan dog på forhånd vide noget om antallet af rødderne, som et polynomium kan have. Vi vil nu se, at hvis et polynomium har graden n, så har det n rødder, hvis vi tæller rødderne på en bestemt måde. Med divisionsalgoritmen som værktøj påbegynder vi her vores undersøgelse af rødder i et polynomium.

Lemma 4.6.2

Lad $p(Z) \in \mathbb{C}[Z]$ være et polynomium af grad $n \geq 1$, og lad $\lambda \in \mathbb{C}$ være et komplekst tal. Tallet λ er en rod i $p(Z)$, hvis og kun hvis $Z - \lambda$ er en faktor i $p(Z)$.

Bevis. Hvis $Z - \lambda$ er en faktor i $p(Z)$, så findes der et polynomium $q(Z) \in \mathbb{C}[Z]$, således at $p(Z) = (Z - \lambda) \cdot q(Z)$. Derfor gælder der, at $p(\lambda) = 0 \cdot q(\lambda) = 0$. Dette viser, at λ er en rod i $p(Z)$, hvis $Z - \lambda$ er en faktor i $p(Z)$.

Antag nu, at λ er en rod i $p(Z)$. Ved hjælp af divisionsalgoritmen kan vi bestemme polynomierne $q(Z)$ og $r(Z)$, således at

$$p(Z) = (Z - \lambda) \cdot q(z) + r(Z), \tag{4.5}$$

hvor enten $r(Z)$ er nulpolynomiet, eller $\deg(r(Z)) < \deg(Z - \lambda) = 1$. Da enten $r(Z) = 0$, eller $\deg(r(Z)) < 1$, ser vi, at $r(Z)$ faktisk er en konstant $r \in \mathbb{C}$. Ved at sætte $Z = \lambda$ i Ligning (4.5) får vi, at $p(\lambda) = r + 0 = r$. Dermed har vi vist, at $p(Z) = (Z - \lambda) \cdot q(Z) + p(\lambda)$. Hvis λ er en rod i $p(Z)$ (det vil sige, $p(\lambda) = 0$), får vi derfor, at $Z - \lambda$ er en faktor i $p(Z)$. $\square$

Ved hjælp af dette lemma kan vi definere multipliciteten af en rod.

Definition 4.6.1

Lad λ være en rod i et polynomium $p(Z)$. Multipliciteten af roden defineres som det største naturlige tal $m \in \mathbb{N}$, således at $(Z - \lambda)^m$ er en faktor i $p(Z)$. λ siges at være rod i $p(Z)$ med *multiplicitet m*.

Bemærk, at Lemma 4.6.2 medfører, at enhver rod i et polynomium har multiplicitet mindst 1. En rod med multiplicitet to kaldes typisk en dobbeltrod.

Eksempel 4.6.1

Afgør, om -3 er en rod i følgende polynomier. Bestem i bekræftende fald dens multiplicitet.

- $p_1(Z) = 2Z^2 + 3Z - 9$.
- $p_2(Z) = Z^2 + 3Z + 1$.
- $p_3(Z) = Z^3 + 3Z^2 - 9Z - 27$.
- $p_4(Z) = (2Z^2 + 3Z - 9) \cdot (Z^3 + 3Z^2 - 9Z - 27) = 2Z^5 + 9Z^4 - 18Z^3 - 108Z^2 + 243$.

Svar:

(a) Vi har $p_1(-3) = 18 - 9 - 9 = 0$. Derfor er -3 en rod i polynomiet $2Z^2 + 3Z - 9$. Vi har i Eksempel 4.5.1 set, at $2Z^2 + 3Z - 9 = (Z + 3) \cdot (2Z - 3)$. Dette betyder, at multipliciteten af roden -3 er lig med 1. Vi kan også se, at faktoren $2Z - 3$ giver anledning til endnu en rod i $p_1(Z)$, nemlig roden $3/2$. Denne rod har også multiplicitet 1.

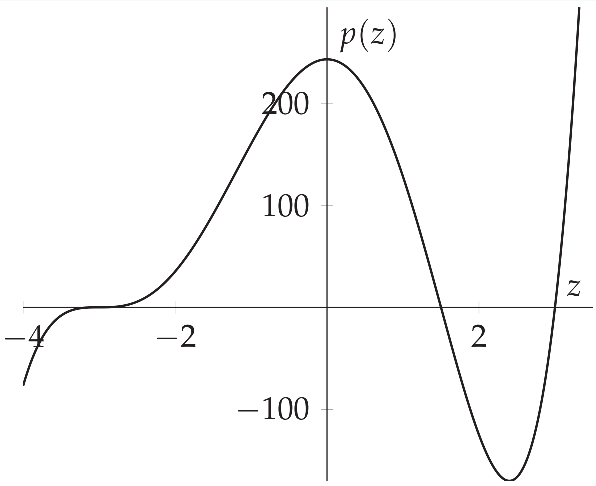
(b) Vi har $p_2(-3) = 1$. Derfor er -3 ikke en rod i $p_2(Z)$.

(c) Denne gang har vi $p_3(-3) = 0$, så -3 er en rod i $p_3(Z)$. Ved brug af divisionsalgoritmen får vi:

$$
\begin{array}{r|r|l}
Z+3 & Z^3+3Z^2-9Z-27 & Z^2-9 \\
\hline
 & Z^3+3Z^2 & \\
\cline{2-2}
 & -9Z-27 & \\
 & -9Z-27 & \\
\cline{2-2}
 & 0 &
\end{array}
$$

Derfor gælder der, at $Z^3 + 3Z^2 - 9Z - 27 = (Z+3) \cdot (Z^2 - 9)$. Tallet -3 er også en rod i polynomiet $Z^2 - 9$, så multipliciteten af roden -3 er mindst 2. Faktisk gælder der, at $Z^2 - 9 = (Z+3) \cdot (Z-3)$, så $Z^3 + 3Z^2 - 9Z - 27 = (Z+3) \cdot (Z^2 - 9) = (Z+3)^2 \cdot (Z-3)$, hvilket betyder, at roden -3 i $p_3(Z)$ har multiplicitet 2. Vi viste også, at 3 er en rod i $p_3(Z)$, og at denne rod har multiplicitet 1.

(d) Vi har $p_4(Z) = p_1(Z)p_3(Z)$. Fra første og tredje punkt i dette eksempel får vi, at $p_4(Z) = (Z+3)^3 \cdot (2Z-3) \cdot (Z-3)$. Dette betyder, at roden -3 har multiplicitet 3. Vi ser også, at tallene $3/2$ og 3 er rødder i $p_4(Z)$, begge med multiplicitet 1. Grafen for den reelle polynomiumsfunktion, som $p_4(Z)$ giver anledning til, ses i Figur 4.3.



Figur 4.3: Grafen for polynomiumsfunktionen $p : \mathbb{R} \to \mathbb{R}$, hvor $p(z) = 2z^5 + 9z^4 - 18z^3 - 108z^2 + 243$.

Ovenstående eksempel illustrerer, at der er en én-til-én-korrespondance mellem faktorer af grad én i et polynomium og rødderne i et polynomium. Algebraens fundamentalsætning (Sætning 4.6.1) siger, at ethvert polynomium af grad mindst 1 har en rod. Dette har følgende konsekvens:

Sætning 4.6.3

Lad $p(Z) = a_n Z^n + a_{n-1} Z^{n-1} + \cdots + a_1 Z + a_0$ være et polynomium af grad $n > 0$. Da eksisterer der $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, således at

$$p(Z) = a_n \cdot (Z - \lambda_1) \cdots (Z - \lambda_n).$$

Bevis. Ifølge algebraens fundamentalsætning eksisterer der en rod $\lambda_1 \in \mathbb{C}$ i polynomiet $p(Z)$. Ved hjælp af Lemma 4.6.2 kan vi skrive $p(Z) = (Z - \lambda_1)q_1(Z)$ for et vist polynomium $q_1(Z)$. Bemærk, at $\deg(q_1(Z)) = \deg(p(Z)) - 1$. Hvis $q_1(Z)$ er en konstant, er vi færdige. Ellers kan vi anvende algebraens fundamentalsætning på polynomiet $q_1(Z)$ og bestemme en rod $\lambda_2 \in \mathbb{C}$ i $q_1(Z)$. Igen ved hjælp af Lemma 4.6.2 kan vi skrive $q_1(Z) = (Z - \lambda_2) \cdot q_2(Z)$. Dette medfører, at $p(Z) = (Z - \lambda_1) \cdot (Z - \lambda_2) \cdot q_2(Z)$. Ved at fortsætte på denne måde kan vi skrive $p(Z)$ som et produkt af polynomier af grad én på formen $Z - \lambda$ gange en konstant c. Da den ledende koefficient i $p(Z)$ er a_n, er denne konstant c lig med a_n. □

Eksempel 4.6.2

Som et eksempel genbesøger vi polynomiet $p_4(Z) = 2Z^5 + 9Z^4 - 18Z^3 - 108Z^2 + 243$ fra Eksempel 4.6.1. Vi ønsker at omskrive dette polynomium som i Sætning 4.6.3. Vi har allerede set, at $p_4(Z) = (Z+3)^3 \cdot (2Z - 3) \cdot (Z - 3)$. Ved at faktorisere 2 ud fra faktoren $2Z - 3$ får vi:

$$p_4(Z) = 2 \cdot (Z+3)^3 \cdot (Z - 3/2) \cdot (Z-3) = 2 \cdot (Z+3) \cdot (Z+3) \cdot (Z+3) \cdot (Z-3/2) \cdot (Z-3).$$

I notationen fra Sætning 4.6.3 ser vi, at $\lambda_1 = -3$, $\lambda_2 = -3$, $\lambda_3 = -3$, $\lambda_4 = 3/2$, og $\lambda_5 = 3$. Dette illustrerer endnu en gang, at multipliciteterne af rødderne -3, $3/2$ og 3 er henholdsvis 3, 1 og 1. Bemærk, at summen af alle multipliciteter er lig med 5, hvilket er graden af $p_4(Z)$.

Faktisk gælder det altid, at summen af alle multipliciteter af rødderne i et polynomium er lig med dets grad. I ord kan man derfor omformulere Sætning 4.6.3 som følger: et polynomium af grad $n \geq 1$ har nøjagtigt n rødder, hvis rødderne tælles med multipliciteter. For polynomier i $\mathbb{R}[Z]$ har Sætning 4.6.3 følgende konsekvens:

Korollar 4.6.4

Ethvert polynomium $p(Z) \in \mathbb{R}[Z]$ af grad mindst én, kan skrives som produktet af polynomier af grad én og grad to fra $\mathbb{R}[Z]$.

Bevis. Ifølge Sætning 4.6.3 kan ethvert polynomium $p(Z)$, der ikke er nulpolynomiet, skrives som produktet af den ledende koefficient i $p(Z)$ og faktorer af grad én på formen $Z - \lambda$. Tallet $\lambda \in \mathbb{C}$ er en rod i polynomiet $p(Z)$. Hvis vi anvender dette på et polynomium $p(Z)$ med reelle koefficienter, ser vi, at mens den ledende koefficient er et reelt tal, behøver roden λ ikke at være et reelt tal. Men enhver reel rod λ giver anledning til en faktor af grad én med reelle koefficienter, nemlig $Z - \lambda$.

Lad nu $\lambda \in \mathbb{C} \setminus \mathbb{R}$ være en rod i $p(Z)$. Lad os skrive $\lambda = a + bi$ på rektangulær form. Da $\lambda \notin \mathbb{R}$, ved vi, at $b \neq 0$. Lemma 4.3.3 medfører, at tallet $\overline{\lambda} = a - bi$ også er en rod i $p(Z)$. Desuden er $\lambda \neq \overline{\lambda}$, da $b \neq 0$. Derfor er $Z - \lambda$ og $Z - \overline{\lambda}$ to forskellige faktorer i $p(Z)$, hvis vi arbejder i $\mathbb{C}[Z]$. Nu er ideen at multiplicere faktorerne $Z - \lambda$ og $Z - \overline{\lambda}$ sammen, da det viser sig, at $(Z-\lambda)\cdot(Z-\overline{\lambda})$ har reelle koefficienter. Vi får

$$\begin{aligned}(Z-\lambda)\cdot(Z-\overline{\lambda}) &= Z^2-(\lambda+\overline{\lambda})Z+\lambda\overline{\lambda}\\ &= Z^2-(a+bi+a-bi)Z+(a+bi)\cdot(a-bi)\\ &= Z^2-2aZ+(a^2+b^2),\end{aligned}$$

der som forventet er et polynomium af grad to i $\mathbb{R}[Z]$, da dets koefficienter er reelle tal. På denne måde kan vi omdanne faktoriseringen af $p(Z)$ i $\mathbb{C}[Z]$ fra Sætning 4.6.3 til en faktorisering af $p(Z)$ i $\mathbb{R}[Z]$ ved brug af faktorer af første og anden grad med reelle koefficienter. $\square$

Eksempel 4.6.3

Skriv følgende polynomier som et produkt af polynomier af grad én og grad to med reelle koefficienter.

(a) $p_1(Z)=Z^3-Z^2+Z-1$

(b) $p_2(Z)=Z^4+4$

Svar:

(a) Tallet 1 er en rod i $p_1(Z)$, da $p(1)=0$. Ved hjælp af divisionsalgoritmen kan man vise, at $p_1(Z)=(Z-1)\cdot(Z^2+1)$. Polynomiet Z^2+1 har ingen reelle rødder og kan derfor ikke faktoriseres yderligere indenfor de reelle tal (indenfor de komplekse tal kan man godt: $Z^2+1=(Z+i)\cdot(Z-i)$). Den ønskede faktorisering er derfor:

$$Z^3-Z^2+Z-1=(Z-1)\cdot(Z^2+1).$$

(b) Ved brug af teorien i Afsnit 4.4 kan vi bestemme alle rødder i polynomiet Z^4+4 til at være $1+i, 1-i, -1+i$ og $-1-i$. Derfor har vi, at

$$Z^4+4=(Z-(1+i))\cdot(Z-(1-i))\cdot(Z-(-1+i))\cdot(Z-(-1-i)).$$

Som i beviset for Korollar 4.6.4 kan vi gange par af komplekst konjugerede faktorer sammen for at slippe af med de komplekse koefficienter. Derefter har vi, at

$$(Z-(1+i))\cdot(Z-(1-i))=Z^2-2Z+2$$

og

$$(Z-(-1+i))\cdot(Z-(-1-i))=Z^2+2Z+2.$$

Den ønskede faktorisering af Z^4+4 er derfor

$$Z^4+4=(Z^2-2Z+2)\cdot(Z^2+2Z+2).$$

Kapitel 5

Rekursion og induktion

5.1 Eksempler på rekursivt definerede funktioner

I dette afsnit introducerer vi ideen om en rekursivt defineret funktion. Pointen i en *rekursion* i denne sammenhæng er simpelthen at definere en funktion eller et udtryk ved hjælp af selv samme funktion eller udtryk for andre inputværdier. Lad os lægge ud med et eksempel:

Eksempel 5.1.1

Fakultets-funktionen $\mathrm{fac} : \mathbb{N} \to \mathbb{N}$ er defineret ved $n \mapsto 1 \cdot 2 \cdots n$. Den afbilder n i produktet af de første n positive heltal. Man vil typisk også skrive $n!$ i stedet for $\mathrm{fac}(n)$. Vi har for eksempel $\mathrm{fac}(1) = 1$, $\mathrm{fac}(2) = 1 \cdot 2 = 2$, $\mathrm{fac}(3) = 1 \cdot 2 \cdot 3 = 6$, $\mathrm{fac}(4) = 1 \cdot 2 \cdot 3 \cdot 4 = 24$ og så videre. Bemærk nu, at hvis vi vil beregne den næste værdi, $\mathrm{fac}(5)$, kan vi udnytte, at vi allerede ved, hvad $\mathrm{fac}(4)$ er. Faktisk har vi:

$$\mathrm{fac}(5) = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = (1 \cdot 2 \cdot 3 \cdot 4) \cdot 5 = \mathrm{fac}(4) \cdot 5 = 24 \cdot 5 = 120.$$

Generelt, hvis vi for et valgt $n > 1$ allerede har udregnet $\mathrm{fac}(n-1)$, kan vi udregne værdien af $\mathrm{fac}(n)$ ved $\mathrm{fac}(n) = \mathrm{fac}(n-1) \cdot n$. Dette fører til følgende algoritmiske beskrivelse af fakultetsfunktionen:

Algoritme 8 $\mathrm{fac}(n)$

Input: $n \in \mathbb{Z}_{\geq 1}$.
if $n = 1$ **then**
  return 1
else
  return $\mathrm{fac}(n-1) \cdot n$.

Denne algoritme benytter sig selv til at udregne $\mathrm{fac}(n)$. Mere udpenslet, hvis $n = 1$, så returneres straks 1 som værdien for $\text{fac}(1)$ som angivet i linje 2 i algoritmen. Grafisk kan vi illustrere dette som følger:

$\text{fac}(1)$

$$\text{fac}(1) = 1$$

returnér 1

Hvis $n = 2$, så vil algoritmen hoppe til linje 4 og forsøge at returnere $\text{fac}(2-1) \cdot 2$. Men dette kræver, at værdien af $\text{fac}(1)$ først udregnes. Derfor vil algoritmen nu starte forfra, denne gang for værdien 1. Vi har allerede set, at algoritmen returnerer 1 i det tilfælde. Da algoritmen er kommet frem til den konklusion, at $\text{fac}(1) = 1$, kan den genbesøge linje 4 og udregne, at $\text{fac}(2) = \text{fac}(1) \cdot 2 = 1 \cdot 2 = 2$. Dermed kan algoritmen nu returnere 2. Grafisk er situationen som følger:

$\text{fac}(2)$

$$\text{fac}(2) = \text{fac}(1) \cdot 2$$

$\text{fac}(1)$

$$\text{fac}(1) = 1$$

returnér 1

$$\text{fac}(2) = 1 \cdot 2 = 2$$

returnér 2

For større værdier af n vil flere "bokse inde i andre bokse" dukke op, da algoritmen nu skal benytte sig selv adskillige gange og udregne output for de lavere inputværdier $n-1, n-2, \ldots, 1$, før den kan returnere det endelige output. Følgende grafiske repræsentation viser, hvad der sker, når denne algoritme får inputværdien $n = 5$:

$\mathrm{fac}(5)$

$$\mathrm{fac}(5) = \mathrm{fac}(4) \cdot 5$$

$\mathrm{fac}(4)$

$$\mathrm{fac}(4) = \mathrm{fac}(3) \cdot 4$$

$\mathrm{fac}(3)$

$$\mathrm{fac}(3) = \mathrm{fac}(2) \cdot 3$$

$\mathrm{fac}(2)$

$$\mathrm{fac}(2) = \mathrm{fac}(1) \cdot 2$$

$\mathrm{fac}(1)$

$$\mathrm{fac}(1) = 1$$

returnér 1

$$\mathrm{fac}(2) = 1 \cdot 2 = 2$$

returnér 2

$$\mathrm{fac}(3) = 2 \cdot 3 = 6$$

returnér 6

$$\mathrm{fac}(4) = 6 \cdot 4 = 24$$

returnér 24

$$\mathrm{fac}(5) = 24 \cdot 5 = 120$$

returnér 120

Da algoritmen benytter sig selv under gennemførslen (i algoritmiske termer siger man typisk, at algoritmen *kalder* sig selv), kaldes den en *rekursiv* algoritme. En rekursiv algoritme er simpelthen en algoritme, der kan kalde sig selv for visse inputværdier for at gennemføre udregningen af sin endelige outputværdi. Rekursioner forekommer også i matematikken. Vi har herunder angivet en rekursiv definition af fakultetsfunktionen fra vores eksempel:

$$\mathrm{fac}(n) = \begin{cases} 1 & \text{hvis } n = 1, \\ \mathrm{fac}(n-1) \cdot n & \text{hvis } n \geq 2. \end{cases} \tag{5.1}$$

Dette eksempel illustrerer princippet i en rekursiv definition: at definere de værdier, en funktion tager, ved brug af funktionen selv. Bemærk i øvrigt, at det også er sædvane at definere $0! = 1$, men det er en anden sag. Her er et andet eksempel på en rekursivt defineret funktion: lad $z \in \mathbb{C}$ være et komplekst tal, og definér $f : \mathbb{N} \to \mathbb{C}$ rekursivt som:

$$f(n) = \begin{cases} z & \text{hvis } n = 1, \\ f(n-1) \cdot z & \text{hvis } n \geq 2. \end{cases} \tag{5.2}$$

Da er $f(1) = z$, da dette svarer til tilfældet $n = 1$ i den rekursive definition. Som det næste er $f(2) = f(1) \cdot z$, da dette er, hvad den rekursive definition angiver for $n = 2$. Ved at udnytte, at vi allerede har bestemt, at $f(1) = z$, kan vi konkludere, at $f(2) = f(1) \cdot z = z \cdot z$. Endelig, ved at udnytte at $z \cdot z = z^2$, ser vi, at $f(2) = f(1) \cdot z = z \cdot z = z^2$. Ligeledes er $f(3) = z^3$. Derfor er det helt rimeligt at *definere* udtrykket z^n for ethvert naturligt tal n rekursivt som $f(n)$. I tidligere kapitler har vi benyttet n'te potenser af komplekse tal flere gange. Nu har vi en mere formel definition af det. I denne forbindelse er det også almindeligt at definere $z^0 = 1$ og $z^{-n} = 1/z^n$ for ethvert naturligt tal n. Vi har hermed præcist defineret, hvad z^n betyder for ethvert heltal $n \in \mathbb{Z}$.

Når man forsøger at definere en funktion rekursivt, skal man efterfølgende sikre sig, at en sådan rekursiv beskrivelse definerer funktionen for alle værdier fra dens definitionsmængde. For de funktioner, der er defineret i Ligningerne (5.1) og (5.2), kan du finde en redegørelse herfor i Eksempel 5.4.2, men vær velkommen til at springe eksemplet over under første gennemlæsning. For nu vil vi vise et eksempel på et forsøg på en rekursiv beskrivelse, der ikke fungerer. Lad $g : \mathbb{N} \to \mathbb{C}$ være en funktion, og antag at

$$g(n) = \begin{cases} 1 & \text{hvis } n = 1, \\ g(n+1) & \text{hvis } n \geq 2. \end{cases}$$

Jævnfør definitionen ser vi, at $g(1) = 1$, men vi har ikke nok information til at bestemme, hvad $g(2)$ er. Hvis vi anvender den rekursive definition, ville vi bare nå frem til, at $g(2) = g(3)$. Derefter vil vi forsøge at beregne $g(3)$, men rekursionen giver os blot $g(3) = g(4)$. Fortsætter vi sådan, får vi blot, at $g(2) = g(3) = g(4) = g(5) = \cdots$, og vi finder aldrig ud af, hvad $g(2)$ rent faktisk er.

Som et sidste eksempel på en rekursiv definition vil vi se på de berømte Fibonacci-tal.

Eksempel 5.1.2

Lad os tage fat på en rekursiv definition, der ser en smule anderledes ud. Vi skal definere en funktion $F : \mathbb{N} \to \mathbb{N}$ rekursivt, hvis værdier $F(1), F(2), F(3), F(4), \ldots$ kaldes *Fibonacci-tal*:

$$F(n) = \begin{cases} 1 & \text{hvis } n = 1, \\ 1 & \text{hvis } n = 2, \\ F(n-1) + F(n-2) & \text{hvis } n \geq 3. \end{cases} \tag{5.3}$$

Lad os se, hvordan denne definition fungerer i praksis ved at udregne de første par Fibonacci-tal. Først og fremmest er $F(1) = 1$, da vi med $n = 1$ skal se på første linje i Ligning (5.3). Hvis $n = 2$, gælder den anden linje i Ligning (5.3), så $F(2) = 1$. For $n = 3$ gælder den tredje linje i Ligning (5.3), og vi får, at $F(3) = F(2) + F(1) = 1 + 1 = 2$. Ligeledes for $n = 4$ får vi, at $F(4) = F(3) + F(2) = 2 + 1 = 3$ ved at udnytte, at vi allerede har udregnet, at $F(3) = 2$.

Når man arbejder med en sekvens af tal, som for eksempel Fibonacci-tallene, vil man typisk ændre en smule i notationen: i stedet for at skrive $F(n)$, skriver man ofte F_n. I denne notation har vi $F_1 = 1$, $F_2 = 1$, $F_3 = 2$, $F_4 = 3$ og så videre. Det viser sig, at det er muligt at udlede et lukket formeludtryk for Fibonacci-tallene:

$$F_n = \frac{1}{\sqrt{5}} \cdot \left(\frac{1+\sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \cdot \left(\frac{1-\sqrt{5}}{2} \right)^n. \tag{5.4}$$

Vi vil senere i dette kapitel få en forklaring på, hvordan vi er kommet til dette udtryk.

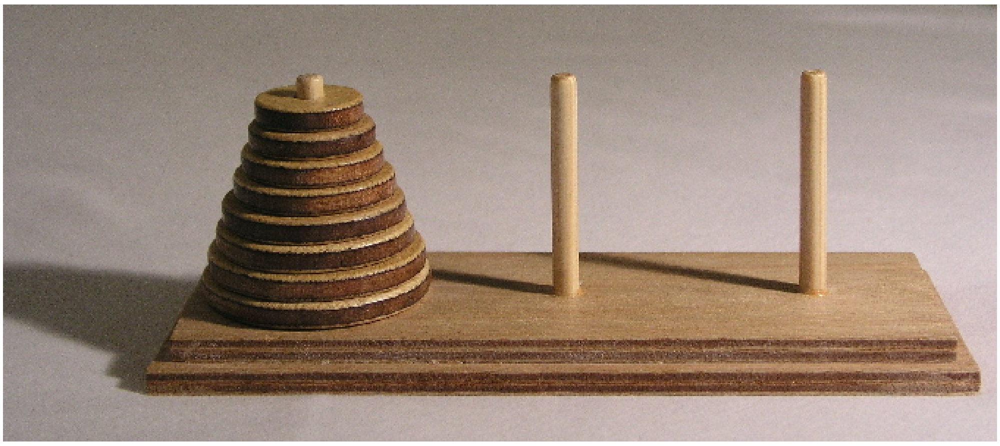
5.2 Hanois tårne

I dette afsnit fortsætter vi med at illustrere nyttigheden af en rekursiv tankegang ved at analysere en gåde kaldet *Hanois tårne.* Hanois tårne er en gåde, der foregår på et bræt påmonteret tre lodrette pinde med samme længder og størrelser. Man har et antal cirkulære skiver til rådighed, alle med forskellig diameter og med et hul i midten, så de kan placeres på en pind. Gådens startopstilling er, at alle skiver er stablet på den første pind. Skiven med den største diameter er nederst i stakken, og de andre skiver følger derpå i faldende diameterstørrelse opefter. Antallet af skiver kan variere. For et eksempel med otte skiver, se Figur 5.1. Billedet i denne figur er fra Wikipedia; se https://commons.wikimedia.org/wiki/File:Tower_of_Hanoi.jpeg for flere detaljer.

Målet med gåden er at flytte stakken af skiver fra den første til den tredje pind, hvor de skal være stablet på samme måde, altså fra stor til lille set fra bunden. Udfordringen er dog, at man skal følge tre regler, når man flytter skiver:

- Kun én skive må flyttes ad gangen.
- Kun en skive øverst på en stak må flyttes.

Figur 5.1: Hanois tårne med otte skiver.



- En skive må kun placeres på en større skive.

Hvis der kun er meget få skiver, er det ikke svært at løse opgaven. Hvis der er mange skiver, bliver spillet mere kompliceret, og det er a priori ikke engang klart, om der altid findes en løsning. For at komme i gang kan vi se på nogle eksempler med kun ganske få skiver. Hvis der kun er én skive, kan vi løse opgaven i ét træk:

Flyt skive 1 fra pind 1 til pind 3

Hvis der er to skiver, kan gåden løses med tre flytninger:

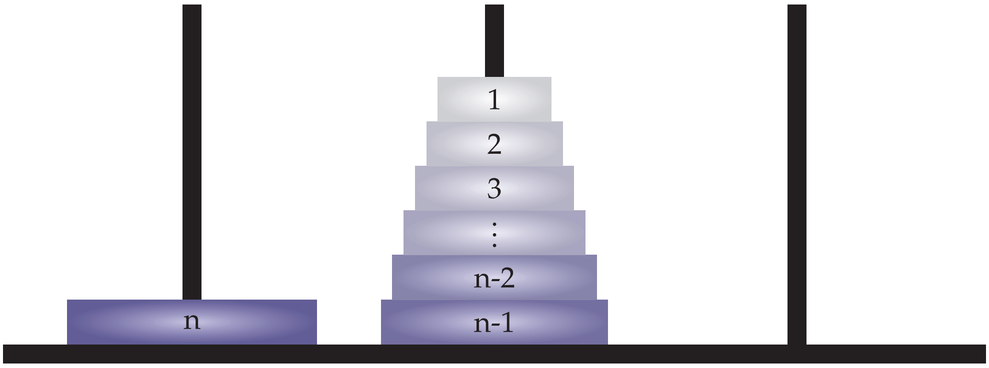
Flyt skive 1 fra pind 1 til pind 2

Flyt skive 2 fra pind 1 til pind 3

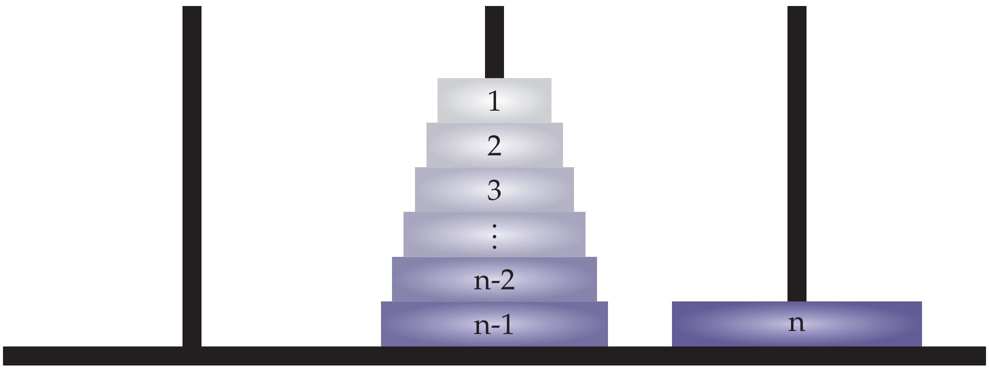
Flyt skive 1 fra pind 2 til pind 3

Med tre skiver er det stadig ikke så svært at løse gåden ved at prøve sig frem, men hvad nu hvis der er ti skiver, eller hundrede? Lad os prøve at tænke på en rekursiv måde for at finde frem til en løsningsmetode. Vi ved allerede, hvordan vi løser opgaven, hvis der kun er én skive (og også hvis der er to skiver). Måske kan vi, som med fakultetsfunktionen, finde ud af, hvad vi skal gøre i tilfælde af et større antal skiver, lad os sige n skiver, hvis vi allerede ved, hvad vi skal gøre, hvis vi har færre end n skiver. Antag derfor, at $n \geq 2$ er et naturligt tal, og at vi allerede ved, hvordan vi løser opgaven, hvis der er $n-1$ skiver. Dette betyder, at vi ved, hvordan vi flytter en stak af $n-1$ skiver fra én pind til en anden. I så fald virker følgende strategi til at flytte n skiver:

Da vi ved, hvordan vi flytter $n-1$ skiver fra én pind til en anden, flytter vi nu alle skiver fra skive nummer 1 til nummer $n-1$ i stakken fra pind 1 til pind 2.



Dernæst flytter vi skive n fra pind 1 til pind 3. Dette tager kun ét træk.



Igen anvender vi vores givne viden om, hvordan man flytter $n-1$ skiver, til at flytte stakken af skiver indeholdende skive nummer 1 til nummer $n-1$ fra pind 2 til pind 3.

Dette viser, at gåden kan løses rekursivt! Og vi har endda vist, at der findes en løsning uanset antallet af skiver.

5.3 Summationssymbolet

Hvis n er et naturligt tal, og $z_1, \ldots, z_n$ er komplekse tal, kan man angive deres sum med et udtryk som $z_1 + z_2 + \cdots + z_n$ eller $z_1 + \cdots + z_n$. Dog er det ofte nemmere med den mere kompakte notation for dette: $\sum_{k=1}^{n} z_k$. Ved at bruge en rekursiv definition kan vi være meget præcise:

$$\sum_{k=1}^{n} z_k = \begin{cases} z_1 & \text{hvis } n = 1, \\ \left(\sum_{k=1}^{n-1} z_k\right) + z_n & \text{hvis } n > 1. \end{cases} \tag{5.5}$$

Ved at bruge denne rekursive definition opnår vi præcis det, vi ønsker. Man kan nemt benytte definitionen og verificere, at man ganske rigtigt for små værdier af n opnår:

n	$\sum_{k=1}^{n} z_k$
1	z_1
2	$z_1 + z_2$
3	$z_1 + z_2 + z_3$
4	$z_1 + z_2 + z_3 + z_4$

(5.6)

Hvis $f : \mathbb{N} \to \mathbb{C}$ er en funktion, kan man på lignende vis erstatte summen $f(1) + f(2) + \cdots + f(n)$ med det mere kompakte udtryk $\sum_{k=1}^{n} f(k)$. Overvej for eksempel udtrykket $\sum_{k=1}^{n} k$, det vil sige summen af de første n naturlige tal. Som i Tabel 5.6 får vi følgende:

n	$\sum_{k=1}^{n} k$
1	1
2	$1 + 2 = 3$
3	$1 + 2 + 3 = 6$
4	$1 + 2 + 3 + 4 = 10$

(5.7)

Denne notation er praktisk at have ved hånden i flere senere kapitler, og den benyttes flittigt indenfor flere områder i matematikken og naturvidenskabben.

Variablen k i et udtryk som $\sum_{k=1}^{n} z_k$ kaldes sumindekset. Der er som sådan ingen grund til at vælge symbolet k, og det er helt fint at benytte et andet symbol. For eksempel har man $\sum_{k=1}^{n} z_k = \sum_{j=1}^{n} z_j$, da begge summer svarer til at lægge tallene $z_1, \ldots, z_n$ sammen. Man kan også indeksere de tal, der skal lægges sammen, på en anden måde. Hvis vi vil lægge tallene $z_2, \ldots, z_{10}$ sammen, kan man simpelthen skrive $\sum_{k=2}^{10} z_k$.

5.4 Induktion

I det foregående afsnit sluttede vi af med at løse gåden om Hanois tårne fuldstændigt med en rekursiv tilgang. Antallet af træk, vores løsning kræver, kan beskrives rekursivt. Hvis vi lader $T(n)$ betegne antallet af træk, som vores strategi vil kræve, når gåden stilles med n skiver, så ved vi, at $T(1) = 1$ (gåden med kun én skive kan løses i ét træk), men også at $T(n) = T(n-1) + 1 + T(n-1) = 2T(n-1) + 1$ for $n \geq 2$ (vores strategi lod os flytte en stak af $n-1$ skiver to gange, mens den n'te skive skulle flyttes én gang). Med andre ord har vi

$$T(n) = \begin{cases} 1 & \text{hvis } n = 1, \\ 2 \cdot T(n-1) + 1 & \text{hvis } n \geq 2. \end{cases} \tag{5.8}$$

For eksempel er $T(2) = 2 \cdot 1 + 1 = 3$, $T(3) = 2 \cdot 3 + 1 = 7$ og $T(4) = 2 \cdot 7 + 1 = 15$. Det slår os nu, at værdien af $T(n)$ for disse små værdier af n ser ud til altid at være én mindre end 2^n. Vi vil derfor "gætte" på, at $T(n) = 2^n - 1$ for alle naturlige tal n. Lad os teste denne formodning, et ord vi hellere benytter frem for "gætte", ved at beregne $T(5)$. Vi har $T(5) = 2 \cdot T(4) + 1 = 2 \cdot 15 + 1 = 31$. Dette bekræfter vores formodning om, at $T(n) = 2^n - 1$ for $n = 5$. Ulempen er, vi nu kun ved, at formodningen er sand for alle n i mængden $\{1, 2, 3, 4, 5\}$. Vi kunne selvfølgelig fortsætte med at bekræfte vores formodning for flere værdier af n ved at beregne $T(6), T(7)$ og så videre, men da der er uendeligt mange naturlige tal, er det umuligt for os at verificere formlen $T(n) = 2^n - 1$ for *alle* naturlige tal n på denne måde. Heldigvis har de naturlige tal en ganske intuitiv egenskab, der kan hjælpe os, og som vi fremsætter uden bevis:

Sætning 5.4.1 Induktionsprincippet
Lad S være en delmængde af de naturlige tal, og antag, at S har følgende to egenskaber:

1. $1 \in S$,

2. hvis $n - 1 \in S$ for et vilkårligt naturligt tal $n \geq 2$, da er også $n \in S$.

I så fald har vi $S = \mathbb{N}$.

Udsagnet i denne sætning kaldes ofte *induktionsprincippet* eller simpelthen *induktion*. Krav 1. ($1 \in S$) kaldes *basistrinnet for induktionen*, mens krav 2. (hvis $n - 1 \in S$ for et naturligt tal n, så er også $n \in S$) kaldes *induktionstrinnet*. Grunden til, at det naturlige tal n i krav 2. skal være mindst to, er, at ellers er $n-1$ måske ikke et naturligt tal. Hvis $n = 1$, så er $n - 1 = 0$, men 0 tilhører ikke $\mathbb{N}$. Krav 2. kan omformuleres ved brug af udsagnslogikken som følger.

2. for alle $n \in \mathbb{N}_{\geq 2}: \quad n-1 \in S \quad \Rightarrow \quad n \in S.$

Verificering af krav 2., det vil sige verificering af induktionstrinnet, gøres typisk ved at vise, at $n \in S$ er sandt, hvis vi antager, at $n-1 \in S$. Når man verificerer induktionstrinnet $n-1 \in S \quad \Rightarrow \quad n \in S$, kaldes antagelsen $n-1 \in S$ for *induktionshypotesen*. Processen, der verificerer de to krav, kaldes typisk *bevis ved induktion* eller, hvis variablen n's rolle skal understreges, et *bevis ved induktion på n*.

Induktionsprincippet er nøglen til at forstå mange udsagn i matematikken og er også centralt i datalogi, da det dér kan bruges til at vise korrektheden af forskellige algoritmer, rekursive definitioner og computerprogrammer.

I matematikken er det bekvemt at benytte en omformulering af induktionsprincippet, så arbejde med en delmængde $S \subseteq \mathbb{N}$ undgås. Det kan nemlig undgås ved udnyttelse af en pæn konsekvens af Sætning 5.4.1. Sådanne konsekvenser kaldes typisk for "korollarer" i matematiske tekster, en terminologi vi vil benytte her.

Korollar 5.4.2

Lad $P(n)$ være et logisk udsagn for ethvert naturligt tal n. Antag, at følgende to udsagn er sande:

1. $P(1)$,
2. for alle $n \in \mathbb{N}_{\geq 2}: \quad P(n-1) \Rightarrow P(n)$.

Da er $P(n)$ sandt for alle $n \in \mathbb{N}$.

Bevis. For at kunne benytte Sætning 5.4.1 gør vi brug af et trick ved at definere $S = \{n \in \mathbb{N} \mid P(n)\}$. Med andre ord, $n \in S$, netop hvis $P(n)$ er sandt. For at kunne konkludere, at $P(n)$ er sandt for alle naturlige tal n, er det nok at vise, at $S = \mathbb{N}$. Hvis der eksisterede et naturligt tal m, således at $P(m)$ er falsk, så ville vi ved definitionen af S få, at $m \notin S$ og derfor, at $S \neq \mathbb{N}$.

Nu benytter vi Sætning 5.4.1 til at vise, at $S = \mathbb{N}$. Antagelsen om, at $P(1)$ er sandt, betyder blot, at $1 \in S$. Antagelsen om, at for alle $n \in \mathbb{N}_{\geq 2}: \quad P(n-1) \Rightarrow P(n)$, betyder, at når $n-1 \in S$, er også $n \in S$. Derfor er de to krav fra Sætning 5.4.1 opfyldt. Dermed kan vi ved brug af Sætning 5.4.1 konkludere, at $S = \mathbb{N}$. Dette betyder blot, som allerede bemærket, at $P(n)$ er sandt for alle naturlige tal n. $\square$

Som i Sætning 5.4.1 kaldes det at tjekke, at $P(1)$ er opfyldt, for basistrinnet for induktionen, mens det at tjekke, at for alle $n \in \mathbb{N}_{\geq 2}: \quad P(n-1) \Rightarrow P(n)$, kaldes for induktionstrinnet. Mens man udfører induktionstrinnet, kaldes det logiske udsagn $P(n-1)$ for induktionshypotesen ligesom før. Også udsagnet i Korollar 5.4.2 kaldes som helhed stadig for induktionsprincippet. Vi kan derfor bevise en påstand af formen "$P(n)$ er sandt for alle naturlige tal n" ved at følge følgende strategi:

(i) Informér læseren om, at du nu vil bevise påstanden, at "$P(n)$ er sandt for alle naturlige tal n," ved brug af induktion på n.

(ii) **Basistrin:** Tjek, at $P(1)$ er opfyldt.

(iii) **Induktionstrin:** Antag for et vilkårligt naturligt tal $n \geq 2$, at $P(n-1)$ er sandt, og benyt denne antagelse (induktionshypotesen) til at vise, at så er også $P(n)$ sandt. Udfordringen her er nogle gange at finde ud af, hvordan man bruger induktionshypotesen $P(n-1)$ til sin fordel.

(iv) Informér læseren, når de foregående punkter er udført, om, at man ved induktionsprincippet nu kan konkludere, at $P(n)$ gælder for alle naturlige tal n.

Lad os nu benytte denne strategi til at bevise vores formodning om, at $T(n) = 2^n - 1$. Med andre ord, lad os bevise følgende:

Påstand: Lad $T : \mathbb{N} \to \mathbb{N}$ opfylde rekursionen

$$T(n) = \begin{cases} 1 & \text{hvis } n = 1, \\ 2 \cdot T(n-1) + 1 & \text{hvis } n \geq 2. \end{cases}$$

Så har vi for alle $n \in \mathbb{N}$, at $T(n) = 2^n - 1$.

Bevis. Lad $P(n)$ være udsagnet $T(n) = 2^n - 1$. Vi vil bevise påstanden ved induktion på n.

Basistrin: Vi har $T(1) = 1$. Da $2^1 - 1 = 1$, ser vi, at $T(1) = 2^1 - 1$. Derfor er $P(1)$ opfyldt.

Induktionstrin: Lad $n \geq 2$ være et vilkårligt naturligt tal. Induktionshypotesen er $P(n-1)$, som i vores tilfælde blot betyder ligningen $T(n-1) = 2^{n-1} - 1$. Hvis vi antager dette, skal vi nu udlede, at $P(n)$ er gyldigt. Med andre ord, hvis vi antager, at $T(n-1) = 2^{n-1} - 1$, skal vi udlede, at $T(n) = 2^n - 1$. Fra den rekursive definition af $T(n)$ for $n \geq 2$ ved vi, at $T(n) = 2 \cdot T(n-1) + 1$. Ved at kombinere dette med induktionshypotesen, ser vi, at

$$T(n) = 2 \cdot T(n-1) + 1 = 2 \cdot (2^{n-1} - 1) + 1 = 2 \cdot 2^{n-1} - 2 \cdot 1 + 1 = 2^n - 1.$$

Dette er præcis, hvad vi skulle vise.

Nu hvor vi har vist basistrinnet for induktionen samt induktionstrinnet, kan vi konkludere ved induktionsprincippet, at udsagnet $T(n) = 2^n - 1$ er sandt for alle naturlige tal n. $\square$

Man kan faktisk vise, at den strategi, vi fandt i Afsnit 5.2, er den bedst mulige. Med andre ord vil enhver løsning af gåden med n skiver tage mindst $T(n)$ træk. Vi ser ved udregning, at løsningen af en ti-skive-version af Hanois tårne ville kræve $2^{10} - 1 = 1023$ træk.

Den bedste måde at lære induktionsbeviser på er at gennemgå adskillige eksempler og derefter prøve at udføre et induktionsbevis selv. Lad os derfor finde nogle flere eksempler frem. Her er et berømt et:

Eksempel 5.4.1

Lad os betegne ved $S(n)$ summen af de første n naturlige tal. Uformelt skriver man ofte $1 + 2 + \cdots + n$ for denne sum, men vi kan benytte os af summationstegnet og skrive $S(n) = \sum_{k=1}^{n} k$. Som vi så i Tabel 5.7, har vi for eksempel $S(1) = 1$, $S(2) = 1 + 2 = 3$, $S(3) = 1 + 2 + 3 = 6$ og $S(4) = 1 + 2 + 3 + 4 = 10$. Påstanden er nu, at følgende formel gælder for ethvert naturligt tal n:

$$S(n) = \frac{n \cdot (n+1)}{2}.$$

Bemærk, at $S(n)$ opfylder følgende rekursion:

$$S(n) = \begin{cases} 1 & \text{hvis } n = 1, \\ S(n-1) + n & \text{hvis } n \geq 2. \end{cases}$$

Faktisk har vi allerede observeret, at $S(1) = 1$. I tilfældet $n \geq 2$ får vi ved brug af Ligning (5.5), at

$$S(n) = 1 + \cdots + n = \sum_{k=1}^{n} k = \left(\sum_{k=1}^{n-1} k \right) + n = S(n-1) + n.$$

Lad os nu bevise følgende påstand.

Påstand: Lad $S(n) = 1 + \cdots + n$ for $n \in \mathbb{N}$, altså summen af de første n naturlige tal. Så er $S(n) = \frac{n \cdot (n+1)}{2}$.

Bevis. Vi beviser påstanden ved induktion på n.

Basistrin: Hvis $n = 1$, så er $S(1) = 1$, mens $\frac{1 \cdot (1+1)}{2} = 1$. Derfor er formlen $S(n) = \frac{n \cdot (n+1)}{2}$ gyldig for $n = 1$.

Induktionstrin: Lad $n \geq 2$ være et vilkårligt naturligt tal, og antag som induktionshypotese, at $S(n-1) = \frac{(n-1) \cdot (n-1+1)}{2}$. Vi kan reducere induktionshypotesen en smule, da der gælder, at

$S(n-1) = \frac{(n-1)\cdot n}{2}$. Hvis vi antager induktionshypotesen og benytter, at $S(n) = S(n-1) + n$, kan vi konkludere, at

$$\begin{aligned} S(n) &= S(n-1) + n \\ &= \frac{(n-1)\cdot n}{2} + n \\ &= \frac{(n-1)\cdot n}{2} + \frac{2\cdot n}{2} \\ &= \frac{n^2 - n}{2} + \frac{2\cdot n}{2} \\ &= \frac{n^2 - n + 2\cdot n}{2} \\ &= \frac{n^2 + n}{2} \\ &= \frac{n\cdot(n+1)}{2}. \end{aligned}$$

Dette er præcis, hvad vi skulle vise, hvilket fuldfører induktionstrinnet.

Ved induktionsprincippet kan vi dermed konkludere, at formlen $S(n) = \frac{n\cdot(n+1)}{2}$ er gyldig for alle naturlige tal n. □

Eksempel 5.4.2

Dette eksempel er af en mere teoretisk karakter og kan springes over ved første læsning. Vi vil sikre os, at den rekursive definition, vi tidligere gav af fakultetsfunktionen $\text{fac} : \mathbb{N} \to \mathbb{N}$ i Ligning (5.1), rent faktisk var korrekt i matematisk forstand. Problemet er, at vi aldrig viste, at fac er defineret ved sin rekursive beskrivelse for *ethvert* naturligt tal n. Med andre ord, når vi skriver $\text{fac} : \mathbb{N} \to \mathbb{N}$, siger vi implicit, at funktionens definitionsmængde er $\mathbb{N}$, men hvordan ved vi det? Det, vi skal gøre, er at vise, at for ethvert naturligt tal n vil den rekursive beskrivelse i Ligning (5.1) give outputværdien $\text{fac}(n)$ efter endeligt mange trin.

Lad derfor $P(n)$ være udsagnet, at $\text{fac}(n)$ kan beregnes i endeligt mange trin ved hjælp af Ligning (5.1) for ethvert naturligt tal n. Vi vil vise, at dette udsagn $P(n)$ er sandt for alle naturlige tal. Basistrinnet for induktionen er hurtigt vist opfyldt, da Ligning (5.1) straks medfører, at $\text{fac}(1) = 1$. Lad nu $n \geq 2$ være et vilkårligt naturligt tal, og antag som induktionshypotese, at $\text{fac}(n-1)$ kan beregnes i endeligt mange trin ved hjælp af Ligning (5.1). Da $n \geq 2$, medfører Ligning (5.1), at $\text{fac}(n) = \text{fac}(n-1)\cdot n$. Givet $\text{fac}(n-1)$ har vi derfor kun brug for en multiplikation med n for at udregne $\text{fac}(n)$. Altså kan $\text{fac}(n)$ udregnes i endeligt mange trin, hvis $\text{fac}(n-1)$ kan. Dette fuldfører induktionstrinnet.

Mere generelt kan en funktion $f : \mathbb{N} \to B$ fra de naturlige tal $\mathbb{N}$ til en givet mængde B defineres rekursivt, så længe $f(1)$ er bestemt, og så længe værdien $f(n)$ kan udregnes ud fra $f(n-1)$ for ethvert $n \geq 2$. Det skyldes, at der i sådanne tilfælde gælder et lignende ræsonnement som det, vi netop har udført for fakultetsfunktionen. For eksempel definerer Ligning (5.2) z^n for ethvert naturligt tal n.

5.5 En variant af induktion

Der findes mange varianter af induktion. I dette afsnit vil vi nævne en af dem: induktion, der starter med et andet basistrin. Indtil videre har basistrinnet for vores induktionsbeviser været for $n = 1$, hvorefter vi har haft at gøre med større naturlige tal n. I nogle tilfælde giver et logisk udsagn dog også mening for andre værdier af n. Betragt for eksempel udsagnet:

Et polynomium $p(Z) \in \mathbb{C}[Z]$ af grad n har højst n rødder i $\mathbb{C}$.

Dette udsagn giver også mening for $n = 0$. Faktisk er udsagnet for $n = 0$ ret let at verificere: et polynomium $p(Z)$ af grad nul er simpelthen en konstant p_0, som ikke er nul. Konstanten p_0 er forskellig fra nul, netop fordi den ledende koefficient i et d'te-gradspolynomium generelt er forskellig fra nul ifølge Definition 4.1.1. Men så er $p(z) = p_0 \neq 0$ for alle $z \in \mathbb{C}$, hvilket medfører, at polynomiet ikke har nogen rødder.

Omvendt er der udsagn, der kun bliver sande for store nok værdier af n. Betragt for eksempel udsagnet:

Der eksisterer n punkter i planen $\mathbb{R}^2$, som ikke ligger på linje.

Hvis $n = 1$, er dette forkert, da der findes mange linjer igennem ethvert givet punkt. Også hvis $n = 2$, er dette forkert, da linjen, der forbinder de to punkter, vil indeholde disse punkter. Men for $n \geq 3$ er udsagnet sandt. Hvis $n \geq 3$, kan vi for eksempel vælge tre af punkterne som hjørner i en ligesidet trekant og de resterende $n - 3$ punkter vilkårligt.

På grund af denne slags eksempler er det praktisk at have en lidt mere fleksibel variant af induktion. For et givet heltal $a \in \mathbb{Z}$ har vi betegnelsen $\mathbb{Z}_{\geq a} = \{n \in \mathbb{Z} \mid n \geq a\}$. For eksempel er $\mathbb{Z}_{\geq -1} = \{-1, 0, 1, 2, \dots\}$. Med denne notation på plads kan vi formulere følgende variant af induktion, kaldet *induktion med basistrin b*:

Sætning 5.5.1

Lad $b \in \mathbb{Z}$ være et heltal, og lad $P(n)$ være et logisk udsagn for ethvert heltal $n \geq b$. Antag, at følgende to udsagn er sande:

1. $P(b)$,
2. for alle $n \in \mathbb{Z}_{\geq b+1}$: $\quad P(n-1) \Rightarrow P(n)$.

Så er $P(n)$ sandt for alle $n \in \mathbb{Z}_{\geq b}$.

Bevis. Lad os definere det logiske udsagn $Q(n)$ til at være $P(n + b - 1)$. Så er $Q(n)$ defineret for ethvert naturligt tal n. Vi ser nemlig, at hvis $n \geq 1$, så er $n + b - 1 \geq b$. Nu anvender vi Korollar 5.4.2 på det logiske udsagn $Q(n)$. Det første krav fra Korollar 5.4.2 er, at $Q(1)$ skal være gyldigt. Dette er opfyldt, da $Q(1) = P(b)$, og det er givet, at $P(b)$ er gyldigt. Det andet krav fra Korollar 5.4.2 bliver, at for alle $n \in \mathbb{N}_{\geq 2}$: $\quad Q(n-1) \Rightarrow Q(n)$. Men da $n \geq 2$, har vi $n + b - 1 \geq b + 1$ og derfor $n + b - 1 \in \mathbb{Z}_{\geq b+1}$. Da $Q(n-1) = P(n+b-2)$

og $Q(n) = P(n+b-1)$, og da implikationen $P(n+b-2) \Rightarrow P(n+b-1)$ er gyldig (vi ved, at $n+b-1 \in \mathbb{Z}_{\geq b+1}$), ser vi, at implikationen $Q(n-1) \Rightarrow Q(n)$ er gyldig. Derfor er det andet krav til det logiske udsagn $Q(n)$, ved anvendelse af Korollar 5.4.2, også opfyldt. Derfor medfører korollaret, at $Q(n)$ er gyldigt for alle naturlige tal n. Da $Q(n) = P(n+b-1)$, betyder dette, at $P(n+b-1)$ er gyldigt for alle naturlige tal n. For eksempel er $P(1+b-1) = P(b)$ gyldigt, $P(2+b-1) = P(b+1)$ er gyldigt og så videre. Dette svarer til udsagnet, at $P(n)$ er gyldigt for alle heltal $n \geq b$, hvilket er, hvad vi ønskede at vise. $\square$

Bemærk, at hvis vi vælger $b = 1$, genskaber vi Korollar 5.4.2. Den overordnede struktur af et bevis ved induktion med basistrin b er den samme som for den sædvanlige induktion. Man har stadig et basistrin og et induktionstrin. Lad os gennemgå et eksempel på et bevis ved induktion af denne type.

Eksempel 5.5.1

Betragt uligheden $n + 10 \leq n^2 - n$. Da et polynomium af anden grad, så som $n^2 - n$, vokser hurtigere end et polynomium af første grad, så som $n + 10$, bør man kunne forvente, at hvis n bliver stor nok, er den givne ulighed sand. Lad os nu ved $P(n)$ betegne udsagnet, at $n + 10 \leq n^2 - n$. I dette tilfælde kan vi definere $P(n)$ for ethvert heltal n. Udsagnet $P(4)$ er for eksempel uligheden $4 + 10 \leq 4^2 - 4$. Dette er dog falsk, da $14 = 4 + 10 > 4^2 - 4 = 12$. På den anden side er udsagnet $P(5)$ sandt, da $15 = 5 + 10 \leq 5^2 - 5 = 20$. Vi hævder derfor, at $P(n)$ er sandt for ethvert $n \in \mathbb{Z}_{\geq 5}$ og giver et bevis ved induktion ved hjælp af Sætning 5.5.1 med $b = 5$:

Basistrin: Vi har allerede verificeret, at $P(5)$ er gyldigt, så basistilfældet er klaret.

Induktionstrin: Lad $n \geq 6$ være et vilkårligt naturligt tal, og antag som induktionshypotese, at $P(n-1)$ er gyldigt. Dette betyder, at vi må antage, at $(n-1) + 10 \leq (n-1)^2 - (n-1)$. Ved hjælp af denne antagelse skal vi udlede, at $P(n)$ er gyldigt. Lad os først omskrive induktionshypotesen til en mere bekvem form. Vi har $(n-1) + 10 = n + 9$, mens $(n-1)^2 - (n-1) = n^2 - 2n + 1 - n + 1 = n^2 - 3n + 2$. Derfor svarer induktionshypotesen til antagelsen, at uligheden $n + 9 \leq n^2 - 3n + 2$ er gyldig. Men så kan vi udlede:

$$\begin{aligned} n + 10 &= (n+9) + 1 \\ &\leq (n^2 - 3n + 2) + 1 \\ &= n^2 - 3n + 3 \\ &= n^2 - n - 2n + 3 \\ &\leq n^2 - n. \end{aligned}$$

Den endelige ulighed holder, da $-2n + 3 \leq 0$ for ethvert $n \geq 6$ (faktisk endda for ethvert $n \geq 2$). Vi konkluderer, at hvis $P(n-1)$ er sandt, så er $n + 10 \leq n^2 - n$, det vil sige $P(n)$, også sandt. Dette er præcis, hvad vi skulle vise, hvilket fuldfører induktionstrinnet.

Ved at bruge induktion med basistrin 5 kan vi konkludere, at uligheden $n + 10 \leq n^2 - n$ er gyldig for alle $n \in \mathbb{Z}_{\geq 5}$.

Kapitel 6

Systemer af lineære ligninger

6.1 Strukturen af lineære ligningssystemer

Når man arbejder med en ligning i én variabel benyttes ofte variablen x. I Eksempel 1.4.2 arbejdede vi for eksempel med ligningen $2|x| = 2x + 1$. Ofte er der ikke kun én variabel, men flere. Hvis der er to variable, bruger man ofte x og y, hvis der er tre x, y og z, men hvad gør man, hvis der er endnu flere variable, for eksempel fem variable? Da vil man typisk benytte sig af variabelnavne så som x_1, x_2 og så videre. Hvis vi for eksempel har behov for fem variable, benytter vi blot x_1, x_2, x_3, x_4 og x_5. Vi kan dermed lade det præcise antal variable være uspecificeret og sige, at vi har n variable for et naturligt tal $n \in \mathbb{N}$. Man siger, at man har en ligning i de n variable $x_1, \dots, x_n$.

En *lineær ligning* i de n variable $x_1, \dots, x_n$ er en ligning på formen

$$a_1 \cdot x_1 + \cdots + a_n \cdot x_n = b,$$

hvor $a_1, \dots, a_n, b$ er konstanter. Disse konstanter vil typisk være reelle eller komplekse tal afhængigt af situationen. For at undgå hele tiden at skulle udspecificere, om vi arbejder med reelle eller komplekse tal, vil vi her introducere følgende definition:

Definition 6.1.1

En mængde $\mathbb{F}$ kaldes et *legeme*, hvis addition $+$ og multiplikation $\cdot$ er defineret for alle par af elementer i $\mathbb{F}$ på en sådan måde, at følgende regler er opfyldt:

(i) Addition og multiplikation er *associative*: $a_1 + (a_2 + a_3) = (a_1 + a_2) + a_3$ og $a_1 \cdot (a_2 \cdot a_3) = (a_1 \cdot a_2) \cdot a_3$ for alle $a_1, a_2, a_3 \in \mathbb{F}$.

(ii) Addition og multiplikation er *kommutative*: $a_1 + a_2 = a_2 + a_1$ og $a_1 \cdot a_2 = a_2 \cdot a_1$ for alle $a_1, a_2, a_3 \in \mathbb{F}$.

(iii) Der gælder *distributivitet* af multiplikation over addition: $a_1 \cdot (a_2 + a_3) = a_1 \cdot a_2 + a_1 \cdot a_3$ for alle $a_1, a_2, a_3 \in \mathbb{F}$.

(iv) Addition og multiplikation har et neutralt element. Det vil sige, der findes to forskellige elementer i $\mathbb{F}$, typisk betegnet ved 0 og 1, som opfylder henholdsvis $a + 0 = a$ og $a \cdot 1 = a$ for alle $a \in \mathbb{F}$.

(v) Additivt inverse eksisterer: for ethvert $a \in \mathbb{F}$ findes der et element i $\mathbb{F}$, betegnet $-a$ og kaldet den additivt inverse af a, således at $a + (-a) = 0$.

(vi) Multiplikativt inverse eksisterer: for ethvert $a \in \mathbb{F} \setminus \{0\}$ findes der et element i $\mathbb{F}$, betegnet a^{-1} eller $1/a$ og kaldet den multiplikativt inverse af a, således at $a \cdot a^{-1} = 1$.

Sætningerne 3.2.2 og 3.2.3 siger tilsammen, at de komplekse tal udgør et legeme. Også de reelle tal $\mathbb{R}$ med den sædvanlige addition og multiplikation udgør et legeme. Der findes mange flere mulige eksempler på legemer, men når vi benytter symbolet $\mathbb{F}$ eller skriver noget i stil med "legemet $\mathbb{F}$", kan du blot tænke på $\mathbb{R}$ eller $\mathbb{C}$. For at vise, at der findes flere legemer end blot disse to, giver vi her to eksempler.

Eksempel 6.1.1

Lad $\mathbb{F} = \mathbb{Q}$ være mængden af rationale tal, se Eksempel 2.1.4. Denne mængde, udstyret med den sædvanlige addition og multiplikation, er et legeme. Det kaldes legemet af rationale tal.

Eksempel 6.1.2

Lad $\mathbb{F}_2 = \{0, 1\}$, og definér addition og multiplikation som følger: $0 + 0 = 0, 0 + 1 = 1, 1 + 0 = 1, 1 + 1 = 0$ og $0 \cdot 0 = 0, 0 \cdot 1 = 0, 1 \cdot 0 = 0, 1 \cdot 1 = 1$. Med disse definitioner af addition og multiplikation er $\mathbb{F}_2$ et legeme. Det kaldes legemet af bits, det binære legeme eller alternativt det endelige legeme med to elementer.

Vender vi tilbage til vores studie af lineære ligninger, kan vi nu give en mere præcis definition.

Definition 6.1.2

En lineær ligning over et legeme $\mathbb{F}$ i de n variable $x_1, \ldots, x_n$ er en ligning på formen

$$a_1 \cdot x_1 + \cdots + a_n \cdot x_n = b,$$

hvor $a_1, \ldots, a_n, b \in \mathbb{F}$.

En løsning til denne lineære ligning er et n-tupel $(v_1, \ldots, v_n) \in \mathbb{F}^n$, således at $a_1 \cdot v_1 + \cdots + a_n \cdot v_n = b$.

Vi har set notationen $\mathbb{F}^n$ i denne definition før, nemlig i Afsnit 2.1 (se Ligning (2.3)). Det er det kartesiske produkt af $\mathbb{F}$ med sig selv n gange. Mere jordnært er $\mathbb{F}^n$ simpelthen mængden af alle n-tupler $(v_1, \ldots, v_n)$, hvor hver koordinat er et element fra $\mathbb{F}$. Nogle gange udelades multiplikationssymbolet mellem konstanten og variablerne. For eksempel har $2x_1$ den samme betydning som $2 \cdot x_1$.

Der er en spidsfindighed i Definition 6.1.2, som er let at overse. Hvis vi siger, at vi betragter en lineær ligning over $\mathbb{F}$, er vi kun interesserede i løsninger $(v_1, \ldots, v_n)$, der ligger i $\mathbb{F}^n$. Med andre ord, ved at specificere, at den lineære ligning er over $\mathbb{F}$, siger vi implicit, at alle koordinaterne af en løsning $(v_1, \ldots, v_n)$ skal ligge i $\mathbb{F}$. Lad os betragte nogle få eksempler.

Eksempel 6.1.3 (a) Find en løsning til den lineære ligning $3x_1 + x_2 = 5$ over $\mathbb{R}$.

(b) Betragt den lineære ligning $x_1 + x_2 = 0$ over $\mathbb{C}$. Er $(i, -i) \in \mathbb{C}^2$ en løsning til denne lineære ligning?

(c) Betragt den lineære ligning $x_1 + x_2 = 0$ over $\mathbb{R}$. Er $(i, -i) \in \mathbb{C}^2$ en løsning til denne lineære ligning?

(d) Betragt den lineære ligning $x_1 + x_2 = 0$ over $\mathbb{R}$. Find en løsning.

Svar:

(a) Der er mange mulige løsninger, men for eksempel er $(x_1, x_2) = (0, 5)$ en løsning, da $3 \cdot 0 + 5 = 5$.

(b) Da $i + (-i) = 0$, er talparret $(i, -i) \in \mathbb{C}^2$ en løsning til den lineære ligning $x_1 + x_2 = 0$ over $\mathbb{C}$.

(c) Selvom $i + (-i) = 0$, er talparret $(i, -i)$ ikke en løsning til den lineære ligning $x_1 + x_2 = 0$ over $\mathbb{R}$. Det skyldes, at talparret $(-i, i)$ ikke er et element i $\mathbb{R}^2$.

(d) En mulig løsning er $(1, -1)$. En anden løsning er $(0, 0)$.

Nu kommer vi til hovedemnet i dette afsnit, nemlig systemer af lineære ligninger. Det er simpelthen en udvidelse af Definition 6.1.2, hvor vi betragter ikke blot én lineær ligning men adskillige lineære ligninger over et legeme $\mathbb{F}$ på samme tid.

Definition 6.1.3

Et system af m lineære ligninger $R_1, \ldots, R_m$ over et legeme $\mathbb{F}$ i de n variable $x_1, \ldots, x_n$ er et system af m ligninger på formen

$$\left\{\begin{array}{lccccccc} R_1: & a_{11} \cdot x_1 & + & \cdots & + & a_{1n} \cdot x_n & = & b_1 \\ R_2: & a_{21} \cdot x_1 & + & \cdots & + & a_{2n} \cdot x_n & = & b_2 \\ & & & \vdots & & & & \vdots \\ R_m: & a_{m1} \cdot x_1 & + & \cdots & + & a_{mn} \cdot x_n & = & b_m \end{array}\right.$$

hvor $a_{11}, \ldots, a_{mn}, b_1, \ldots, b_m \in \mathbb{F}$.

En løsning til dette system af lineære ligninger er et n-tupel $(v_1, \ldots, v_n) \in \mathbb{F}^n$, således at der for alle j mellem 1 og m gælder, at $a_{j1} \cdot v_1 + \cdots + a_{jn} \cdot v_n = b_j$.

Lidt forklaring af notationen er på sin plads her. Først og fremmest blev der tilføjet et dobbeltindeks på konstanterne foran variablerne. a_{ij} betegner konstanten, der optræder i ligning i foran variablen x_j. Har vi for eksempel mindst to ligninger og mindst tre variable, vil a_{23} betegne konstanten i ligning to foran variablen x_3. Hvis $m = 1$ i Definition 6.1.3, får vi blot tilfældet med én lineær ligning som beskrevet i Definition 6.1.2.

Brugen af krølleparentesen $\{$ foran ligningerne er blot for at understrege, at alle ligninger betragtes samtidigt, og at en løsning til systemet skal opfylde alle ligninger på samme tid. I logiske termer kan vi derfor skrive, at et n-tupel $(v_1, \ldots, v_n) \in \mathbb{F}^n$ er en løsning til systemet af ligninger som angivet i Definition 6.1.3, netop hvis:

$$a_{11} \cdot v_1 + \cdots + a_{1n} \cdot v_n = b_1 \quad \wedge \quad \cdots \quad \wedge \quad a_{m1} \cdot v_1 + \cdots + a_{mn} \cdot v_n = b_m.$$

Brugen af $R_1, \ldots, R_m$ som "etiketter" for ligningerne er ikke nødvendig, og ofte udelades disse etiketter. Vi vil også oftest udelade disse etiketter fremover, men mens vi udvikler teorien om, hvordan man løser systemer af lineære ligninger, kan de være praktiske. For at fordøje denne definition bør vi straks gennemgå nogle eksempler.

Eksempel 6.1.4

Bestem løsningsmængden til følgende system af to lineære ligninger i to variable over $\mathbb{R}$:

$$\begin{cases} x_1 + 2x_2 & = & 1 \\ \quad x_2 & = & 2 \end{cases}$$

Dette system er temmelig simpelt at løse, da ligning to allerede fastsætter værdien af x_2 (nemlig $x_2 = 2$). Ved at benytte dette i den første ligning ser vi, at ethvert talpar (x_1, x_2), der opfylder *begge* lineære ligninger, vil opfylde $x_2 = 2$ og $x_1 = 1 - 2x_2 = 1 - 2 \cdot 2 = -3$. Derfor har systemet kun én løsning, nemlig $(x_1, x_2) = (-3, 2)$. Løsningsmængden er derfor givet ved $\{(-3, 2)\}$.

Eksempel 6.1.5

Betragt følgende system af lineære ligninger over $\mathbb{R}$ i variablerne $x_1, \ldots, x_4$:

$$\begin{cases} 2x_1 + 5x_2 + x_4 & = & 0 \\ \quad 3x_1 - x_3 & = & 6 \end{cases}$$

Lad os overveje, hvordan dette eksempel passer med Definition 6.1.3. Først og fremmest har vi to lineære ligninger og derfor $m = 2$. Desuden er de eneste variable, der optræder i disse to ligninger, x_1, x_2, x_3 og x_4. Derfor kan vi vælge $n = 4$. At bestemme a_{ij} er nu et spørgsmål om at aflæse konstanterne foran variablerne. Men før vi gør dette, er det praktisk at omskrive ligningssystemet en smule som følger:

$$\begin{cases} 2 \cdot x_1 & + & 5 \cdot x_2 & + & 0 \cdot x_3 & + & 1 \cdot x_4 & = & 0 \\ 3 \cdot x_1 & + & 0 \cdot x_2 & + & (-1) \cdot x_3 & + & 0 \cdot x_4 & = & 6 \end{cases}$$

Vi kan nu direkte aflæse, at $a_{11} = 2$, $a_{12} = 5$, $a_{13} = 0$, $a_{14} = 1$, $b_1 = 0$, $a_{21} = 3$, $a_{22} = 0$, $a_{23} = -1$, $a_{24} = 0$, og $b_2 = 6$. Vi vil bestemme løsningerne til dette system af lineære ligninger senere.

Et system af m lineære ligninger over et legeme $\mathbb{F}$ i de n variable $x_1, \ldots, x_n$ kaldes *homogent*, hvis det for alle i mellem 1 og m gælder, at $b_i = 0$. Ellers kaldes systemet *inhomogent*. Systemet givet i Eksempel 6.1.5 er inhomogent, da vi i det eksempel har $b_2 \neq 0$. Et eksempel på et homogent system af lineære ligninger i tre variable er:

$$\begin{cases} 3 \cdot x_1 & + & 5 \cdot x_2 & + & 10 \cdot x_3 & = & 0 \\ 5 \cdot x_1 & + & 2 \cdot x_2 & - & 2 \cdot x_3 & = & 0 \end{cases}$$

Bemærk, at tuplet $(0, 0, 0)$ er en mulig løsning til dette system. Mere generelt kan man vise, at et homogent system af lineære ligninger i n variable har tuplet $(0, \ldots, 0)$ som løsning. Lad os afslutte dette afsnit ved at formulere to struktursætninger, der vedrører løsningsmængderne til systemer af lineære ligninger. Den ene vil være for homogene systemer, den anden for inhomogene systemer.

Sætning 6.1.1

Lad et homogent system af m lineære ligninger $R_1, \ldots, R_m$ over et legeme $\mathbb{F}$ i de n variable $x_1, \ldots, x_n$ være givet,

$$\begin{cases} a_{11} \cdot x_1 & + & \cdots & + & a_{1n} \cdot x_n & = & 0 \\ a_{21} \cdot x_1 & + & \cdots & + & a_{2n} \cdot x_n & = & 0 \\ & & \vdots & & & & \vdots \\ a_{m1} \cdot x_1 & + & \cdots & + & a_{mn} \cdot x_n & = & 0 \end{cases}$$

hvor $a_{11}, \ldots, a_{mn} \in \mathbb{F}$. Da gælder der følgende:

(i) Tuplet $(0, \ldots, 0) \in \mathbb{F}^n$ er en løsning til systemet.

(ii) Hvis $(v_1, \ldots, v_n) \in \mathbb{F}^n$ er en løsning, og $c \in \mathbb{F}$, så er $(c \cdot v_1, \ldots, c \cdot v_n)$ også en løsning.

(iii) Hvis $(v_1, \ldots, v_n), (w_1, \ldots, w_n) \in \mathbb{F}^n$ er løsninger, så er $(v_1 + w_1, \ldots, v_n + w_n)$ også en løsning.

Bevis. Vi har allerede beskrevet, hvorfor tuplet $(0, \ldots, 0)$ er en løsning til et homogent system. Vi vil bevise det tredje udsagn og overlade beviset på det andet udsagn til læseren. Hvis $(v_1, \ldots, v_n), (w_1, \ldots, w_n) \in \mathbb{F}^n$ er løsninger, gælder der for alle j mellem 1 og m, at:

$$a_{j1} \cdot v_1 + \cdots + a_{jn} \cdot v_n = 0, \text{ og } a_{j1} \cdot w_1 + \cdots + a_{jn} \cdot w_n = 0.$$

Lægger vi disse ligninger sammen, får vi, at

$$a_{j1} \cdot v_1 + a_{j1} \cdot w_1 + \cdots + a_{jn} \cdot v_n + a_{jn} \cdot w_n = 0,$$

som kan omskrives til

$$a_{j1} \cdot (v_1 + w_1) + \cdots + a_{jn} \cdot (v_n + w_n) = 0.$$

Læseren opfordres til at tænke over, hvilke egenskaber ved et legeme fra Definition 6.1.1 vi har benyttet os af her. Da dette er sandt for ethvert j, kan vi konkludere, at $(v_1 + w_1, \ldots, v_n + w_n)$ også er en løsning til det givne homogene system af lineære ligninger. □

Sætning 6.1.2

Lad et inhomogent system af m lineære ligninger $R_1, \ldots, R_m$ over et legeme $\mathbb{F}$ i de n variable $x_1, \ldots, x_n$ være givet,

$$\begin{cases} a_{11} \cdot x_1 & + & \cdots & + & a_{1n} \cdot x_n & = & b_1 \\ a_{21} \cdot x_1 & + & \cdots & + & a_{2n} \cdot x_n & = & b_2 \\ & & \vdots & & & & \vdots \\ a_{m1} \cdot x_1 & + & \cdots & + & a_{mn} \cdot x_n & = & b_m \end{cases}$$

hvor $a_{11}, \ldots, a_{mn}, b_1, \ldots, b_m \in \mathbb{F}$, og ikke alle b_i er nul. Hvis systemet har en løsning $(v_1, \ldots, v_n) \in \mathbb{F}^n$, så vil enhver anden løsning kunne skrives på formen $(v_1 + w_1, \ldots, v_n + w_n)$, hvor $(w_1, \ldots, w_n) \in \mathbb{F}^n$ er en løsning til det tilsvarende homogene system:

$$\begin{cases} a_{11} \cdot x_1 & + & \cdots & + & a_{1n} \cdot x_n & = & 0 \\ a_{21} \cdot x_1 & + & \cdots & + & a_{2n} \cdot x_n & = & 0 \\ & & \vdots & & & & \vdots \\ a_{m1} \cdot x_1 & + & \cdots & + & a_{mn} \cdot x_n & = & 0 \end{cases}$$

Bevis. Antag, at systemet har en løsning $(v_1, \ldots, v_n) \in \mathbb{F}^n$. Lad $(v_1', \ldots, v_n') \in \mathbb{F}^n$ være en hvilken som helst anden løsning. Hvis vi definerer $w_i = v_i' - v_i$, får vi fra definitionen af w_i, at $(v_1', \ldots, v_n') = (v_1 + w_1, \ldots, v_n + w_n)$. Derfor skal vi vise, at tuplet $(w_1, \ldots, w_n)$ er en løsning til det homogene system angivet i sætningen. Men vi ved, at der for alle j gælder, at:

$$a_{j1} \cdot v_1' + \cdots + a_{jn} \cdot v_n' = b_j, \text{og } a_{j1} \cdot v_1 + \cdots + a_{jn} \cdot v_n = b_j.$$

Hvis vi tager forskellen mellem disse to ligninger, får vi:

$$a_{j1} \cdot v_1' - a_{j1} \cdot v_1 + \cdots + a_{jn} \cdot v_n' - a_{jn} \cdot v_n = b_j - b_j,$$

som kan omskrives til

$$a_{j1} \cdot (v_1' - v_1) + \cdots + a_{jn} \cdot (v_n' - v_n) = 0.$$

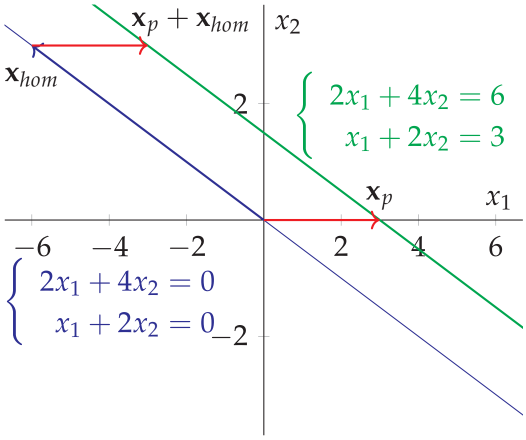
Da $w_i = v_i' - v_i$, får vi, at der for alle j gælder, at

$$a_{j1} \cdot w_1 + \cdots + a_{jn} \cdot w_n = 0.$$

Dette er præcis det samme som at sige, at $(w_1, \ldots, w_n)$ er en løsning til det homogene system angivet i sætningen. □

Det er ikke en tilfældighed, at Sætning 6.1.2 er formuleret, som den er. Sætningen gælder, hvis der findes en løsning til det inhomogene system, men der er ingen garanti for, at en sådan løsning faktisk findes. En løsning til et inhomogent system af lineære ligninger kaldes også for en *partikulær løsning*. I ord fastslår Sætning 6.1.2 således, at alle løsninger til et inhomogent system kan opnås som summen af en given partikulær løsning (hvis den findes) og løsningerne til det tilsvarende homogene system.

Vi illustrerer Sætning 6.1.2 i Figur 6.1 for et mindre inhomogent ligningssystem over legemet $\mathbb{R}$. I denne figur indikerer den grønne linje alle løsninger til systemet, mens den blå linje indikerer alle løsninger til det tilsvarende homogene system.



Figur 6.1: Løsningerne til et inhomogent system kan opnås ved at lægge en partikulær løsning $\mathbf{x}_p$ til løsningerne $\mathbf{x}_{hom}$ til det tilsvarende homogene system.

Lad os for fuldstændighedens skyld give et lille eksempel på et inhomogent system af lineære ligninger, der ikke har nogen løsninger:

Eksempel 6.1.6

Overvej følgende system af to lineære ligninger i to variable over $\mathbb{R}$:

$$\begin{cases} x_1 + x_2 = 1 \\ x_1 + x_2 = 0 \end{cases}.$$

Dette system er inhomogent, da højresiden af den første ligning ikke er nul. Dette system har ingen løsninger, da det ikke er muligt, at $x_1 + x_2$ er lig med både 1 og 0 på samme tid! Hvis det var muligt, kunne vi konkludere, at $0 = 1$, hvilket ville være en modstrid.

Før Sætningerne 6.1.1 og 6.1.2 er konstruktive, skal vi finde en måde at besvare følgende tre spørgsmål på:

(i) Hvordan beskriver vi alle løsninger til et homogent system af lineære ligninger?

(ii) Hvordan afgør vi, om et inhomogent system af lineære ligninger har en løsning?

(iii) Hvordan finder vi en løsning til et inhomogent system af lineære ligninger, hvis en sådan eksisterer?

Hvis vi kan besvare disse spørgsmål, kan Sætning 6.1.2 benyttes til at beskrive alle løsninger til et inhomogent system af lineære ligninger, der har mindst én løsning. I de næste afsnit vil vi besvare disse spørgsmål.

6.2 Transformering af et system af lineære ligninger

I dette afsnit vil vi udvikle en procedure, der transformerer et givet system af lineære ligninger til et simplere system, uden at der ændres i deres løsninger. Med andre ord ønsker vi at finde en måde, hvorpå vi kan erstatte et muligvis kompliceret system af lineære ligninger med et andet, meget simplere system af lineære ligninger, mens vi sikrer os, at det oprindelige, muligvis komplicerede system har præcis de samme løsninger som det nye, simplere system.

Før vi går i gang, vil vi først introducere en kompakt måde at opskrive et system af lineære ligninger på ved hjælp af det, der kendes som *matricer*. Som en start kan du tænke på en matrix som et rektangulært skema, der indeholder elementer fra det legeme $\mathbb{F}$, man arbejder over. I et senere kapitel vil vi give en mere dybdegående diskussion af matricer.

Definition 6.2.1

For et givet lineært ligningssystem

$$\begin{cases} a_{11} \cdot x_1 & + & \cdots & + & a_{1n} \cdot x_n & = & b_1 \\ a_{21} \cdot x_1 & + & \cdots & + & a_{2n} \cdot x_n & = & b_2 \\ & & \vdots & & & & \vdots \\ a_{m1} \cdot x_1 & + & \cdots & + & a_{mn} \cdot x_n & = & b_m \end{cases}$$

betegner

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

koefficientmatricen for systemet af lineære ligninger. Matricen

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{bmatrix}$$

kaldes *totalmatricen* for systemet af lineære ligninger.

Bemærk, at begrebet totalmatrix på engelsk kaldes *augmented matrix*.

Eksempel 6.2.1

Lad os se på systemet af lineære ligninger som givet i Eksempel 6.1.5. Koefficientmatricen for dette system er givet ved

$$\begin{bmatrix} 2 & 5 & 0 & 1 \\ 3 & 0 & -1 & 0 \end{bmatrix},$$

mens totalmatricen for dette system er

$$\begin{bmatrix} 2 & 5 & 0 & 1 & 0 \\ 3 & 0 & -1 & 0 & 6 \end{bmatrix}.$$

Nogle gange tegner man en lodret streg

$$\left[\begin{array}{cccc|c} 2 & 5 & 0 & 1 & 0 \\ 3 & 0 & -1 & 0 & 6 \end{array}\right]$$

i totalmatricen for at tydeliggøre, at de sidste tal 0 og 6 kommer fra ligningssystemets højreside. Dette er dog kun et æstetisk valg.

En matrix siges at have *rækker* og *søjler*. En række er et horisontalt snit af en matrix, en søjle et vertikalt snit. For eksempel har matricen

$$\begin{bmatrix} 2 & 6 & 0 & 1 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix}$$

to rækker: den første række er givet ved $[2\ 6\ 0\ 1\ 0]$, mens den anden række er givet ved $[4\ 0\ -1\ 0\ 6]$. Tilsvarende har den fem søjler, nemlig

$$\begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 6 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ og } \begin{bmatrix} 0 \\ 6 \end{bmatrix}.$$

En matrix siges at være en $m \times n$-matrix, hvis den har præcis m rækker og præcis n søjler. Derfor er matricen, vi lige har set, en 2×5-matrix. Ud fra matrixdefinitionerne vist i Definition 6.2.1 er koefficientmatricen for et system af m lineære ligninger i n variable en $m \times n$-matrix. Tilsvarende er dets totalmatrix en $m \times (n+1)$-matrix. Den har nemlig en søjle mere end koefficientmatricen, der indeholder b_i'erne fra højresiderne af de lineære ligninger.

Lad os nu vende tilbage til vores primære mål: at transformere et system af lineære ligninger over et legeme $\mathbb{F}$ om til et simplere system uden at ændre på løsningsmængden. Ideen hertil er gradvist at transformere et hvilket som helst givet system over $\mathbb{F}$ til et meget simplere system, hvor vi ved hvert trin sikrer, at løsningsmængden ikke ændres. De operationer, vi vil benyttes os af til at transformere systemerne, er af tre typer:

1. Byt to ligninger.
2. Multiplicér en given ligning med en konstant forskellig fra nul fra $\mathbb{F}$.

3. Læg et multiplum af en ligning til en anden.

Lad os forklare i flere detaljer, hvad disse tre operationer gør. Den første tager to lineære ligninger fra et givet system, lad os sige R_i og R_j, og bytter om på dem. Efter denne operation finder man derfor R_j på position i og R_i på position j. Vi betegner denne operation ved $R_i \leftrightarrow R_j$.

Eksempel 6.2.2

Lad os illustrere bytteoperationen på systemet fra Eksempel 6.1.5:

$$\begin{cases} 2 \cdot x_1 & + & 6 \cdot x_2 & + & 0 \cdot x_3 & + & 1 \cdot x_4 & = & 0 \\ 4 \cdot x_1 & + & 0 \cdot x_2 & + & (-1) \cdot x_3 & + & 0 \cdot x_4 & = & 6 \end{cases}.$$

Her kan vi udføre operationen $R_1 \leftrightarrow R_2$ og opnå systemet

$$\begin{cases} 4 \cdot x_1 & + & 0 \cdot x_2 & + & (-1) \cdot x_3 & + & 0 \cdot x_4 & = & 6 \\ 2 \cdot x_1 & + & 6 \cdot x_2 & + & 0 \cdot x_3 & + & 1 \cdot x_4 & = & 0 \end{cases}.$$

Hvis vi, hvilket typisk vil være nemmere, arbejder med systemets totalmatrix, er effekten af operationen $R_1 \leftrightarrow R_2$, at totalmatricen

$$\begin{bmatrix} 2 & 6 & 0 & 1 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix}$$

erstattes med

$$\begin{bmatrix} 4 & 0 & -1 & 0 & 6 \\ 2 & 6 & 0 & 1 & 0 \end{bmatrix}.$$

Operationen $R_1 \leftrightarrow R_2$ bytter simpelthen rundt på den første og den anden række i totalmatricen. Vi vil normalt skrive dette som følger:

$$\begin{bmatrix} 2 & 6 & 0 & 1 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix} \xrightarrow[R_1 \leftrightarrow R_2]{} \begin{bmatrix} 4 & 0 & -1 & 0 & 6 \\ 2 & 6 & 0 & 1 & 0 \end{bmatrix}.$$

Den anden operation, vi vil benytte til at reducere et ligningssystem, multiplicerer en af de givne lineære ligninger med en konstant $c \in \mathbb{F}$, som er *forskellig fra nul* (med andre ord $c \in \mathbb{F} \setminus \{0\}$). Dette betyder simpelthen, at man erstatter den lineære ligning R_j, for eksempel givet ved $a_{j1}x_1 + \cdots + a_{jn}x_n = b_j$, med den lineære ligning $ca_{j1}x_1 + \cdots + ca_{jn}x_n = cb_j$ (som for nemheds skyld blot betegnes $c \cdot R_j$). Vi betegner denne operation ved $R_j \leftarrow c \cdot R_j$.

Eksempel 6.2.3

Lad os illustrere operationen $R_1 \leftarrow (1/2) \cdot R_1$ på systemet fra Eksempel 6.1.5. Dette svarer til at erstatte systemet

$$\begin{cases} 2 \cdot x_1 & + & 6 \cdot x_2 & + & 0 \cdot x_3 & + & 1 \cdot x_4 & = & 0 \\ 4 \cdot x_1 & + & 0 \cdot x_2 & + & (-1) \cdot x_3 & + & 0 \cdot x_4 & = & 6 \end{cases}$$

med

$$\begin{cases} 1 \cdot x_1 & + & 3 \cdot x_2 & + & 0 \cdot x_3 & + & 1/2 \cdot x_4 & = & 0 \\ 4 \cdot x_1 & + & 0 \cdot x_2 & + & (-1) \cdot x_3 & + & 0 \cdot x_4 & = & 6 \end{cases}.$$

I matrixnotation opnår vi

$$\begin{bmatrix} 2 & 6 & 0 & 1 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix} \underset{R_1 \leftarrow (1/2) \cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix}.$$

Effekten af operationen $R_1 \leftarrow (1/2) \cdot R_1$ på totalmatricen er, at alle elementer i den første række ganges med $1/2$. Vi bruger pilen $\longrightarrow$ til at indikere et trin, når vi ændrer matricen. Fremover vil vi bruge pilen $\longrightarrow$, hver gang en operation udføres, mens vi gradvist ændrer matricen. Under pilen skriver vi, hvilken operation der er i brug (i dette tilfælde $R_1 \leftarrow (1/2) \cdot R_1$).

Slutteligt betyder den tredje operation, hvor ligning R_j lægges d gange til ligning R_i (hvor $i \neq j$, og $d \in \mathbb{F}$) simpelthen, at den lineære ligning R_i givet ved $a_{i1}x_1 + \cdots + a_{in}x_n = b_i$ erstattes af ligningen $(a_{i1} + da_{j1})x_1 + \cdots + (a_{in} + da_{jn})x_n = b_i + db_j$. Man kan kort sige, at den lineære ligning R_i erstattes af $R_i + d \cdot R_j$, med andre ord $R_i \leftarrow R_i + d \cdot R_j$.

Eksempel 6.2.4

Lad os igen arbejde på systemet fra Eksempel 6.1.5 for at illustrere effekten af operationen $R_1 \leftarrow R_1 + 2 \cdot R_2$. Dette svarer til at erstatte systemet

$$\begin{cases} 2 \cdot x_1 & + & 6 \cdot x_2 & + & 0 \cdot x_3 & + & 1 \cdot x_4 & = & 0 \\ 4 \cdot x_1 & + & 0 \cdot x_2 & + & (-1) \cdot x_3 & + & 0 \cdot x_4 & = & 6 \end{cases}$$

med

$$\begin{cases} 10 \cdot x_1 & + & 6 \cdot x_2 & + & (-2) \cdot x_3 & + & 1 \cdot x_4 & = & 12 \\ 4 \cdot x_1 & + & 0 \cdot x_2 & + & (-1) \cdot x_3 & + & 0 \cdot x_4 & = & 6 \end{cases}.$$

I matrixnotation opnår vi

$$\begin{bmatrix} 2 & 6 & 0 & 1 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix} \underset{R_1 \leftarrow R_1 + 2 \cdot R_2}{\longrightarrow} \begin{bmatrix} 10 & 6 & -2 & 1 & 12 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix}.$$

Effekten af operationen $R_1 \leftarrow R_1 + 2 \cdot R_2$ på totalmatricen er derfor, at den første række erstattes af den første række plus to gange den anden række.

Som det fremgår af eksemplerne, kan effekten af de tre operationer $R_i \leftrightarrow R_j$, $R_j \leftarrow c \cdot R_j$ og $R_i \leftarrow R_i + d \cdot R_j$ opfattes som elementære operationer udført på elementerne i rækkerne i totalmatricen for det system af lineære ligninger, vi startede med. Derfor kaldes de *elementære rækkeoperationer*. Dette er også grunden til, at vi benyttede stort R i symbolerne $R_1, \ldots, R_m$ for de lineære ligninger i vores system: R var simpelthen inspireret af det første bogstav i ordet "række".

Lad os nu sikre os, at løsningsmængden til det nye system rent faktisk er identisk med løsningsmængden til det oprindelige system af lineære ligninger, når vi benytter disse elementære rækkeoperationer. Lad os opstille dette som en sætning.

Sætning 6.2.1

Lad $R_1, \ldots, R_m$ være et system af m lineære ligninger i n variable over et legeme $\mathbb{F}$. Lad desuden i og j være to forskellige heltal mellem 1 og m. Da har ethvert system, der opnås ved anvendelse af operationerne $R_i \leftrightarrow R_j$, $R_j \leftarrow c \cdot R_j$ med $c \in \mathbb{F} \setminus \{0\}$ eller $R_i \leftarrow R_i + d \cdot R_j$ med $d \in \mathbb{F}$, samme løsningsmængde som det oprindelige system.

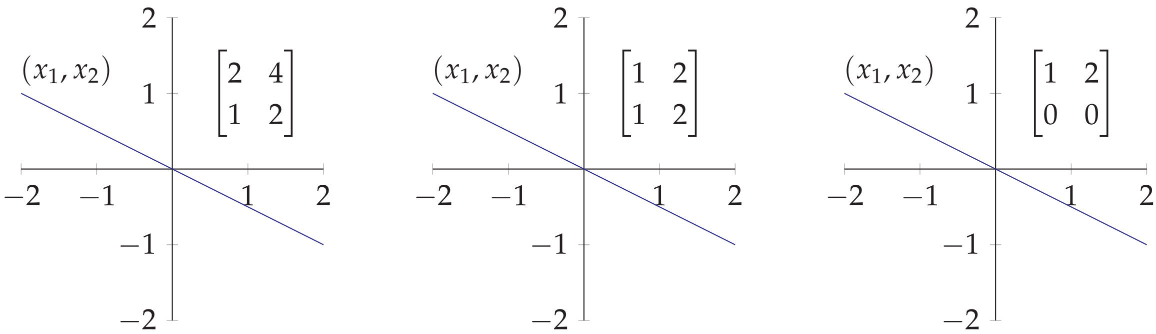
Bevis. Vi beviser kun sætningen for den elementære operation $R_i \leftarrow R_i + d \cdot R_j$. Læseren opfordres til at kontrollere, at sætningen også er sand for de to resterende elementære operationer. Vi skal vise, at løsningsmængden til systemet af lineære ligninger $R_1, \ldots, R_{i-1}, R_i, R_{i+1}, \ldots, R_m$ er den samme som løsningsmængden til systemet givet ved $R_1, \ldots, R_{i-1}, R_i + d \cdot R_j, R_{i+1}, \ldots, R_m$. Lad os betegne den første løsningsmængde ved S og den anden ved T. Vi ønsker at vise, at $S = T$.

Først og fremmest hævder vi, at $S \subseteq T$. Derfor vælger vi vilkårligt $(v_1, \ldots, v_n) \in S$. Vi vil vise, at $(v_1, \ldots, v_n) \in T$. Med andre ord, hvis vi antager, at $(v_1, \ldots, v_n) \in \mathbb{F}^n$ er en fælles løsning til de lineære ligninger $R_1, \ldots, R_m$, skal vi vise, at den også er en fælles løsning til de lineære ligninger $R_1, \ldots, R_{i-1}, R_i + d \cdot R_j, R_{i+1}, \ldots, R_m$. Men så skal vi kun vise, at $(v_1, \ldots, v_n)$ er en løsning til $R_i + d \cdot R_j$. At dette er sandt, skyldes, at hvis $(v_1, \ldots, v_n)$ er en fælles løsning til R_i og R_j, så er den også en løsning til $R_i + d \cdot R_j$ for enhver konstant $d \in \mathbb{F}$. Derfor har vi, at $(v_1, \ldots, v_n) \in T$. Da vi valgte $(v_1, \ldots, v_n) \in S$ vilkårligt, medfører dette, at $S \subseteq T$.

Nu hævder vi, at $T \subseteq S$. Vi vælger vilkårligt $(v_1, \ldots, v_n) \in T$ og vil nu vise, at $(v_1, \ldots, v_n) \in S$. Dette betyder, at vi kan antage, at $(v_1, \ldots, v_n) \in \mathbb{F}^n$ er en fælles løsning til de lineære ligninger $R_1, \ldots, R_{i-1}, R_i + d \cdot R_j, R_{i+1}, \ldots, R_m$. Vi skal vise, at $(v_1, \ldots, v_n)$ er en løsning til R_i. Dette er dog sandt, da $R_i = (R_i + d \cdot R_j) - d \cdot R_j$. Derfor har vi, at $(v_1, \ldots, v_n) \in S$. Da vi valgte $(v_1, \ldots, v_n) \in T$ vilkårligt, medfører dette, at $T \subseteq S$.

Nu hvor vi har vist, at $S \subseteq T$ og $T \subseteq S$, medfører Lemma 2.1.1, at $S = T$, hvilket er, hvad vi ønskede at vise. $\square$

Sætning 6.2.1 er illustreret i Figur 6.2. I denne figur angiver den blå linje løsningsmængden $(x_1, x_2) \in \mathbb{R}^2$ til det homogene system af ligninger givet ved $\left\{ \begin{array}{rcrcr} 2x_1 & + & 4x_2 & = & 0 \\ x_1 & + & 2x_2 & = & 0 \end{array} \right.$, som har koefficientmatricen $\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$. Når dette ligningssystem transformeres til et nyt system via elementære rækkeoperationer, forbliver løsningsmængden den samme.



(a) Det oprindelige system (b) $R_1 \leftarrow (1/2)R_1$. (c) $R_2 \leftarrow R_1 + (-1)R_2$.

Figur 6.2: Løsningsmængden til et ligningssystem $\begin{cases} 2x_1 & + & 4x_2 = & 0 \\ x_1 & + & 2x_2 = & 0 \end{cases}$ ændres ikke, når systemet transformeres ved brug af elementære rækkeoperationer.

Det viser sig, at med disse tre temmeligt elementære operationer ved hånden kan vi finde løsningsmængden til ethvert system af lineære ligninger. Brugen af blot én elementær rækkeoperation vil sikkert ikke simplificere et system af lineære ligninger særligt meget, men ideen er, at hvis vi benytter adskillige elementære rækkeoperationer i træk, kan vi transformere ethvert givet system til et væsentligt simplere system. I de næste afsnit vil vi se hvordan, men inden da tager vi et kig på et eksempel.

Eksempel 6.2.5

Lad os genbesøge Eksempel 6.1.5. Dér betragtede vi følgende system af 2 ligninger i 4 variable over $\mathbb{R}$:

$$\begin{cases} 2 \cdot x_1 & + & 6 \cdot x_2 & + & 0 \cdot x_3 & + & 1 \cdot x_4 & = & 0 \\ 4 \cdot x_1 & + & 0 \cdot x_2 & + & (-1) \cdot x_3 & + & 0 \cdot x_4 & = & 6 \end{cases}.$$

Lad os reducere dette system ved at anvende elementære rækkeoperationer. Som Sætning 6.2.1 fortæller, ændrer dette ikke systemets løsningsmængde. Da det er langt mere kompakt at arbejde med systemets totalmatrix, vil vi gøre det.

Vi lægger ud med at anvende transformationen $R_1 \leftarrow (1/2) \cdot R_1$, hvorved vi opnår totalmatricen:

$$\begin{bmatrix} 2 & 6 & 0 & 1 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix} \underset{R_1 \leftarrow (1/2) \cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix}.$$

Vi valgte denne operation på denne række, da vi ønskede at opnå en et-tal som første element i første række. Dette gør det nemlig nemt at eliminere variablen x_1 fra anden ligning. I det næste trin ønsker vi altså at opnå et nul som første element i anden række. Dette opnås ved anvendelse af den elementære rækkeoperation $R_2 \leftarrow R_2 - 4 \cdot R_1$, da vi derved får

$$\begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix} \underset{R_2 \leftarrow R_2 - 4 \cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 0 & -12 & -1 & -2 & 6 \end{bmatrix}.$$

Nu reducerer vi yderligere ved at omdanne koefficienten foran x_2 i anden ligning til et et-tal. Altså ønsker vi nu at opnå, at andet element i anden række bliver lig med ét. For at opnå dette, anvendes $R_2 \leftarrow (-1/12) \cdot R_2$:

$$\begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 0 & -12 & -1 & -2 & 6 \end{bmatrix} \underset{R_2 \leftarrow (-1/12) \cdot R_2}{\longrightarrow} \begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 0 & 1 & 1/12 & 2/12 & -6/12 \end{bmatrix}.$$

Brøkerne i den resulterende matrix kan reduceres en smule, så vi kunne også have skrevet:

$$\begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 0 & -12 & -1 & -2 & 6 \end{bmatrix} \underset{R_2 \leftarrow (-1/12) \cdot R_2}{\longrightarrow} \begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 0 & 1 & 1/12 & 1/6 & -1/2 \end{bmatrix}.$$

Det transformerede system er nu snart reduceret mest muligt, men vi kan stadig bruge anden ligning til at fjerne x_2-leddet i første ligning ved hjælp af $R_1 \leftarrow R_1 - 3 \cdot R_2$:

$$\begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 0 & 1 & 1/12 & 1/6 & -1/2 \end{bmatrix} \underset{R_1 \leftarrow R_1 - 3 \cdot R_2}{\longrightarrow} \begin{bmatrix} 1 & 0 & -1/4 & 0 & 3/2 \\ 0 & 1 & 1/12 & 1/6 & -1/2 \end{bmatrix}.$$

Det transformerede men tilsvarende system af lineære ligninger er nu:

$$\begin{cases} x_1 + (-1/4) \cdot x_3 = 3/2 \\ x_2 + (1/12) \cdot x_3 + (1/6) \cdot x_4 = -1/2 \end{cases}.$$

Det er vigtigt at huske på, at løsningsmængden til dette sidste system ifølge Sætning 6.2.1 er nøjagtig den samme som løsningsmængden til det system, vi startede med.

Det er nemt at finde løsninger $(v_1, v_2, v_3, v_4) \in \mathbb{R}^4$ til dette sidste system: vælg blot $v_3, v_4 \in \mathbb{R}$ som du vil, og brug derefter de lineære ligninger til at løse for v_1 og v_2. For eksempel, hvis vi vælger $v_3 = 0$ og $v_4 = 3$, får vi, at $v_1 = (1/4)v_3 + 3/2 = 3/2$ og $v_2 = -(1/12)v_3 + (-1/6)v_4 - 1/2 = -1$. Derfor er $(3/2, -1, 0, 3)$ en løsning til systemet. Flere (faktisk alle) løsninger kan opnås på denne måde: vælg en hvilken som helst værdi for v_3 og v_4, som du ønsker, og bestem de tilsvarende v_1 og v_2 ved hjælp af ligningerne $v_1 = (1/4)v_3 + 3/2$ og $v_2 = -(1/12)v_3 + (-1/6)v_4 - 1/2$.

Dette eksempel viser, at det kan hjælpe meget at reducere et givet system af lineære ligninger først, før man forsøger at løse det.

6.3 Den reducerede trappeform af en matrix

Vi har set i Eksempel 6.2.5, at brugen af elementære rækkeoperationer kan hjælpe os med at finde frem til løsningsmængden til et system af lineære ligninger. Hvad vi nu vil gøre, er at vise, at denne tilgang altid virker. I stedet for at arbejde med systemer af lineære ligninger, vil vi arbejde med koefficient- og totalmatricer for systemerne. Vi har set, at hvis ligningssystemet består af m lineære ligninger i n variable, så er koefficientmatricen en $m \times n$-matrix, mens totalmatricen er en $m \times (n+1)$-matrix. Elementerne i disse matricer kommer fra $\mathbb{F}$, some er det legeme, vi arbejder over. Som nævnt tidligere vil vi typisk arbejde med enten $\mathbb{F} = \mathbb{R}$, de reelle tal, eller $\mathbb{F} = \mathbb{C}$, de komplekse tal. Mængden af alle $m \times n$-matricer med elementer i $\mathbb{F}$ vil blive betegnet $\mathbb{F}^{m \times n}$. I formler og udtryk vil vi typisk bruge fed skrift for matricer, så som $\mathbf{A}, \mathbf{B}, \ldots$.

Vi lægger ud med at definere en særlig type matrix:

Definition 6.3.1

Lad $\mathbb{F}$ være et legeme og $\mathbf{A} \in \mathbb{F}^{m \times n}$ en $m \times n$-matrix med elementer i $\mathbb{F}$. Man siger, at $\mathbf{A}$ er på *reduceret trappeform*, hvis alle følgende regler er opfyldt.

(i) Hvis en række i matricen kun indeholder nuller, er den placeret nederst i matricen. Sådanne rækker kaldes nulrækker.

(ii) Det første element fra venstre, der ikke er nul, i en række, der ikke er en nulrække, er lig med 1. Dette element kaldes rækkens *pivot*.

(iii) Pivoterne i to rækker, der ikke er nulrækker, forekommer ikke i samme søjle. Desuden er pivoten i en øvre række længere til venstre end pivoten i en nedre række.

(iv) Hvis en søjle i matricen indeholder en pivot, så er alle andre elementer i den søjle 0.

Bemærk, at begrebet reduceret trappematrix på engelsk kaldes *reduced row-echelon form*. En matrix, der opfylder de førstnævnte tre regler men ikke nødvendigvis den fjerde regel, siges blot at være på *trappeform*.

Eksempel 6.3.1

1×4-matricerne $[0\ 0\ 0\ 0]$ og $[0\ 0\ 1\ 5]$ er begge på reduceret trappeform. Også 2×5-matricen

$$\begin{bmatrix} 1 & 0 & -1/4 & 0 & 3/2 \\ 0 & 1 & 1/12 & 1/6 & -1/2 \end{bmatrix},$$

som vi nåede frem til sidst i Eksempel 6.2.5, er på reduceret trappeform.

Et eksempel på en 1×4-matrix, der ikke er på reduceret trappeform, er: $[0\ 0\ 2\ 0]$. Det første element fra venstre, der er forskelligt fra nul, i den første (og eneste) række er ikke lig med 1. Et eksempel på en 3×4-matrix, der ikke er på reduceret trappeform, er:

$$\begin{bmatrix} 1 & 0 & 2 & 0 & 4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Denne matrix er på trappeform men ikke på reduceret trappeform. Problemet her er den tredje søjle. Denne søjle indeholder en pivot, nemlig pivoten i anden række, men udover pivoten indeholder denne søjle endnu et element, der ikke er nul (nemlig 2-tallet).

Grunden til, at reducerede trappeformer er så vigtige, ses af følgende resultat:

Sætning 6.3.1

Lad $\mathbf{A} \in \mathbb{F}^{m \times n}$ være en matrix. Da kan $\mathbf{A}$ bringes på reduceret trappeform ved hjælp af elementære rækkeoperationer.

Bevis. Lad os skitsere beviset her. Strategien i beviset er først at bringe matricen på trappeform og derefter på reduceret trappeform. Lad os derfor først vise, at vi kan benytte elementære rækkeoperationer til at bringe matricen $\mathbf{A}$ på trappeform. For at gøre dette vil vi udføre induktion på m, antallet af rækker.

Hvis $m = 1$ (basistrinnet af induktionen), er det kun muligt for $\mathbf{A}$ ikke at være på trappeform, hvis rækken indeholder et element forskelligt fra nul, der er det første element fra venstre forskelligt fra nul, lad os kalde det c, som ikke er lig med ét. Men i så fald vil operationen $R_1 \leftarrow c^{-1} \cdot R_1$ bringe $\mathbf{A}$ på trappeform.

Antag til induktionstrinnet, at $m > 1$ er givet, og at sætningen er sand for $(m-1) \times n$-matricer. Hvis alle elementer i matricen $\mathbf{A}$ er nul, er den allerede på trappeform (og faktisk også på reduceret trappeform), og vi er færdige. Lad os derfor antage, at matricen $\mathbf{A}$ har mindst ét element forskelligt fra nul. Vi starter med at vælge den mindst mulige værdi af j, der opfylder, at den j'te søjle i $\mathbf{A}$ indeholder et element forskelligt fra nul. Hvis $j > 1$, så er de første $j-1$ søjler i $\mathbf{A}$ udelukkende nulsøjler. Herefter vælger vi den mindst mulige værdi af i, der opfylder, at a_{ij}, altså det (i, j)'te element i $\mathbf{A}$, ikke er nul. Nu udfører vi operationen $R_1 \leftrightarrow R_i$. Den første række i den resulterende matrix har et element forskelligt fra nul på sin j'te position, som vi kan kalde c, og nuller som elementerne fra position 1 til $j-1$. Dernæst udfører vi operationen $R_1 \leftarrow c^{-1}R_1$, hvilket omdanner det j'te element i den første række til 1. Hvis ikke alle elementer under dette 1-tal er nuller, bruger vi elementære operationer af formen $R_j \leftarrow R_j + dR_1$ for passende valgte $d \in \mathbb{F}$ til at transformere matricen yderligere, indtil vi har en matrix, hvor der kun optræder nuller under pivoten i første række. Vi har nu transformeret matricen **A** til en matrix **B** på formen

$$\mathbf{B} = \begin{bmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{bmatrix}.$$

Med denne notation er den første del af matricen **B**

$$\begin{bmatrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix}.$$

Dette afspejler det faktum, at de første $j-1$ søjler i **B** er nulsøjler. Det er dog ikke hensigten med notationen, at den første del af **B** indeholder mindst to nulsøjler. Faktisk består denne del kun af én nulsøjle, hvis $j=2$, da vi så har $j-1=1$. I det tilfælde, at $j=1$, er den første søjle i matricen **B** faktisk ikke en nulsøjle overhovedet, da dens første koordinat er 1 med nuller under.

Uanset hvad den præcise værdi af j er, fortsætter vi ved at fjerne den første række i matricen **B**. Den $(m-1) \times n$-matrix, vi dermed opnår, betegnes **C**. Ved at benytte induktionshypotesen kan vi konkludere, at vi kan bruge elementære rækkeoperationer til at transformere matricen **C** om til en matrix $\hat{\mathbf{C}}$, der er på trappeform. Sættes den første række fra **B** tilbage, opnår vi en $m \times n$-matrix, som vi kan betegne $\hat{\mathbf{A}}$, der er på trappeform.

Dette afslutter det induktive bevis for, at enhver matrix kan bringes på trappeform ved hjælp af elementære rækkeoperationer. Det, vi stadig mangler at gøre, er at bringe denne matrix på reduceret trappeform. Vi ved ifølge definitionen på trappeformen, at pivoterne i to rækker, som ikke er nulrækker, i matricen $\hat{\mathbf{A}}$ ikke forekommer i samme søjle, og desuden at pivoten i en øvre række er længere til venstre end pivoten i en lavere række. Derfor er elementerne under en pivot i matricen $\hat{\mathbf{A}}$ lig nul. Dog er elementerne over en pivot i denne matrix ikke nødvendigvis nul. Vi kan opnå, at disse også bliver nul, ved at benytte elementære rækkeoperationer af typen $R_i \leftarrow R_i + dR_j$, hvor række R_j indeholder en pivot, og $i < j$. Vi starter med rækken, der indeholder den pivot, der er længst til højre, og ønsker at opnå udelukkende nuller over denne pivot. Derefter vil vi arbejde os mod venstre, idet vi håndterer én pivot ad gangen. Når vi når til pivoten længst til venstre og har udført den skitserede procedure for denne pivot også, vil den opnåede matrix være på reduceret trappeform. □

Som et eksempel kan vi se på Eksempel 6.2.5. Dér benyttede vi os af elementære rækkeoperationer for at bringe en matrix på reduceret trappeform. Der er i princippet mange forskellige måder at benytte elementære rækkeoperationer på for at transformere en given matrix **A** til sin reducerede trappeform. Dog viser det sig, at resultatet altid er det samme for en given matrix **A**. Derfor kan vi tale om *den* reducerede trappeform af en matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$. Særligt er følgende definition nu berettiget:

Definition 6.3.2

Lad $\mathbb{F}$ være et legeme og $\mathbf{A} \in \mathbb{F}^{m \times n}$ en matrix. Da defineres *rangen* af $\mathbf{A}$, betegnet ved $\rho(\mathbf{A})$, som antallet af pivoter i den reducerede trappeform af $\mathbf{A}$.

Beviset for Sætning 6.3.1 var algoritmisk i sin natur og kan angives som en algoritme. Lad os opskrive pseudo-koden for en algoritme, der bestemmer en trappeform af en matrix. I pseudo-koden er betegnelsen ref en forkortelse for 'row-echelon form', den engelske betegnelse for trappeform. Bemærk, hvor tæt algoritmen følger den første del af beviset for Sætning 6.3.1. Man kunne udvide algoritmen og opnå pseudo-koden for en algoritme, der bestemmer den reducerede trappeform af en matrix, men det vil vi ikke gøre her.

Algoritme 9 til bestemmelse af en trappeform af en matrix

Input: Positive heltal m, n og en $m \times n$-matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$
Output: $\mathrm{ref}(\mathbf{A})$, en trappeform af $\mathbf{A}$
if $\mathbf{A} = \mathbf{0}$ **then**
 $\mathrm{ref}(\mathbf{A}) \leftarrow \mathbf{0}$,
if $m = 1$ og $\mathbf{A} \neq \mathbf{0}$ **then**
 $j \leftarrow$ mindste søjleindeks, således at $\mathbf{A}_{1j} \neq 0$
 $\mathrm{ref}(\mathbf{A}) \leftarrow (\mathbf{A}_{1j})^{-1} \cdot \mathbf{A}$
if $m > 1$ og $\mathbf{A} \neq \mathbf{0}$ **then**
 $j \leftarrow$ mindste ℓ, således at en række i $\mathbf{A}$ har et ℓ'te element forskelligt fra nul
 $i \leftarrow$ mindste i, således at den i'te række i $\mathbf{A}$ har et j'te element forskelligt fra nul
 $\mathbf{B} \leftarrow$ matricen opnået fra $\mathbf{A}$ ved anvendelse af $R_1 \leftrightarrow R_i$
 $b \leftarrow$ det i'te element i den første række i $\mathbf{B}$
 $\mathbf{B} \leftarrow$ matricen opnået fra $\mathbf{B}$ ved anvendelse af $R_1 \leftarrow b^{-1} \cdot R_1$
 $\mathbf{r} \leftarrow$ den første række i $\mathbf{B}$
 for $i = 2...m$ **do**
  $b \leftarrow$ det første element i den i'te række i $\mathbf{B}$
  $\mathbf{B} \leftarrow$ matricen opnået fra $\mathbf{B}$ ved anvendelse af $R_i \leftarrow R_i - bR_1$
 $\mathbf{C} \leftarrow$ matricen opnået fra $\mathbf{B}$ ved sletning af den første række
 $\mathbf{C} \leftarrow \mathrm{ref}(\mathbf{C})$ (her kalder algoritmen sig selv rekursivt)
 $\mathrm{ref}(\mathbf{A}) \leftarrow$ matricen opnået ved tilføjelse af $\mathbf{r}$ øverst i $\mathbf{C}$

Vi har i pseudo-koden benyttet en såkaldt for-løkke, der gentager kodelinjerne for hver heltalsforøgelse i det angivne variabelinterval.

6.4 Bestemmelse af alle løsninger til systemer af lineære ligninger

Indtil nu har vi typisk skrevet elementer fra $\mathbb{F}^n$ som n-tupler $(a_1, \ldots, a_n)$. Det er ret almindeligt at identificere $\mathbb{F}^n$ med $\mathbb{F}^{n \times 1}$, det vil sige at identificere et n-tupel med en $n \times 1$-matrix. En sådan matrix indeholder kun én søjle. Dette betyder for eksempel, at:

$$(1, 2, 4, 7) \text{ identificeres med } \begin{bmatrix} 1 \\ 2 \\ 4 \\ 7 \end{bmatrix}.$$

En lille advarsel er på sin plads. Selvom vi altid vil opfatte $\mathbb{F}^n$ og $\mathbb{F}^{n\times 1}$ som samme objekt, foretrækker nogle bøger at opfatte $\mathbb{F}^n$ som $\mathbb{F}^{1\times n}$.

Når vi udfører elementære rækkeoperationer, ganger vi en gang imellem rækker i en matrix med et element c fra $\mathbb{F}$ eller lægger en række til en anden række. Lignende operationer kan udføres på søjler i en matrix. Det er sædvane at definere

$$c \cdot \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} c \cdot a_1 \\ \vdots \\ c \cdot a_n \end{bmatrix} \text{ and } \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} a_1' \\ \vdots \\ a_n' \end{bmatrix} = \begin{bmatrix} a_1 + a_1' \\ \vdots \\ a_n + a_n' \end{bmatrix}.$$

Denne notation, kombineret med teorien om reducerede trappeformer, gør det muligt at afgøre, om et givet system af lineære ligninger har løsninger, og i så fald at skrive alle løsninger ned på en systematisk måde. Lad os starte med at bestemme, hvornår et system har en løsning.

Sætning 6.4.1

Lad et system af m lineære ligninger i n variable over et legeme $\mathbb{F}$ være givet. Betegn ved $\mathbf{A}$ systemets koefficientmatrix og ved $[\mathbf{A}|\mathbf{b}]$ dets totalmatrix. Hvis ikke $\mathbf{A}$ og $[\mathbf{A}|\mathbf{b}]$ har samme rang, så har systemet ingen løsning.

Bevis. Vi ved fra Sætning 6.3.1, at der findes en sekvens af elementære rækkeoperationer, der bringer matricen $\mathbf{A}$ på sin reducerede trappeform, betegnet $\hat{\mathbf{A}}$. Da de første n søjler i totalmatricen $[\mathbf{A}|\mathbf{b}]$ er identiske med de tilsvarende søjler i koefficientmatricen $\mathbf{A}$, vil anvendelse af præcis de samme elementære rækkeoperationer på $[\mathbf{A}|\mathbf{b}]$ frembringe en matrix, som vi kan kalde $\mathbf{B}$, hvis første n søjler er identiske med de tilsvarende i den reducerede trappeform af $\mathbf{A}$. Derfor kan vi skrive $\mathbf{B} = [\hat{\mathbf{A}}|\hat{\mathbf{b}}]$ for enhver $\hat{\mathbf{b}} \in \mathbb{F}^m$. Lad os betegne det nederste element i $\hat{\mathbf{b}}$ ved $\hat{b}_m$. Hvis den nederste række i $\hat{\mathbf{A}}$ indeholder en pivot, så er matricen $[\hat{\mathbf{A}}|\hat{\mathbf{b}}]$ på reduceret trappeform. Men så ser vi, at matricerne $\mathbf{A}$ og $[\mathbf{A}|\mathbf{b}]$ har samme rang, i modstrid med antagelsen i sætningen om, at $\mathbf{A}$ og $[\mathbf{A}|\mathbf{b}]$ ikke har samme rang. Derfor kan vi antage, at den nederste række i $\hat{\mathbf{A}}$ ikke indeholder en pivot, hvilket simpelthen betyder, at denne række er en nulrække. Hvis den sidste række i $\hat{\mathbf{A}}$ ikke indeholder en pivot, og $\hat{b}_m = 0$, så er matricen $[\hat{\mathbf{A}}|\hat{\mathbf{b}}]$ på reduceret trappeform, og vi kan konkludere, at $\rho(\mathbf{A}) = \rho([\mathbf{A}|\mathbf{b}])$, hvilket igen fører til en modstrid. Derfor kan vi antage, at den nederste række i $\hat{\mathbf{A}}$ ikke indeholder en pivot, og at $\hat{b}_m \neq 0$. Men så svarer den nederste række i matricen $[\hat{\mathbf{A}}|\hat{\mathbf{b}}]$ til ligningen $0 \cdot x_1 + \cdots 0 \cdot x_m = \hat{b}_m$. Da denne ligning ikke har nogen løsning, medfører Sætning 6.2.1, at det system, vi startede med, heller ikke har nogen løsning. □

Eksempel 6.4.1

Som i Eksempel 6.1.6 vil vi se på følgende system af to lineære ligninger i to variable over $\mathbb{R}$:

$$\begin{cases} x_1 + x_2 = 1 \\ x_1 + x_2 = 0 \end{cases}.$$

Vi har allerede set i Eksempel 6.1.6, at dette system ikke har nogen løsninger. Lad os nu prøve at bekræfte dette ved hjælp af Sætning 6.4.1. Totalmatricen $[\mathbf{A}|\mathbf{b}]$ er givet ved

$$[\mathbf{A}|\mathbf{b}] = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Ved anvendelse af rækkeoperationen $R_1 \leftrightarrow R_2$ efterfulgt af $R_2 \leftarrow R_2 - R_1$ opnår vi totalmatricens reducerede trappeform:

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Derfor er $\rho([\mathbf{A}|\mathbf{b}]) = 2$. Koefficientmatricens reducerede trappeform er matricen

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix},$$

som kan opnås fra $\mathbf{A}$ ved anvendelse af operationen $R_2 \leftarrow R_2 - R_1$. Derfor er $\rho(\mathbf{A}) = 1$. Da $\rho(\mathbf{A}) \neq \rho([\mathbf{A}|\mathbf{b}])$, medfører Sætning 6.4.1, at det system, vi startede med, ganske rigtigt ikke har nogen løsning.

Når $\mathbf{A}$ og $[\mathbf{A}|\mathbf{b}]$ har samme rang, kan vi bruge teorien om reducerede trappeformer til at beskrive en løsning eksplicit. Lad os først se på et konkret eksempel.

Eksempel 6.4.2

Lad os se på et system af tre lineære ligninger i fire variable over $\mathbb{R}$, hvis totalmatrix allerede er på reduceret trappeform:

$$\begin{cases} x_1 + 2 \cdot x_2 + 3 \cdot x_4 = 5 \\ x_3 + 4 \cdot x_4 = 6 \\ 0 = 0 \end{cases}.$$

Vi ser her, at koefficientmatricen $\mathbf{A}$ og totalmatricen $[\mathbf{A}|\mathbf{b}]$ er

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \text{ henholdsvis } [\mathbf{A}|\mathbf{b}] = \begin{bmatrix} 1 & 2 & 0 & 3 & 5 \\ 0 & 0 & 1 & 4 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Da begge allerede er på reduceret trappeform, kan vi straks bestemme rangene af disse matricer og konkludere, at $\rho(\mathbf{A}) = \rho([\mathbf{A}|\mathbf{b}]) = 2$. Sætning 6.4.1 er derfor ikke gældende her, og vi kan endnu ikke konkludere noget om eksistensen af løsninger. Men en løsning kan nemt bestemmes på følgende måde: først omskrives ligningerne som følger:

$$\begin{cases} x_1 &= 5 - 2 \cdot x_2 - 3 \cdot x_4 \\ x_3 &= 6 - 4 \cdot x_4 \end{cases}.$$

Nu kan vi frit vælge værdier af $x_2 = v_2$ og $x_4 = v_4$ for $v_2, v_4 \in \mathbb{R}$, og derefter beregne de resulterende værdier for x_1 og x_3. Vælger vi $v_2 = v_4 = 0$, får vi for eksempel løsningen

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ 6 \\ 0 \end{bmatrix}.$$

Nøjagtig samme tilgang kan benyttes generelt til at finde en løsning til et system af lineære ligninger, forudsat at koefficient- og totalmatricen har samme rang. Resultatet er følgende:

Sætning 6.4.2

Lad et system af m lineære ligninger i n variable over et legeme $\mathbb{F}$ være givet. Lad $\mathbf{A}$ være koefficientmatricen af systemet og $[\mathbf{A}|\mathbf{b}]$ dets totalmatrix, og antag, at disse matricer har samme rang ρ. Antag desuden, at pivoterne i $\mathbf{A}$'s reducerede trappeform er at finde på positionerne $(1, j_1), \ldots, (\rho, j_\rho)$, samt at de øverste ρ elementer i den sidste søjle i $[\mathbf{A}|\mathbf{b}]$'s reducerede trappeform er givet ved $\hat{b}_1, \ldots, \hat{b}_\rho$. Da er m-tuplet $(v_1, \ldots, v_n)$ defineret som

$$v_j = \begin{cases} \hat{b}_\ell & \text{hvis } j = j_\ell \text{ for } \ell = 1, \ldots, \rho, \\ 0 & \text{ellers,} \end{cases}$$

en mulig løsning på systemet.

Bevis. Ideen med beviset er simpelthen at generalisere tilgangen benyttet i Eksempel 6.4.2. Først og fremmest benytter vi ligningerne svarende til rækkerne i den reducerede trappeform af totalmatricen $[\mathbf{A}|\mathbf{b}]$ til at udtrykke variablerne x_j for $j \in \{j_1, \ldots, j_\ell\}$ ved de resterende $n - \rho$ variabler. Ved derefter at sætte alle disse resterende variabler $x_j, j \notin \{j_1, \ldots, j_\ell\}$ lig med nul, får vi, at $x_j = \hat{b}_\ell$ for $j = j_\ell$ og $\ell = 1, \ldots, \rho$. Derfor er n-tuplet $(v_1, \ldots, v_n)$ en løsning til det system, hvis totalmatrix er den reducerede trappeform af $[\mathbf{A}|\mathbf{b}]$. Ved anvendelse af Sætning 6.2.1 ser vi nu, at dette n-tupel også er en løsning på det system, vi startede med. □

Sætning 6.4.2 siger på ingen måde, at den angivne løsning er den eneste løsning. Faktisk ved vi fra Sætning 6.1.2, at der kan være flere løsninger. Husk på, at en løsning til et inhomogent system af lineære ligninger blev kaldt en partikulær løsning. Hvis systemet af lineære ligninger er inhomogent, giver Sætning 6.4.2 derfor en sådan partikulær løsning, forudsat at den eksisterer.

Korollar 6.4.3

Lad et system bestående af m lineære ligninger i n variable over et legeme $\mathbb{F}$ være givet. Lad $\mathbf{A}$ betegne koefficientmatricen for systemet og $[\mathbf{A}|\mathbf{b}]$ dets totalmatrix. Da har systemet ingen løsning, hvis og kun hvis $\mathbf{A}$ og $[\mathbf{A}|\mathbf{b}]$ ikke har samme rang.

Bevis. Det første "hvis" er faktisk Sætning 6.4.1. Med andre ord, vi har allerede set i Sætning 6.4.1, at hvis $\rho(\mathbf{A}) \neq \rho([\mathbf{A}|\mathbf{b}])$, så har systemet ingen løsninger. Omvendt, hvis $\rho(\mathbf{A}) = \rho([\mathbf{A}|\mathbf{b}])$, så medfører Sætning 6.4.2, at systemet har mindst én løsning. $\square$

Med Korollar 6.4.3 kan vi nu nøjagtigt bestemme, om et givet system af lineære ligninger har en løsning. Desuden kan vi, ved hjælp af Sætning 6.4.2, bestemme mindst én løsning, hvis sådanne løsninger eksisterer. Husk på, at vi i Sætning 6.1.2 så, at for at bestemme alle løsninger til et inhomogent system af lineære ligninger er det nok at bestemme alle løsninger til det tilsvarende homogene system af lineære ligninger samt én partikulær løsning til det inhomogene system. Derfor mangler vi nu kun at beskrive, hvordan man bestemmer alle løsninger til et homogent system af lineære ligninger. Dette er netop målet med den næste sætning, men lad os først se på et eksempel for at varme op til det.

Eksempel 6.4.3

Givet er et system af tre lineære ligninger i fire variable over $\mathbb{R}$, hvis totalmatrix allerede er på reduceret trappeform:

$$\begin{cases} x_1 + 2\cdot x_2 + 3\cdot x_4 = 0 \\ x_3 + 4\cdot x_4 = 0 \\ 0 = 0 \end{cases}.$$

Dette system ligner systemet af lineære ligninger, vi arbejdede på i Eksempel 6.4.2, men denne gang er det homogent. Særligt er koefficientmatricen for systemet ovenfor og for systemet fra Eksempel 6.4.2 de samme, og som vi så i Eksempel 6.4.2, er den på reduceret trappeform.

Det er ikke svært at bestemme alle løsninger til systemet. Da koefficientmatricen for systemet er på reduceret trappeform med pivoter i første og tredje søjle, kan vi udtrykke x_1 og x_3 ved x_2 og x_4. Mere konkret kan vi omskrive ligningerne som

$$\begin{cases} x_1 = -2\cdot x_2 - 3\cdot x_4 \\ x_3 = -4\cdot x_4 \end{cases}.$$

Så enhver løsning $(v_1, v_2, v_3, v_4) \in \mathbb{R}^4$ til systemet opfylder

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} -2\cdot v_2 - 3\cdot v_4 \\ v_2 \\ -4\cdot v_4 \\ v_4 \end{bmatrix} = v_2 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + v_4 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix}.$$

Derfor kan vi betragte $v_2, v_4 \in \mathbb{R}$ som parametre, vi kan vælge vilkårligt, hvor hvert valg giver os en løsning til det system af lineære ligninger, vi startede med. Ved at ændre notation fra v_2 til t_1 og fra v_4 til t_2 har vi, at enhver løsning til systemet er på formen

$$t_1 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \qquad (t_1, t_2 \in \mathbb{R})$$

Omvendt, da en direkte kontrol viser, at

$$\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \text{ og } \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix}$$

er løsninger til systemet, medfører Sætning 6.1.1, at for ethvert $t_1, t_2 \in \mathbb{R}$ er udtrykket

$$t_1 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix}$$

også en løsning. Samlet set ser vi, at løsningerne til det homogene system af lineære ligninger, vi startede med, netop er de $(v_1, v_2, v_3, v_4) \in \mathbb{R}^4$, der opfylder

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = t_1 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \qquad (t_1, t_2 \in \mathbb{R}).$$

En sådan beskrivelse af løsningerne kaldes den *fuldstændige løsning* til det homogene system. Mængden af løsninger til det homogene system af lineære ligninger

$$\begin{cases} x_1 & + & 2 \cdot x_2 & + & & & 3 \cdot x_4 & = & 0 \\ & & & & x_3 & + & 4 \cdot x_4 & = & 0 \\ & & & & & & 0 & = & 0 \end{cases}$$

er præcist givet ved

$$\left\{ t_1 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \;\middle|\; t_1, t_2 \in \mathbb{R} \right\}.$$

I dette eksempel startede vi med et homogent system af lineære ligninger, hvis koefficientmatrix var på reduceret trappeform. Dette var grunden til, at vi relativt hurtigt kunne bestemme alle løsninger. Fra de tidligere afsnit ved vi dog, at selv hvis vi starter med et mere kompliceret system, kan vi altid benytte elementære rækkeoperationer til at transformere det på en sådan måde, at den resulterende koefficientmatrix er på reduceret trappeform. Grundlæggende beskriver Eksempel 6.4.3, hvordan man bestemmer alle løsninger, når koefficientmatricen for systemet af lineære ligninger er på reduceret trappeform. Samme idéer gælder for ethvert homogent system af lineære ligninger: reducér først systemet ved at bringe dets koefficientmatrix på reduceret trappeform, og følg derefter proceduren eksemplificeret i Eksempel 6.4.3. Det er muligt at beskrive resultatet i det generelle tilfælde, og for fuldstændighedens skyld gør vi det i følgende sætning. Når man bliver bedt om at løse et homogent system af lineære ligninger i praksis, er det dog ofte nemmere direkte at anvende en procedure som den i Eksempel 6.4.3 fremfor at benytte denne sætning.

Sætning 6.4.4

Lad et homogent system af m lineære ligninger i n variable over et legeme $\mathbb{F}$ være givet. Lad koefficientmatricen for dette system være $\mathbf{A}$, og lad $\hat{\mathbf{A}}$ betegne $\mathbf{A}$'s reducerede trappeform. Antag videre, at $\hat{\mathbf{A}}$ har ρ pivoter i søjlerne $j_1, \dots, j_\rho$, og betegn ved

$$\mathbf{c}_1 = \begin{bmatrix} c_{11} \\ \vdots \\ c_{m1} \end{bmatrix}, \dots, \mathbf{c}_{n-\rho} = \begin{bmatrix} c_{1\,n-\rho} \\ \vdots \\ c_{m\,n-\rho} \end{bmatrix}$$

de $n-\rho$ søjler i $\hat{\mathbf{A}}$, der ikke indeholder en pivot. Definér afslutningsvis

$$\mathbf{v}_1 = \begin{bmatrix} v_{11} \\ \vdots \\ v_{n1} \end{bmatrix}, \dots, \mathbf{v}_{n-\rho} = \begin{bmatrix} v_{1\,n-\rho} \\ \vdots \\ v_{n\,n-\rho} \end{bmatrix}$$

ved

$$v_{ji} = \begin{cases} -c_{\ell i} & \text{hvis } j = j_\ell \text{ for nogle } \ell = 1, \dots, \rho, \\ 1 & \text{hvis } \mathbf{c}_i \text{ er den } j\text{'te søjle i } \hat{\mathbf{A}}, \\ 0 & \text{ellers.} \end{cases}$$

Da er løsningsmængden til det givne homogene system af lineære ligninger givet ved

$$\left\{ t_1 \cdot \begin{bmatrix} v_{11} \\ \vdots \\ v_{n1} \end{bmatrix} + \cdots + t_{n-\rho} \cdot \begin{bmatrix} v_{1\,n-\rho} \\ \vdots \\ v_{n\,n-\rho} \end{bmatrix} \mid t_1, \dots, t_{n-\rho} \in \mathbb{F} \right\}.$$

Bevis. Vi vil ikke bevise denne sætning men blot angive idéen bag beviset. Først anvendes Sætning 6.2.1 til at konkludere, at det homogene system med koefficientmatrix $\mathbf{A}$ har præcis de samme løsninger som det homogene system med koefficientmatrix $\hat{\mathbf{A}}$. Derefter anvendes den samme metode som i Eksempel 6.4.3 til at beskrive alle løsninger til det homogene system med koefficientmatrix $\hat{\mathbf{A}}$. □

Udtrykket

$$t_1 \cdot \begin{bmatrix} v_{11} \\ \vdots \\ v_{n1} \end{bmatrix} + \cdots + t_{n-\rho} \cdot \begin{bmatrix} v_{1\,n-\rho} \\ \vdots \\ v_{n\,n-\rho} \end{bmatrix} \quad (t_1, \ldots, t_{n-\rho} \in \mathbb{F})$$

kaldes *den fuldstændige løsning* til det homogene system med koefficientmatrix $\mathbf{A}$. Når vi ser tilbage på Eksempel 6.4.3, ser vi, at den fuldstændige løsning til det homogene system af lineære ligninger, der blev studeret i det eksempel, blev bestemt til at være lig med

$$t_1 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \quad (t_1, t_2 \in \mathbb{R}).$$

Korollar 6.4.5

Lad et homogent system af m lineære ligninger i n variable over et legeme $\mathbb{F}$ være givet. Betegn koefficientmatricen for dette system ved $\mathbf{A}$. Da har det homogene system ikke andre løsninger end nulløsningen $(0, \ldots, 0) \in \mathbb{F}^n$, hvis og kun hvis $\rho(\mathbf{A}) = n$.

Bevis. Sætning 6.4.4 medfører, at hvis rangen af $\mathbf{A}$ er mindre end n, så findes der en løsning, der ikke er nulløsningen. Omvendt, hvis rangen af $\mathbf{A}$ er lig med n, er antallet af parametre t_i i beskrivelsen af løsningsmængden i Sætning 6.4.4 nul. Det betyder, at kun nulløsningen $(0, \ldots, 0)$ er en løsning. □

Status er nu, at vi kan bestemme alle løsninger til ethvert homogent system af lineære ligninger (kaldet den fuldstændige løsning til det homogene system), kan bestemme, om et inhomogent system har en løsning, og i så fald kan finde en sådan løsning (kaldet en partikulær løsning). Ved at anvende Sætning 6.1.2 kan vi i dette tilfælde også bestemme en formel, der beskriver alle løsninger til et inhomogent system af lineære ligninger: det er simpelthen summen af en partikulær løsning og den fuldstændige løsning til det tilsvarende homogene system. Denne sum kaldes *den fuldstændige løsning* til det inhomogene system. Derfor har vi på en konstruktiv måde besvaret alle tre spørgsmål, der blev stillet i slutningen af Afsnit 6.1.

Lad os afslutte dette afsnit med et eksempel, hvor vi bestemmer den fuldstændige løsning til et inhomogent system af lineære ligninger.

Eksempel 6.4.4

Lad os vende tilbage til det inhomogene system af lineære ligninger, der blev behandlet i Eksempel 6.4.2:

$$\begin{cases} x_1 & + & 2 \cdot x_2 & + & & & 3 \cdot x_4 & = & 5 \\ & & & & x_3 & + & 4 \cdot x_4 & = & 6 \\ & & & & & & 0 & = & 0 \end{cases} .$$

Vi har bestemt en partikulær løsning i Eksempel 6.4.2 og den fuldstændige løsning til det tilsvarende homogene system i Eksempel 6.4.3. Ved at anvende disse tidligere resultater i kombination med Sætning 6.1.2, konkluderer vi, at den fuldstændige løsning til det inhomogene system er givet ved:

$$\begin{bmatrix} 5 \\ 0 \\ 6 \\ 0 \end{bmatrix} + t_1 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \quad (t_1, t_2 \in \mathbb{R}).$$

Løsningsmængden til det inhomogene system er derfor:

$$\left\{ \begin{bmatrix} 5 \\ 0 \\ 6 \\ 0 \end{bmatrix} + t_1 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \mid t_1, t_2 \in \mathbb{R} \right\}.$$

6.5 Entydighed af den reducerede trappeform

Tidligere har vi udspecificeret, at en given matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ har en entydig reduceret trappeform. Eksistensen af den blev vist i Sætning 6.3.1, og i dette afsnit ønsker vi at vise entydigheden. Afsnittet kan springes over og er kun tiltænkt læsere, der ønsker at se et bevis for entydigheden af den reducerede trappeform.

Sætning 6.5.1

Lad $\mathbb{F}$ være et legeme og $\mathbf{A} \in \mathbb{F}^{m \times n}$ en matrix. Antag, at $\mathbf{A}$ kan transformeres ved hjælp af en sekvens af elementære rækkeoperationer til en matrix $\mathbf{B}_1$ på reduceret trappeform og ved hjælp af en anden sekvens af elementære rækkeoperationer til en matrix $\mathbf{B}_2$ på reduceret trappeform. Da gælder der, at $\mathbf{B}_1 = \mathbf{B}_2$.

Bevis. Fra Sætning 6.2.1 ved vi, at de homogene systemer af lineære ligninger med koefficientmatricerne $\mathbf{A}$, $\mathbf{B}_1$ og $\mathbf{B}_2$ alle har præcis samme løsning. Ideen med beviset er at vise, at de homogene systemer af lineære ligninger med koefficientmatricerne $\mathbf{B}_1$ og $\mathbf{B}_2$ kun kan have de samme løsninger, hvis $\mathbf{B}_1 = \mathbf{B}_2$. Desuden anvender vi induktion på n, antallet af søjler.

Lad os starte med induktionens basistrin. Hvis $n = 1$, er der kun to mulige reducerede trappeformer: $m \times 1$-matricerne

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{og} \quad \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Den første kan kun være en reduceret trappeform af $\mathbf{A}$, hvis $\mathbf{A}$ oprindeligt var en $m \times 1$-nulmatrix. Anvendelse af en hvilken som helst elementær rækkeoperation på nulmatricen resulterer i nulmatricen igen. Hvis $\mathbf{B}_1$ eller $\mathbf{B}_2$ er nulmatricen, har vi derfor $\mathbf{A} = \mathbf{B}_1 = \mathbf{B}_2$, da de i så fald alle er lig med nulmatricen. Antag nu, at $\mathbf{B}_1$ eller $\mathbf{B}_2$ er lig med den anden mulige $m \times 1$-matrix på reduceret trappeform. Hvis $\mathbf{B}_1 \neq \mathbf{B}_2$, så er mindst en af dem lig med den eneste anden $m \times 1$-matrix på reduceret trappeform, nemlig nulmatricen. Men vi har lige set, at dette ville betyde, at både $\mathbf{B}_1$ og $\mathbf{B}_2$ er lig med nulmatricen. Denne modstrid viser, at hvis $\mathbf{B}_1$ eller $\mathbf{B}_2$ er lig med den anden $m \times 1$-matrix på reduceret trappeform, så er $\mathbf{B}_1 = \mathbf{B}_2$.

Vi fortsætter med induktionstrinnet. Antag $n > 1$, samt at sætningen er sand for $n - 1$. Lad os for enhver $m \times n$-matrix $\mathbf{A}$ betegne ved $\mathbf{A}|_{n-1}$ den $m \times (n-1)$-matrix, man opnår ved at fjerne den sidste søjle fra $\mathbf{A}$. Induktionshypotesen betyder, at $\mathbf{A}|_{n-1}$ har en unik reduceret trappeform. Desuden, hvis $\mathbf{B}$ er en $m \times n$-matrix på reduceret trappeform, så er matricen $\mathbf{B}|_{n-1}$ også på reduceret trappeform. Dette medfører, at hvis $\mathbf{B}_1$ og $\mathbf{B}_2$ er to mulige reducerede trappeformer af $\mathbf{A}$, så betyder induktionshypotesen, at $\mathbf{B}_1|_{n-1} = \mathbf{B}_2|_{n-1}$. Med andre ord: de første $n - 1$ søjler i $\mathbf{B}_1$ og $\mathbf{B}_2$ er identiske. Kun de n'te (altså de sidste) søjler i matricerne kan være forskellige. Lad nu ρ betegne antallet af pivoter, der forekommer i $\mathbf{B}_1|_{n-1}$. Hvis den n'te søjle i $\mathbf{B}_1$ indeholder en pivot, indeholder denne søjle kun nuller undtagen i den $(\rho+1)$'te position, hvor den indeholder et ét-tal. Derfor opfylder enhver løsning $(v_1, \ldots, v_n) \in \mathbb{F}^n$ til det homogene system af lineære ligninger med koefficientmatrix $\mathbf{B}_1$, at $v_n = 0$. Omvendt jævnfør Sætning 6.4.4, hvis den n'te søjle i $\mathbf{B}_1$ ikke indeholder en pivot, findes der en løsning $(v_1, \ldots, v_n)$, således at $v_n = 1$. Et lignende ræsonnement gælder for den sidste søjle i $\mathbf{B}_2$. Ved at benytte Sætning 6.2.1 kan vi konkludere, at de homogene systemer af lineære ligninger med koefficientmatricerne $\mathbf{B}_1$, $\mathbf{B}_2$ og $\mathbf{A}$ alle har præcis de samme løsningmængder. Det følger, at enten forekommer der en pivot i den n'te søjle i både $\mathbf{B}_1$ og $\mathbf{B}_2$, eller der forekommer ikke nogen pivot i den n'te søjle i både $\mathbf{B}_1$ og $\mathbf{B}_2$. I det første tilfælde har vi allerede set, at den n'te søjle er fuldstændig bestemt, hvilket medfører, at $\mathbf{B}_1 = \mathbf{B}_2$. I det andet tilfælde kan vi konkludere, at der er præcis én løsning til det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A}$, som har udelukkende nuller i søjlerne, der ikke indeholder pivoter, pånær i den n'te søjle, hvor den har et ét-tal. Ved brug af Sætning 6.4.4 ser man, at koefficienterne for denne løsning fuldstændigt bestemmer den n'te søjle i $\mathbf{A}$'s reducerede trappeform. Vi konkluderer, at $\mathbf{B}_1 = \mathbf{B}_2$ også gælder i det andet tilfælde, hvor der ikke forekommer nogen pivot i den n'te søjle i både $\mathbf{B}_1$ og $\mathbf{B}_2$. $\square$

Kapitel 7

Vektorer og matricer

7.1 Vektorer

Som i det forrige kapitel vil vi betegne et tallegeme ved $\mathbb{F}$. Det, vi vil arbejde med nu, virker over ethvert legeme, men læseren kan blot tænke på $\mathbb{F} = \mathbb{R}$ eller $\mathbb{F} = \mathbb{C}$. Når vi beskriver løsninger til systemer af lineære ligninger, arbejder vi allerede med $\mathbb{F}^n$, mængden af alle n-tupler med værdier i $\mathbb{F}$. Vi har også allerede forklaret, at et sådant n-tupel for nemheds skyld ofte betragtes som en $n \times 1$-matrix. Det betyder blot, at:

$$(v_1, \dots, v_n) \text{ også kan skrives som } \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}.$$

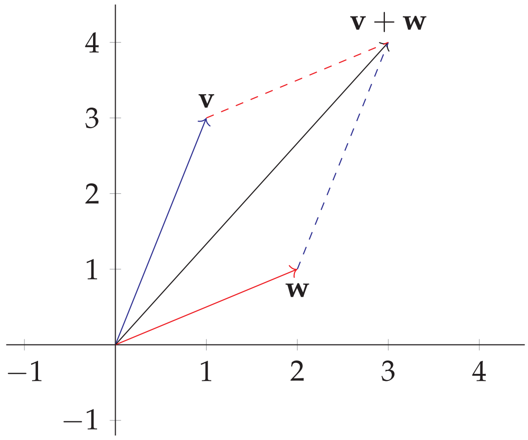
Når et n-tupel skrives som en $n \times 1$-matrix, siger vi, at n-tuplet er skrevet på *vektor*-form. Elementer i $\mathbb{F}^n$ kaldes derfor *vektorer* med n elementer fra $\mathbb{F}$. Hvis alle elementer i en sådan vektor er nul, kalder vi denne vektor *nulvektoren* i $\mathbb{F}^n$.

Bemærkning 7.1.1
Elementer i $\mathbb{F}^{n \times 1}$ kaldes typisk *søjlevektorer*, mens elementer i $\mathbb{F}^{1 \times n}$ kaldes *rækkevektorer*.

Vi har i det forrige kapitel allerede benyttet os af, at der findes en naturlig måde at lægge to vektorer $\mathbf{v}$ og $\mathbf{w}$ fra $\mathbb{F}^n$ sammen på, og at man også kan multiplicere en vektor fra $\mathbb{F}^n$ med et element $c \in \mathbb{F}$, ofte kaldet en *skalar* i denne sammenhæng, da multiplikation af en vektor med en konstant kan betragtes som en skalering af vektoren. Mere præcist er addition af vektorer i $\mathbb{F}^n$ defineret som:

$$\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{bmatrix}. \tag{7.1}$$

For $\mathbb{F} = \mathbb{R}$ og $n = 2$ er addition af to vektorer illustreret i Figur 7.1.

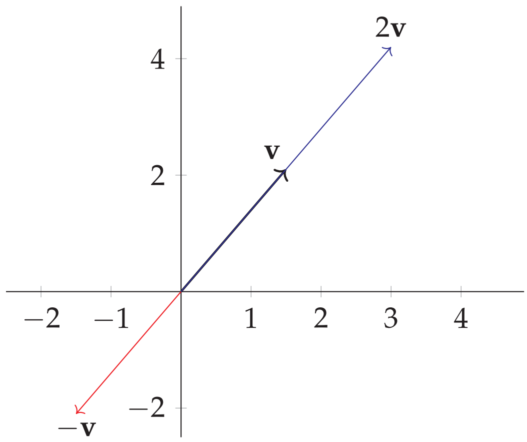


Figur 7.1: Addition af to vektorer $\mathbf{v}$ og $\mathbf{w}$ i $\mathbb{R}^2$.

Produktet af en skalar c fra $\mathbb{F}$ med en vektor i $\mathbb{F}^n$ defineres som:

$$c \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} c \cdot v_1 \\ \vdots \\ c \cdot v_n \end{bmatrix}. \tag{7.2}$$

I stedet for $(-1) \cdot \mathbf{v}$ kan man også skrive $-\mathbf{v}$. Ligeledes skrives et udtryk som $\mathbf{v} + (-1) \cdot \mathbf{w}$ ofte som $\mathbf{v} - \mathbf{w}$. Man udelader også normalt multiplikationstegnet mellem skalar og vektor. Med andre ord, et udtryk som $c\mathbf{v}$ skal blot læses som $c \cdot \mathbf{v}$. For $\mathbb{F} = \mathbb{R}$ og $n = 2$ er skalering af en vektor illustreret i Figur 7.2.



Figur 7.2: Skalering af en vektor $\mathbf{v} \in \mathbb{R}^2$.

Som med matricer vil vi ofte benytte fed skrift til vektorer og typisk bogstaver som $\mathbf{u}, \mathbf{v}, \mathbf{w}$. Lad os for fremtidig reference opstille følgende samling af egenskaber for vektoraddition og skalarmultiplikation:

Sætning 7.1.1

Lad $\mathbb{F}$ være et legeme, og lad $c, d \in \mathbb{F}$ samt $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{F}^n$. Da gælder:

(i) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$

(ii) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

(iii) $c \cdot (d \cdot \mathbf{u}) = (c \cdot d) \cdot \mathbf{u}$

(iv) $c \cdot (\mathbf{u} + \mathbf{v}) = c \cdot \mathbf{u} + c \cdot \mathbf{v}$

(v) $(c + d) \cdot \mathbf{u} = c \cdot \mathbf{u} + d \cdot \mathbf{u}$

Vi udelader beviset herfor.

Da vi nu har vektorer til rådighed, kan vi se nærmere på flere af deres egenskaber. Vi starter med et eksempel.

Eksempel 7.1.1

Vi er givet vektorerne

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ og } \mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \in \mathbb{R}^2.$$

(a) Udregn $4 \cdot \mathbf{u} + 3 \cdot \mathbf{v}$.

(b) Find c og d, således at $c \cdot \mathbf{u} + d \cdot \mathbf{v} = \mathbf{0}$, hvor $\mathbf{0}$ betegner nulvektoren i $\mathbb{R}^2$.

Svar:

(a) Ved brug af definitionen af skalarmultiplikation og vektoraddition får vi

$$4 \cdot \mathbf{u} + 3 \cdot \mathbf{v} = 4 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3 \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix} + \begin{bmatrix} 6 \\ 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 11 \end{bmatrix}.$$

(b) Vi har

$$c \cdot \mathbf{u} + d \cdot \mathbf{v} = c \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} + d \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} c \\ 2c \end{bmatrix} + \begin{bmatrix} 2d \\ d \end{bmatrix} = \begin{bmatrix} c + 2d \\ 2c + d \end{bmatrix}.$$

Hvis vi ønsker, at resultatet skal være nulvektoren, skal vi løse det homogene system af lineære ligninger:

$$\begin{cases} c + 2d = 0 \\ 2c + d = 0 \end{cases}.$$

Ved at trække den første ligning to gange fra den anden ligning, hvilket svarer til at udføre den elementære rækkeoperation $R_2 \leftarrow R_2 - 2 \cdot R_1$, får vi systemet:

$$\begin{cases} c + 2d = 0 \\ 0c - 3d = 0 \end{cases}.$$

Vi kunne fortsætte og bringe systemet på reduceret trappeform, men det er allerede klart nu, at den eneste løsning er $c = d = 0$.

Et udtryk som $4 \cdot \mathbf{u} + 3 \cdot \mathbf{v}$ kaldes en *linearkombination* af vektorerne $\mathbf{u}$ og $\mathbf{v}$. Mere generelt, givet vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ og skalarer $c_1, \ldots, c_n \in \mathbb{F}$ er et udtryk på formen

$$c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n$$

en linearkombination af vektorerne $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Anden del af eksemplet viser, at den eneste linearkombination af vektorerne $\mathbf{u}$ og $\mathbf{v}$, der er lig med nulvektoren, er linearkombinationen $0 \cdot \mathbf{u} + 0 \cdot \mathbf{v}$. En sekvens af vektorer kan have denne egenskab generelt, hvilket beskrives i følgende.

Definition 7.1.1

En sekvens af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ kaldes *lineært uafhængig*, hvis og kun hvis ligningen $c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n = \mathbf{0}$ kun kan være sand for $c_1 = \cdots = c_n = 0$.

Hvis sekvensen af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ ikke er lineært uafhængig, siges den at være *lineært afhængig*.

Med andre ord er en sekvens af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ lineært uafhængig, hvis og kun hvis den eneste linearkombination af vektorerne, der er lig med nulvektoren, opnås for $c_1 = \cdots = c_n = 0$. Ved hjælp af nogle logiske udtryk kan lineær uafhængighed af en sekvens af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ formuleres som følger:

$$\text{for alle } c_1, \ldots, c_n \in \mathbb{F} \text{ gælder: } c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n = \mathbf{0} \Rightarrow c_1 = \cdots = c_n = 0. \tag{7.3}$$

På samme måde kan lineær afhængighed af sekvensen af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ formuleres på følgende måde:

$$\text{der eksisterer } c_1, \ldots, c_n \in \mathbb{F}, \text{ således at: } c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n = \mathbf{0} \;\wedge\; \text{ ikke alle } c_i \text{ er nul.} \tag{7.4}$$

I stedet for at sige, at en sekvens af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært (u)afhængig, er det også ganske almindeligt blot at sige, at vektorerne $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært (u)afhængige. Vi vil ofte bruge denne måde at formulere tingene på.

Eksempel 7.1.2

Sekvensen af vektorer bestående af

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ og } \mathbf{v} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \in \mathbb{R}^2$$

er lineært afhængig. Da $\mathbf{v} = 2 \cdot \mathbf{u}$, ser vi nemlig, at $(-2) \cdot \mathbf{u} + \mathbf{v} = \mathbf{0}$.

Dette eksempel illustrerer et mere generelt princip: To vektorer $\mathbf{u}$ og $\mathbf{v}$ er lineært afhængige, hvis og kun hvis den ene er en skalering af den anden. Hvis for eksempel $\mathbf{u} = c \cdot \mathbf{v}$, så er $1 \cdot \mathbf{u} + (-c) \cdot \mathbf{v} = \mathbf{0}$, hvilket viser, at vektorerne er lineært afhængige. På samme måde, hvis

$\mathbf{v} = c \cdot \mathbf{u}$, så er $(-c) \cdot \mathbf{u} + 1 \cdot \mathbf{v} = \mathbf{0}$, hvilket igen viser, at vektorerne er lineært afhængige. Omvendt, hvis vektorerne er lineært afhængige, findes der $c, d \in \mathbb{F}$, som ikke begge er nul, således at $c \cdot \mathbf{u} + d \cdot \mathbf{v} = \mathbf{0}$. Hvis $c \neq 0$, så får vi $\mathbf{u} = (-d/c) \cdot \mathbf{v}$, således at $\mathbf{v}$ er en skalering af $\mathbf{u}$. Hvis $d \neq 0$, opnår vi på samme måde, at $\mathbf{v} = (-c/d) \cdot \mathbf{u}$, hvilket viser at $\mathbf{u}$ i dette tilfælde er en skalering af $\mathbf{v}$. Intuitivt kan man derfor sige, at to vektorer $\mathbf{u}$ og $\mathbf{v}$ er lineært afhængige, hvis og kun hvis der findes en linje gennem origo, som indeholder både $\mathbf{u}$ og $\mathbf{v}$.

Eksempel 7.1.3

Sekvensen af vektorer bestående af

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ og } \mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \in \mathbb{R}^2$$

er lineært uafhængig. Vi har nemlig set i Eksempel 7.1.1, at ligningen $c \cdot \mathbf{u} + d \cdot \mathbf{v} = \mathbf{0}$ medfører, at $c = d = 0$.

Dette eksempel illustrerer, at den lineære uafhængighed af en sekvens af vektorer kan undersøges ved brug af teorien om systemer af lineære ligninger. Dette er faktisk tilfældet generelt, og det generelle resultat er følgende:

Lemma 7.1.2

Lad vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ være givet, og lad $\mathbf{A} \in \mathbb{F}^{m \times n}$ være den $m \times n$-matrix, hvis søjler er $\mathbf{v}_1, \ldots, \mathbf{v}_n$, det vil sige

$$\mathbf{A} = \begin{bmatrix} | & & | \\ \mathbf{v}_1 & \ldots & \mathbf{v}_n \\ | & & | \end{bmatrix}.$$

Sekvensen af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært uafhængig, hvis og kun hvis det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A}$ kun har nulvektoren $\mathbf{0} \in \mathbb{F}^n$ som løsning.

Bevis. Antag først, at sekvensen af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært uafhængig, og lad $(c_1, \ldots, c_n) \in \mathbb{F}^n$ være en løsning til det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A}$. Dette system kan direkte omskrives til ligningen $c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n = \mathbf{0}$. Ved at bruge vores antagelse om, at sekvensen af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært uafhængig, ser vi, at $(c_1, \ldots, c_n) = (0, \ldots, 0)$.

Antag nu omvendt, at det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A}$ kun har nulvektoren $\mathbf{0} \in \mathbb{F}^n$ som løsning. Hvis $(c_1, \ldots, c_n) \in \mathbb{F}^n$ opfylder $c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n = \mathbf{0}$, kan vi straks konkludere, at $(c_1, \ldots, c_n)$ også er en løsning til det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A}$. Men så kan vi konkludere, at $(c_1, \ldots, c_n) = (0, \ldots, 0)$. $\square$

Dette lemma fører til en enkel karakterisering af lineær uafhængighed:

Sætning 7.1.3
Lad $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ være givet, og lad $\mathbf{A} \in \mathbb{F}^{m \times n}$ være matricen med søjlerne $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Sekvensen af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært uafhængig, hvis og kun hvis matricen $\mathbf{A}$ har rang n.

Bevis. Dette følger af Korollar 6.4.5 og Lemma 7.1.2. □

Eksempel 7.1.4
Betragt de følgende tre vektorer i $\mathbb{C}^3$:

$$\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 1+i \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 0 \\ 1+i \\ 0 \end{bmatrix} \text{ og } \mathbf{w} = \begin{bmatrix} 1+i \\ -1+5i \\ 2i \end{bmatrix}.$$

(a) Er vektorerne $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ lineært uafhængige?

(b) Er vektorerne $\mathbf{u}$, $\mathbf{v}$ lineært uafhængige?

(c) Er vektoren $\mathbf{u}$ lineært uafhængig?

Svar: Den generelle strategi for denne type spørgsmål er at bruge Sætning 7.1.3. Husk, at for at beregne rangen af en matrix er det ifølge Definition 6.3.2, definitionen på rangen af en matrix, tilstrækkeligt at bestemme dens reducerede trappeform. Lad os nu besvare de tre spørgsmål ét ad gangen.

(a) Sætning 7.1.3 medfører, at vi for at finde svaret skal bestemme rangen af matricen

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1+i \\ 0 & 1+i & -1+5i \\ 1+i & 0 & 2i \end{bmatrix}.$$

Vi har

$$\begin{bmatrix} 1 & 0 & 1+i \\ 0 & 1+i & -1+5i \\ 1+i & 0 & 2i \end{bmatrix} \underset{R_3 \leftarrow R_3 - (1+i) \cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1+i \\ 0 & 1+i & -1+5i \\ 0 & 0 & 0 \end{bmatrix}$$
$$\underset{R_2 \leftarrow (1+i)^{-1} \cdot R_2}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1+i \\ 0 & 1 & 2+3i \\ 0 & 0 & 0 \end{bmatrix}.$$

Vi kan konkludere, at $\rho(\mathbf{A}) = 2$, hvilket er mindre end tre, altså antallet af vektorer vi arbejder med. Derfor er vektorerne $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ lineært afhængige.

(b) I dette tilfælde skal vi beregne rangen af matricen

$$\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1+i \\ 1+i & 0 \end{bmatrix}.$$

Anvendes de samme elementære rækkeoperationer som ved løsning af første spørgsmål, får vi, at den reducerede trappeform af **B** er matricen

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Vi ser, at $\rho(\mathbf{B}) = 2$, hvilket er lig med antallet af vektorer, vi arbejder med. Derfor er vektorerne $\mathbf{u}$, $\mathbf{v}$ lineært uafhængige.

(c) Hvis vi kun overvejer vektoren $\mathbf{u}$, skal vi bestemme rangen af matricen

$$\mathbf{C} = \begin{bmatrix} 1 \\ 0 \\ 1+i \end{bmatrix}.$$

Denne matrix har rang ét, da den ene søjle, som denne matrix består af, ikke er en nulsøjle. Vi kan konkludere, at sekvensen bestående af vektoren $\mathbf{u}$ er lineært uafhængig. Generelt er en sekvens bestående af kun én vektor $\mathbf{u} \in \mathbb{F}^m$ lineært uafhængig, hvis og kun hvis $\mathbf{u} \neq \mathbf{0}$.

7.2 Matricer og vektorer

Da vi studerede systemer af lineære ligninger, introducerede vi begrebet matrix. En matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ blev introduceret som et rektangulært skema indeholdende $m \times n$ elementer fra et givet legeme $\mathbb{F}$:

$$\mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}.$$

Nogle gange skriver man blot $\mathbf{A} = [a_{ij}]_{1 \leq i \leq m, 1 \leq j \leq m}$ for at holde det kort. Når en matrix gives på denne form, er elementet a_{ij}, som nogle gange også skrives som $a_{i,j}$, det element, der befinder sig i række i og søjle j i matricen $\mathbf{A}$. Det er også normal notation at betegne dette element ved $\mathbf{A}_{ij}$ eller $\mathbf{A}_{i,j}$. Matricen $\mathbf{A}$, der er givet ovenfor, har m rækker: $[a_{i1} \ \dots \ a_{in}]$ for

$i = 1, \ldots, m$ og n søjler:

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \text{ for } j = 1, \ldots, n.$$

Vi vil kalde rækker i en matrix for *rækkevektorer* og på tilsvarende vis søjler i en matrix for *søjlevektorer*.

Det viser sig at være ganske nyttigt at kunne multiplicere en matrix med en vektor. Vi definerer følgende:

Definition 7.2.1

Lad $\mathbf{A} = (a_{ij})_{1 \le i \le m, 1 \le j \le m} \in \mathbb{F}^{m \times n}$ være en matrix og $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{F}^n$ en vektor. Da definerer vi $\mathbf{A} \cdot \mathbf{v} \in \mathbb{F}^m$ som følger:

$$\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} a_{11} \cdot v_1 + \cdots a_{1n} \cdot v_n \\ \vdots \\ a_{m1} \cdot v_1 + \cdots a_{mn} \cdot v_n \end{bmatrix}$$

Bemærk, at vi ikke kan multiplicere en hvilken som helst matrix med en hvilken som helst vektor. Deres størrelser skal "passe" sammen: antallet af søjler i matricen skal være lige med antallet af elementer i vektoren. Hvis dette ikke er tilfældet, er den tilsvarende matrix-vektor-multiplikation ikke defineret.

Eksempel 7.2.1

Lad

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \text{ og } \mathbf{v} = \begin{bmatrix} 7 \\ -1 \\ -2 \end{bmatrix}.$$

Udregn $\mathbf{A} \cdot \mathbf{v}$.

Svar: Ved at benytte Definition 7.2.1 får vi, at:

$$\mathbf{A} \cdot \mathbf{v} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 7 \\ -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \cdot 7 + 2 \cdot (-1) + 3 \cdot (-2) \\ 4 \cdot 7 + 5 \cdot (-1) + 6 \cdot (-2) \end{bmatrix} = \begin{bmatrix} -1 \\ 11 \end{bmatrix}.$$

Bemærk, at matrix-vektor-produktet opstår meget naturligt, når man betragter et system af lineære ligninger. Et system af lineære ligninger

$$\begin{cases} a_{11} \cdot x_1 & + & \cdots & + & a_{1n} \cdot x_n & = & b_1 \\ & & \vdots & & & & \vdots \\ a_{m1} \cdot x_1 & + & \cdots & + & a_{mn} \cdot x_n & = & b_m \end{cases}$$

kan udtrykkes som

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}. \tag{7.5}$$

Nu hvor vi har defineret et matrix-vektor-produkt, kan man stille spørgsmålet, om man mere generelt også kan multiplicere matricer med hinanden. Svaret viser sig at være ja, forudsat at deres størrelser passer sammen. Mere præcist kan vi gøre følgende:

Definition 7.2.2

Lad $\mathbf{A} \in \mathbb{F}^{m\times n}$ og $\mathbf{B} \in \mathbb{F}^{n\times \ell}$. Antag, at søjlerne i $\mathbf{B}$ er givet ved $\mathbf{b}_1, \ldots, \mathbf{b}_\ell \in \mathbb{F}^n$, altså antag, at

$$\mathbf{B} = \begin{bmatrix} | & & | \\ \mathbf{b}_1 & \ldots & \mathbf{b}_\ell \\ | & & | \end{bmatrix}.$$

Så definerer vi

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} | & & | \\ \mathbf{A} \cdot \mathbf{b}_1 & \ldots & \mathbf{A} \cdot \mathbf{b}_\ell \\ | & & | \end{bmatrix}.$$

Bemærk, at matrixproduktet $\mathbf{A} \cdot \mathbf{B}$ kun er defineret, hvis antallet af søjler i $\mathbf{A}$ er lig med antallet af rækker i $\mathbf{B}$. Hvis disse tal stemmer overens, er $\mathbf{A} \cdot \mathbf{B}$ en matrix med m rækker og ℓ søjler. Med andre ord, hvis $\mathbf{A} \in \mathbb{F}^{m\times n}$, og $\mathbf{B} \in \mathbb{F}^{n\times \ell}$, så er $\mathbf{A} \cdot \mathbf{B} \in \mathbb{F}^{m\times \ell}$.

En anden måde at tilgå definitionen af matrixproduktet på er at opstille en formel for elementerne i produktet $\mathbf{A} \cdot \mathbf{B}$ én ad gangen. Lad os sige, at vi ønsker at finde en formel for det (i, j)'te element i produktet, $(\mathbf{A} \cdot \mathbf{B})_{i,j}$, altså elementet i række i og søjle j. Dette svarer til at bestemme det i'te element i produktet $\mathbf{A} \cdot \mathbf{b}_j$, hvor $\mathbf{b}_j$ er den j'te søjle i $\mathbf{B}$. Dette er i sig selv præcis det samme som resultatet af at gange den i'te række i matricen $\mathbf{A}$ med den j'te søjle i matricen $\mathbf{B}$. Da den i'te række i $\mathbf{A}$ kan skrives som $[a_{i1} \ \ldots \ a_{in}]$ og den j'te søjle i $\mathbf{B}$ som

$$\mathbf{b}_j = \begin{bmatrix} b_{1j} \\ \vdots \\ b_{mj} \end{bmatrix},$$

ser vi, at

$$(\mathbf{A} \cdot \mathbf{B})_{i,j} = [a_{i1} \ \ldots \ a_{in}] \cdot \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix} = a_{i1} \cdot b_{1j} + \cdots + a_{in} \cdot b_{nj}.$$

Ved at bruge summationssymbolet fra Afsnit 5.3 kan vi omskrive denne formel som følger:

$$(\mathbf{A} \cdot \mathbf{B})_{i,j} = \sum_{r=1}^{n} a_{ir} \cdot b_{rj}. \tag{7.6}$$

Eksempel 7.2.2

I dette eksempel sætter vi $\mathbb{F} = \mathbb{R}$, og vi har givet

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \text{ og } \mathbf{B} = \begin{bmatrix} 7 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}.$$

Udregn, hvis det er muligt, matrixprodukterne $\mathbf{A} \cdot \mathbf{B}$ og $\mathbf{B} \cdot \mathbf{A}$.

Svar: Først betragtes matrixproduktet $\mathbf{A} \cdot \mathbf{B}$. Da $\mathbf{A} \in \mathbb{R}^{2\times 3}$, og $\mathbf{B} \in \mathbb{R}^{3\times 3}$, er produktet $\mathbf{A} \cdot \mathbf{B}$ defineret. Vi har allerede udregnet produktet af $\mathbf{A}$ med den første søjle i $\mathbf{B}$ i Eksempel 7.2.1, så vi vil ikke gentage de beregninger her. Tages det i betragtning, får vi:

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 7 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} &= \begin{bmatrix} -1 & 1\cdot 0 + 2 \cdot 1 + 3 \cdot 0 & 1 \cdot 0 + 2 \cdot 0 + 3 \cdot 1 \\ 11 & 4 \cdot 0 + 5 \cdot 1 + 6 \cdot 0 & 4 \cdot 0 + 5 \cdot 0 + 6 \cdot 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 2 & 3 \\ 11 & 5 & 6 \end{bmatrix}. \end{aligned}$$

Nu ser vi på matrixproduktet $\mathbf{B} \cdot \mathbf{A}$. Da $\mathbf{B}$ har tre søjler, og $\mathbf{A}$ har to rækker, er matrixproduktet $\mathbf{B} \cdot \mathbf{A}$ ikke defineret.

Dette eksempel viser, at der gælder generelt, at $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$. Med andre ord er matrixmultiplikation ikke kommutativ. Faktisk kan det endda ske, som vi lige har set, at et af produkterne ikke er defineret. Selv hvis begge produkter er defineret, betyder rækkefølgen af matricerne stadig noget, og $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$ generelt. Tag for eksempel et kig på $\mathbf{A} = [1\, 0]$ og $\mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Så

$$\mathbf{A} \cdot \mathbf{B} = [1\,0] \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1 \cdot 0 + 0 \cdot 1 = 0, \text{ og } \mathbf{B} \cdot \mathbf{A} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cdot [1\,0] = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Lad os også definere addition af matricer:

Definition 7.2.3

Lad $\mathbf{A}, \mathbf{A}' \in \mathbb{F}^{m\times n}$ være givet, altså

$$\mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \text{ og } \mathbf{A}' = \begin{bmatrix} a'_{11} & \dots & a'_{1n} \\ \vdots & & \vdots \\ a'_{m1} & \dots & a'_{mn} \end{bmatrix}.$$

Så definerer vi $\mathbf{A}+\mathbf{A}'$ som følger:

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} + \begin{bmatrix} a'_{11} & \dots & a'_{1n} \\ \vdots & & \vdots \\ a'_{m1} & \dots & a'_{mn} \end{bmatrix} = \begin{bmatrix} a_{11}+a'_{11} & \dots & a_{1n}+a'_{1n} \\ \vdots & & \vdots \\ a_{m1}+a'_{m1} & \dots & a_{mn}+a'_{mn} \end{bmatrix}.$$

Addition af matricer er kun defineret, hvis de har samme størrelse. På elementniveauet ser vi, at $(\mathbf{A}+\mathbf{A}')_{i,j} = a_{ij}+a'_{ij}$. Addition og multiplikation af matricer opfylder mange af de samme regler som addition og multiplikation af reelle eller komplekse tal. Vi samler nogle af disse i følgende sætning. Den primære forskel, som allerede blev nævnt tidligere, er, at matrixmultiplikation generelt ikke er kommutativ.

Sætning 7.2.1

Lad $\mathbb{F}$ være et legeme. Da gælder

(i) $\mathbf{A}+\mathbf{A}' = \mathbf{A}'+\mathbf{A}$ for alle $\mathbf{A}, \mathbf{A}' \in \mathbb{F}^{m\times n}$.

(ii) $(\mathbf{A}+\mathbf{A}')+\mathbf{A}'' = \mathbf{A}+(\mathbf{A}'+\mathbf{A}'')$ for alle $\mathbf{A}, \mathbf{A}', \mathbf{A}'' \in \mathbb{F}^{m\times n}$.

(iii) $\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C}) = (\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C}$ for alle $\mathbf{A} \in \mathbb{F}^{m\times n}$, $\mathbf{B} \in \mathbb{F}^{n\times \ell}$ og $\mathbf{C} \in \mathbb{F}^{\ell\times k}$.

(iv) $\mathbf{A}\cdot(\mathbf{B}+\mathbf{B}') = \mathbf{A}\cdot\mathbf{B}+\mathbf{A}\cdot\mathbf{B}'$ for alle $\mathbf{A} \in \mathbb{F}^{m\times n}$ og $\mathbf{B}, \mathbf{B}' \in \mathbb{F}^{n\times \ell}$.

(v) $(\mathbf{A}+\mathbf{A}')\cdot\mathbf{B} = \mathbf{A}\cdot\mathbf{B}+\mathbf{A}'\cdot\mathbf{B}$ for alle $\mathbf{A}, \mathbf{A}' \in \mathbb{F}^{m\times n}$ og $\mathbf{B} \in \mathbb{F}^{n\times \ell}$.

Bevis. Vi vil bevise det tredje punkt og overlade de andre dele til læseren. Anvendes Ligning (7.6) på produktet $(\mathbf{B}\cdot\mathbf{C})$ får vi, at $(\mathbf{B}\cdot\mathbf{C})_{s,j} = \sum_{r=1}^{\ell} b_{sr}\cdot c_{rj}$. Ved at benytte dette samt Ligning (7.6) på produktet $\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C})$ opnår vi efter en smule omskrivning, at:

$$\begin{aligned}
(\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C}))_{i,j} &= \sum_{s=1}^{n} a_{is}\cdot(\mathbf{B}\cdot\mathbf{C})_{s,j} \\
&= \sum_{s=1}^{n} a_{is}\cdot\sum_{r=1}^{\ell} b_{sr}\cdot c_{rj} \\
&= \sum_{s=1}^{n}\sum_{r=1}^{\ell} a_{is}\cdot(b_{sr}\cdot c_{rj}) \\
&= \sum_{s=1}^{n}\sum_{r=1}^{\ell} (a_{is}\cdot b_{sr})\cdot c_{rj} \\
&= \sum_{r=1}^{\ell}\sum_{s=1}^{n} (a_{is}\cdot b_{sr})\cdot c_{rj} \\
&= \sum_{r=1}^{\ell}\left(\sum_{s=1}^{n} a_{is}\cdot b_{sr}\right)\cdot c_{rj}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{r=1}^{\ell} (\mathbf{B} \cdot \mathbf{C})_{i,r} \cdot c_{rj} \\
&= ((\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C})_{i,j}.
\end{aligned}
$$

□

Vi afslutter dette afsnit med at forklare to yderligere matrixoperationer. Vi har allerede set, at vektorer kan multipliceres med en skalar. Generaliseringen til matricer er umiddelbar: for $c \in \mathbb{F}$ og

$$
\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{F}^{m \times n} \text{ defineres } c \cdot \mathbf{A} = \begin{bmatrix} c \cdot a_{11} & \cdots & c \cdot a_{1n} \\ \vdots & & \vdots \\ c \cdot a_{m1} & \cdots & c \cdot a_{mn} \end{bmatrix}. \tag{7.7}
$$

Afslutningsvis findes der en måde at ombytte rollerne af rækker og søjler i en matrix $\mathbf{A}$ på. Dette kaldes *transponering* af matricen og kan betegnes $\mathbf{A}^T$. Mere præcist, hvis vi har givet

$$
\mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \in \mathbb{F}^{m \times n}, \text{ så defineres } \mathbf{A}^T = \begin{bmatrix} a_{11} & \dots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{bmatrix}. \tag{7.8}
$$

Eksempel 7.2.3

Lad matricen

$$
\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \in \mathbb{R}^{2 \times 3}
$$

være givet. Bestem $\mathbf{A}^T$.

Svar:

Vi har

$$
\mathbf{A}^T = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.
$$

Bemærk, at hvis $\mathbf{A} \in \mathbb{F}^{m \times n}$, så er $\mathbf{A}^T \in \mathbb{F}^{n \times m}$. På elementniveauet har vi ganske enkelt, at det (i, j)'te element i $\mathbf{A}^T$ er lig med det (j, i)'te element i $\mathbf{A}$.

Transponeringen opfører sig fint i forbindelse med matrixaddition og matrixprodukter. Dette præciseres i følgende sætning.

Sætning 7.2.2

Lad $\mathbb{F}$ være et legeme. Da gælder:

(i) $(\mathbf{A}^T)^T = \mathbf{A}$ for alle $\mathbf{A} \in \mathbb{F}^{m \times n}$.

(ii) $(\mathbf{A} + \mathbf{A}')^T = \mathbf{A}^T + (\mathbf{A}')^T$ for alle $\mathbf{A}, \mathbf{A}' \in \mathbb{F}^{m \times n}$.

(iii) $(\mathbf{A} \cdot \mathbf{B})^T = \mathbf{B}^T \cdot \mathbf{A}^T$ for alle $\mathbf{A} \in \mathbb{F}^{m \times n}$ og $\mathbf{B} \in \mathbb{F}^{n \times \ell}$.

Bevis. Vi beviser kun den første påstand. Generelt er det (i, j)'te element i $\mathbf{B}^T$ lig med det (j, i)'te element i $\mathbf{B}$ for enhver matrix $\mathbf{B}$. Anvendes dette først på matricen $\mathbf{A}^T$ og derefter på matricen $\mathbf{A}$, får vi, at $((\mathbf{A}^T)^T)_{i,j} = (\mathbf{A}^T)_{j,i} = (\mathbf{A})_{i,j}$. Dette viser, at matricerne $(\mathbf{A}^T)^T$ og $\mathbf{A}$ har præcis de samme elementer, og dermed at de er ens. □

Det er vigtigt at bemærke sig rækkefølgen af multiplikationen i punkt 3 før og efter transponering. Transponering så at sige vender rækkefølgen af elementer i et produkt om. Der er en god grund til dette. Givet matricer $\mathbf{A} \in \mathbb{F}^{m \times n}$ og $\mathbf{B} \in \mathbb{F}^{n \times \ell}$, så er produktet $\mathbf{A}^T \cdot \mathbf{B}^T$ generelt set ikke engang defineret! Antallet af søjler i $\mathbf{A}^T$ er m, mens antallet af rækker i $\mathbf{B}^T$ er ℓ. Dog er der ingen problemer med produktet $\mathbf{B}^T \cdot \mathbf{A}^T$, da antallet af søjler i $\mathbf{B}^T$ er n, hvilket er det samme som antallet af rækker i $\mathbf{A}^T$. Selvom disse observationer ikke direkte beviser punkt tre fra Sætning 7.2.2, så forklarer de, hvorfor det er ganske naturligt, at multiplikationsrækkefølgen er som angivet.

7.3 Kvadratiske matricer

Hvis antallet af rækker og af søjler i en matrix er ens, kaldes den en *kvadratisk* matrix. Med andre ord, en matrix $\mathbf{A}$ er en kvadratisk matrix, hvis $\mathbf{A} \in \mathbb{F}^{n \times n}$ for et positivt heltal n. De elementer, der befinder sig på de (i, i)'te positioner i en kvadratisk matrix, kaldes matricens *diagonalelementer*. For eksempel er diagonalelementerne i matricen

$$\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

tallene 1, 5 og 9. Samlet udgør diagonalelementerne, hvad der kaldes *diagonalen* i en kvadratisk matrix.

For et givet n kaldes den $n \times n$-matrix, der har 1'ere i sin diagonal og 0'er overalt derudover, for *identitetsmatricen*, og den betegnes $\mathbf{I}_n$. For $n = 4$ som et eksempel har vi dermed

$$\mathbf{I}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Denne matrix kaldes identitetsmatricen fordi den ikke har nogen effekt på en vektor, når den multipliceres med vektoren. En direkte udregning viser, at $\mathbf{I}_n \cdot \mathbf{v} = \mathbf{v}$ for alle $\mathbf{v} \in \mathbb{F}^n$. Derfor er funktionen $L : \mathbb{F}^n \to \mathbb{F}^n$ defineret ved $L(\mathbf{v}) = \mathbf{I}_n \cdot \mathbf{v}$ blot identitetsfunktionen. Med denne matrix på plads er følgende definition mulig:

Definition 7.3.1

En kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ kaldes *invertibel*, hvis der findes en matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$, således at

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A} = \mathbf{I}_n.$$

Matricen $\mathbf{B}$, hvis den eksisterer, kaldes den inverse matrix af $\mathbf{A}$ og betegnes ved $\mathbf{A}^{-1}$.

Invertible matricer vil dukke op i mange situationer senere, men allerede når man løser visse systemer af lineære ligninger, kan de være nyttige. Antag for eksempel, at man ønsker at løse systemet af lineære ligninger $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$, hvor vi har en kvadratisk koefficientmatrix $\mathbf{A}$, en vektor indeholdende de variable $\mathbf{x} = (x_1, \ldots, x_n)$ samt en højreside $\mathbf{b} = (b_1, \ldots, b_n)$. Hvis koefficientmatricen $\mathbf{A}$ har en invers, kan vi multiplicere fra venstre med $\mathbf{A}^{-1}$ og reducere: $\mathbf{A}^{-1} \cdot (\mathbf{A} \cdot \mathbf{x}) = (\mathbf{A}^{-1} \cdot \mathbf{A}) \cdot \mathbf{x} = \mathbf{I}_n \cdot \mathbf{x}$. Dette betyder, at ligningen $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ medfører, at $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$. Omvendt, hvis $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$, så vil en multiplikation med $\mathbf{A}$ fra venstre give $\mathbf{A} \cdot \mathbf{x} = \mathbf{A} \cdot (\mathbf{A}^{-1} \cdot \mathbf{b}) = (\mathbf{A} \cdot \mathbf{A}^{-1}) \cdot \mathbf{b} = \mathbf{I}_n \cdot \mathbf{b} = \mathbf{b}$. Dermed har vi vist, at:

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \text{ hvis og kun hvis } \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}, \text{ forudsat at } \mathbf{A}^{-1} \text{ eksisterer.} \tag{7.9}$$

Denne observation leder til en fin konsekvens for rangen af invertible matricer:

Lemma 7.3.1

Lad $\mathbf{A} \in \mathbb{F}^{n\times n}$ være givet, og antag, at dens inverse matrix eksisterer. Da er $\rho(\mathbf{A}) = n$.

Bevis. Ligning (7.9) medfører, at det homogene system af lineære ligninger $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$ kun har løsningen $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{0} = \mathbf{0}$. Men ifølge Korollar 6.4.5 vil rangen af $\mathbf{A}$ da være lig med n. $\square$

Vi kunne nævne endnu flere sammenhænge, men det vil vi vende tilbage til senere. Spørgsmålet er nu, hvordan man finder ud af, hvornår en matrix har en invers, og hvis den gør, hvordan man bestemmer den. Vi vil først finde et algoritmisk svar og derefter beskrive en teoretisk karakterisering af invertible matricer.

Det første vi vil gøre, er at finde en algoritme, der for en given $n \times n$-matrix $\mathbf{A}$ bestemmer en $n \times n$-matrix $\mathbf{B}$, således at $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$, hvis en sådan eksisterer. Derfor vil outputtet fra algoritmen enten være, at en sådan $\mathbf{B}$ ikke eksisterer, eller den vil returnere en sådan $\mathbf{B}$. Bemærk, at ifølge Definition 7.3.1 skal den inverse af $\mathbf{A}$, her betegnet $\mathbf{B}$, opfylde $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$ *og* $\mathbf{B} \cdot \mathbf{A} = \mathbf{I}_n$. Heldigvis viser det sig, at $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$ *medfører* $\mathbf{B} \cdot \mathbf{A} = \mathbf{I}_n$, således at den algoritme, vi er ved at beskrive, rent faktisk vil bestemme den inverse matrix $\mathbf{B} = \mathbf{A}^{-1}$, forudsat at den eksisterer.

Lad os betegne den i'te søjle i enhedsmatricen $\mathbf{I}_n$ ved $\mathbf{e}_i$ for $i = 1, \ldots, n$. Da har vi for eksempel for $n = 4$, at

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \text{ og } \mathbf{e}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Idéen med algoritmen, der finder inverse matricer, er følgende: Vi forsøger at finde en matrix $\mathbf{B} \in \mathbb{F}^{n \times n}$, således at $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$ for en given $\mathbf{A} \in \mathbb{F}^{n \times n}$. Lad os nu betegne søjlerne i $\mathbf{B}$ ved $\mathbf{b}_1, \ldots, \mathbf{b}_n$. Den i'te søjle i $\mathbf{A} \cdot \mathbf{B}$ er jævnfør definitionen af matrixproduktet lig med $\mathbf{A} \cdot \mathbf{b}_i$, mens den i'te søjle i $\mathbf{I}_n$ er lig med $\mathbf{e}_i$. Derfor er $\mathbf{A} \cdot \mathbf{b}_i = \mathbf{e}_i$ for alle i mellem 1 og n. Omvendt, hvis $\mathbf{A} \cdot \mathbf{b}_i = \mathbf{e}_i$ for alle i mellem 1 og n, så har matricerne $\mathbf{A} \cdot \mathbf{B}$ og $\mathbf{I}_n$ de samme søjler, og derfor gælder $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$. Dermed ser vi, at

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n, \text{ hvis og kun hvis } \mathbf{A} \cdot \mathbf{b}_i = \mathbf{e}_i \text{ for alle } i \text{ mellem 1 og } n.$$

Derfor kan vi bestemme $\mathbf{b}_i$ ved at løse det inhomogene system af lineære ligninger $\mathbf{A} \cdot \mathbf{x} = \mathbf{e}_i$.

Fra teorien fra det foregående kapitel ser vi, at for at finde ud af, om ligningssystemet $\mathbf{A} \cdot \mathbf{x} = \mathbf{e}_i$ har en løsning, er det tilstrækkeligt at bestemme den reducerede trappeform af totalmatricen $[\mathbf{A}|\mathbf{e}_i]$. Hvis $\rho(\mathbf{A}) = \rho([\mathbf{A}|\mathbf{e}_i])$, så findes der ifølge Korollar 6.4.3 en løsning, og ellers ikke. Hvis der for alle i mellem 1 og n gælder, at $\rho(\mathbf{A}) = \rho([\mathbf{A}|\mathbf{e}_i])$, vil vi derfor være i stand til at bestemme en matrix $\mathbf{B} \in \mathbb{F}^{n \times n}$, således at $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$.

Man kunne nu vælge at arbejde sig igennem ligningssystemet $\mathbf{A} \cdot \mathbf{x} = \mathbf{e}_i$ ét i ad gangen og dermed beregne én søjle i matricen $\mathbf{B}$ ad gangen, hvis den eksisterer. Men da første del af de tilsvarende totalmatricer altid er ens, nemlig $\mathbf{A}$, vil det være langt hurtigere at håndtere alle n systemer på én gang ved at bestemme den reducerede trappeform af matricen $[\mathbf{A}|\mathbf{e}_1|\mathbf{e}_2|\ldots|\mathbf{e}_n] = [\mathbf{A}|\mathbf{I}_n]$.

Derfor er en algoritme, der kan afgøre, om en kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$ har en invers og i bekræftende fald bestemme den, følgende:

(i) Bestem den reducerede trappeform af $n \times 2n$-matricen $[\mathbf{A}|\mathbf{I}_n]$. Dette kan gøres ved hjælp af elementære rækkeoperationer, ligesom vi gjorde i Afsnit 6.3.

(ii) Hvis den resulterende reducerede trappeform ikke er på formen $[\mathbf{I}_n|\mathbf{B}]$, konkludér da, at $\mathbf{A}$ ikke har en invers.

(iii) Hvis den er på formen $[\mathbf{I}_n|\mathbf{B}]$, konkludér da, at $\mathbf{A}$ har en invers, nemlig $\mathbf{A}^{-1} = \mathbf{B}$.

For at se hvordan dette virker i praksis, vil vi gennemgå to eksempler.

Eksempel 7.3.1
Lad $\mathbb{F} = \mathbb{R}$ og

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Afgør, om denne matrix har en invers, og bestem den i så fald.

Svar: Først bestemmer vi den reducerede trappeform af matricen $[\mathbf{A}|\mathbf{I}_2]$. Vi får:

$$\begin{aligned}
[\mathbf{A}|\mathbf{I}_2] &= \begin{bmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{bmatrix} \underset{R_2 \leftarrow R_2 - 3 \cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{bmatrix} \\
&\underset{R_2 \leftarrow (-1/2) \cdot R_2}{\longrightarrow} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 3/2 & -1/2 \end{bmatrix} \underset{R_1 \leftarrow R_1 - 2 \cdot R_2}{\longrightarrow} \begin{bmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{bmatrix}.
\end{aligned}$$

Derfor konkluderer vi, at $\mathbf{A}$ har en invers, nemlig

$$\mathbf{A}^{-1} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}.$$

Eksempel 7.3.2
Lad $\mathbb{F} = \mathbb{R}$ og

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix}.$$

Afgør, om denne matrix har en invers, og bestem den i så fald.

Svar: Vi starter med at bestemme den reducerede trappeform af matricen $[\mathbf{A}|\mathbf{I}_3]$. Vi får:

$$\begin{aligned}
[\mathbf{A}|\mathbf{I}_3] &= \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 5 & 7 & 9 & 0 & 0 & 1 \end{bmatrix} \\
\underset{R_2 \leftarrow R_2 - 4 \cdot R_1}{\longrightarrow} & \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 5 & 7 & 9 & 0 & 0 & 1 \end{bmatrix} \\
\underset{R_3 \leftarrow R_3 - 5 \cdot R_1}{\longrightarrow} & \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & -3 & -6 & -5 & 0 & 1 \end{bmatrix} \\
\underset{R_3 \leftarrow R_3 - R_2}{\longrightarrow} & \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 \end{bmatrix}.
\end{aligned}$$

Selvom vi endnu ikke er færdige med bestemmelsen af den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_3]$, har vi allerede her fundet en trappeform af den. Pivotelementerne kan aflæses på dette stadie, og vi ser, at de er placeret i matricens første, anden og fjerde søjle. Fortsætter vi reduktionen til den reducerede trappeform, vil de første tre elementer i tredje række forblive nul. Læseren opfordres til at udregne den reducerede trappeform for at sikre sig, at dette faktisk er korrekt. Dermed vil den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_3]$ ikke være på formen $[\mathbf{I}_3|\mathbf{B}]$. Vi konkluderer, at matricen $\mathbf{A}$ ikke har en invers.

I princippet har vi nu en algoritme, der kan afgøre, om en kvadratisk matrix har en invers, og som i bekræftende fald kan bestemme den. Vi har dog ikke vist, at algoritmen er korrekt. Med andre ord, hvis vi følger trinnene i algoritmen, vil resultatet så altid være, som det skal være? Først og fremmest skal vi sikre os, at hvis den reducerede trappeform af $n \times 2n$-matricen $[\mathbf{A}|\mathbf{I}_n]$ ikke er på formen $[\mathbf{I}_n|\mathbf{B}]$, så har $\mathbf{A}$ ikke en invers. Og for det andet skal vi sikre os, at hvis en matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$ opfylder $\mathbf{A}\cdot\mathbf{B} = \mathbf{I}_n$, så gælder der også, at $\mathbf{B}\cdot\mathbf{A} = \mathbf{I}_n$, således at vi kan konkludere, at $\mathbf{B}$ er den inverse af $\mathbf{A}$. Vi vil adressere disse spørgsmål i resten af dette afsnit. Det vil vise sig, at alt er, som det skal være, og man kan vise, at:

$$\begin{aligned}\mathbf{A}^{-1} \text{ eksisterer } &\Leftrightarrow \text{ den reducerede trappeform af } [\mathbf{A}|\mathbf{I}_n] \text{ er på formen } [\mathbf{I}_n|\mathbf{B}]\\ &\Leftrightarrow \rho(\mathbf{A}) = n \text{ (dvs. rangen af } \mathbf{A} \text{ er } n).\end{aligned} \tag{7.10}$$

En læser, der er villig til at acceptere dette uden bevis, kan springe resten af dette afsnit over, men for andre læsere vil vi give et bevis nedenfor.

Sætning 7.3.2

Lad $\mathbf{A} \in \mathbb{F}^{n\times n}$ være en kvadratisk matrix. Følgende udsagn er logisk ækvivalente:

(i) Den reducerede trappeform af $n \times 2n$-matricen $[\mathbf{A}|\mathbf{I}_n]$ er på formen $[\mathbf{I}_n|\mathbf{B}]$ for en kvadratisk matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$.

(ii) Der eksisterer en kvadratisk matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$, således at $\mathbf{A}\cdot\mathbf{B} = \mathbf{I}_n$.

Bevis. Antag, at den reducerede trappeform af matricen $[\mathbf{A}|\mathbf{I}_n]$ er på formen $[\mathbf{I}_n|\mathbf{B}]$ for en kvadratisk matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$. Lad $\mathbf{b}_i$ være den i'te søjle i matricen $\mathbf{B}$. Så kan de samme elementære rækkeoperationer, der anvendes til at transformere matricen $[\mathbf{A}|\mathbf{I}_n]$ til formen $[\mathbf{I}_n|\mathbf{B}]$, også anvendes til at transformere matricen $[\mathbf{A}|\mathbf{e}_i]$ til $[\mathbf{I}_n|\mathbf{b}_i]$. Da $[\mathbf{I}_n|\mathbf{b}_i]$ er på reduceret trappeform, kan vi konkludere, at den reducerede trappeform af matricen $[\mathbf{A}|\mathbf{e}_i]$ er lig med $[\mathbf{I}_n|\mathbf{b}_i]$. Dette medfører, at $\mathbf{b}_i$ er en løsning til systemet af lineære ligninger $\mathbf{A}\cdot\mathbf{x} = \mathbf{e}_i$. Derfor gælder $\mathbf{A}\cdot\mathbf{B} = \mathbf{I}_n$. Specifikt gælder $\mathbf{A}\cdot\mathbf{B} = \mathbf{I}_n$ for en kvadratisk matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$, nemlig matricen i højre del af den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_n]$.

Antag omvendt, at $\mathbf{A}\cdot\mathbf{B} = \mathbf{I}_n$ for en kvadratisk matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$. Da har systemet af lineære ligninger $\mathbf{A}\cdot\mathbf{x} = \mathbf{e}_i$ for alle i fra 1 til n en løsning, nemlig den i'te søjle i matricen $\mathbf{B}$. Vi påstår, at den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_n]$ kun indeholder pivotelementer i sine første n søjler. Vi vil bevise denne påstand ved et modstridbevis. Antag derfor, at den reducerede trappeform af

$[\mathbf{A}|\mathbf{I}_n]$ har en pivot i en søjle med indeks $n+i$ for et $i > 0$. Da vil den reducerede trappeform af matricen $[\mathbf{A}|\mathbf{e}_i]$ indeholde en pivot i sin $(n+1)$'te søjle. Specifikt vil $\mathbf{A}$ og $[\mathbf{A}|\mathbf{e}_i]$ ikke have samme rang. Men så har systemet $\mathbf{A} \cdot \mathbf{x} = \mathbf{e}_i$ ifølge Korollar 6.4.3 ingen løsning. Da vi allerede har observeret, at det har en løsning, opnår vi her en modstrid. Dette beviser påstanden om, at den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_n]$ kun indeholder pivotelementer i sine første n søjler. Næste påstand er, at rangen af $[\mathbf{A}|\mathbf{I}_n]$ er lig med n. For at opnå en modstrid vil vi antage, at den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_n]$ indeholder en nulrække. Ved at se på anden del af matricen, $\mathbf{I}_n$, kan vi konkludere, at der findes en sekvens af elementære rækkeoperationer, der kan transformere $\mathbf{I}_n$ om til en matrix med en nulrække. Men mens $\mathbf{I}_n$ er en matrix med rang n, kan en $n \times n$-matrix med en nulrække højst have rang $n-1$. Dette beviser den anden påstand. Ved at kombinere de to påstande konkluderer vi, at den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_n]$ indeholder en pivot i hver af sine første n søjler. Så er den på formen $[\mathbf{I}_n|\mathbf{C}]$, hvor højre del er en kvadratisk matrix $\mathbf{C} \in \mathbb{F}^{n\times n}$. □

Korollar 7.3.3

Lad $\mathbf{A} \in \mathbb{F}^{n\times n}$ være givet. Da eksisterer der en matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$, således at $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$, hvis og kun hvis $\rho(\mathbf{A}) = n$.

Bevis. Hvis $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$ for en $\mathbf{B} \in \mathbb{F}^{n\times n}$, så er den reducerede trappeform af $n \times 2n$-matricen $[\mathbf{A}|\mathbf{I}_n]$ ifølge Sætning 7.3.2 på formen $[\mathbf{I}_n|\mathbf{C}]$ for en $\mathbf{C} \in \mathbb{F}^{n\times n}$. Men så er den reducerede trappeform af $\mathbf{A}$ selv $\mathbf{I}_n$, hvilket medfører, at $\rho(\mathbf{A}) = n$.

Omvendt, hvis $\rho(\mathbf{A}) = n$, er den reducerede trappeform af $\mathbf{A}$ lig med $\mathbf{I}_n$. Men så er den reducerede trappeform af $[\mathbf{A}|\mathbf{I}_n]$ på formen $[\mathbf{I}_n|\mathbf{C}]$ for en kvadratisk matrix $\mathbf{C} \in \mathbb{F}^{n\times n}$. Ifølge Sætning 7.3.2 kan vi konkludere, at der eksisterer en $\mathbf{B} \in \mathbb{F}^{n\times n}$, således at $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$. □

Korollar 7.3.4

Lad $\mathbf{A} \in \mathbb{F}^{n\times n}$ være en kvadratisk matrix, og antag, at der eksisterer en kvadratisk matrix $\mathbf{B} \in \mathbb{F}^{n\times n}$, således at $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$. Da gælder $\mathbf{B} \cdot \mathbf{A} = \mathbf{I}_n$, og derfor er $\mathbf{B} = \mathbf{A}^{-1}$ den inverse af $\mathbf{A}$.

Bevis. For at konkludere, at $\mathbf{B}$ er den inverse af $\mathbf{A}$, skal vi vise, at $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A} = \mathbf{I}_n$. Da det er givet, at $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_n$, har vi tilbage at vise, at $\mathbf{B} \cdot \mathbf{A} = \mathbf{I}_n$.

Bemærk nu, at $\mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{A}) = (\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{A} = \mathbf{I}_n \cdot \mathbf{A} = \mathbf{A} = \mathbf{A} \cdot \mathbf{I}_n$. Derfor gælder $\mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{A} - \mathbf{I}_n) = \mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{A}) - \mathbf{A} \cdot \mathbf{I}_n = \mathbf{0}$, hvor $\mathbf{0}$ her betegner $n \times n$-nulmatricen.

Bemærk, at den forrige ligning medfører, at enhver søjle i $\mathbf{B} \cdot \mathbf{A} - \mathbf{I}_n$ er en løsning til det homogene system af ligninger $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$. Ifølge det forrige korollar har matricen $\mathbf{A}$ dog rang n. Derfor ved vi fra Korollar 6.4.5, at systemet $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$ kun har løsningen $\mathbf{x} = \mathbf{0}$. Derfor er alle søjler i $\mathbf{B} \cdot \mathbf{A} - \mathbf{I}_n$ nulsøjler, hvilket medfører, at $\mathbf{B} \cdot \mathbf{A} = \mathbf{I}_n$. Dette er netop, hvad vi skulle vise. □

Korollar 7.3.5

Lad $\mathbf{A} \in \mathbb{F}^{n \times n}$ være givet. Da eksisterer den inverse matrix, hvis og kun hvis $\rho(\mathbf{A}) = n$.

Bevis. Dette følger af de to tidligere korollarer. $\square$

Kapitel 8

Determinanter

8.1 Determinanten af en kvadratisk matrix

I dette afsnit vil vi introducere *determinanten* af en kvadratisk matrix. Determinanter bliver nyttige, når vi undersøger, om en given matrix er invertibel, og vil også blive særdeles nyttige i senere kapitler. Vi starter med en notationskonvention:

Definition 8.1.1

Lad $\mathbf{A} = (a_{ij})_{1\le i\le n, 1\le j\le n} \in \mathbb{F}^{n\times n}$ være en givet kvadratisk matrix. Da definerer vi matricen $\mathbf{A}(i;j) \in \mathbb{F}^{(n-1)\times(n-1)}$ som:

$$\mathbf{A}(i;j) = \begin{bmatrix} a_{11} & \dots & a_{1\,j-1} & a_{1\,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1\,1} & \dots & a_{i-1\,j-1} & a_{i-1\,j+1} & \dots & a_{i-1\,n} \\ a_{i+1\,1} & \dots & a_{i+1\,j-1} & a_{i+1\,j+1} & \dots & a_{i+1\,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n\,j-1} & a_{n\,j+1} & \dots & a_{nn} \end{bmatrix}.$$

I ord: Matricen $\mathbf{A}(i;j)$ opnås fra $\mathbf{A}$ ved at den i'te række og den j'te søjle i $\mathbf{A}$ fjernes. Med dette på plads kan vi definere determinanten af en kvadratisk matrix rekursivt som følger:

Definition 8.1.2

Lad $\mathbf{A} \in \mathbb{F}^{n\times n}$ være en kvadratisk matrix. Vi definerer da

$$\det(\mathbf{A}) = \begin{cases} \mathbf{A} & \text{hvis } n = 1, \\ \sum_{i=1}^{n} (-1)^{i+1} \cdot a_{i1} \cdot \det(\mathbf{A}(i;1)) & \text{hvis } n \geq 2. \end{cases}$$

Uden brug af summationssymbolet kan man også skrive:

$$\det(\mathbf{A}) = a_{11} \cdot \det(\mathbf{A}(1;1)) - a_{21} \cdot \det(\mathbf{A}(2;1)) + \cdots + (-1)^{n+1} \cdot a_{n1} \cdot \det(\mathbf{A}(n;1)).$$

Eksempel 8.1.1

Lad

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

For at beregne determinanten af denne matrix vil vi benytte Definition 8.1.2. Vi ser, at $\mathbf{A}(1;1) = a_{22}$, og at $\mathbf{A}(2;1) = a_{12}$. Derfor har vi

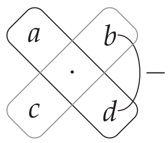
$$\det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = a_{11} \cdot \det(a_{22}) - a_{21} \cdot \det(a_{12}) = a_{11}a_{22} - a_{21}a_{12}.$$

Benyttes bogstaver a, b, c og d som matricens elementer, kan vi skrive dette som:

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc. \tag{8.1}$$

Når man får til opgave at beregne determinanten af en 2×2-matrix, er denne formel særdeles praktisk. Figur 8.1 visualiserer formlen for determinanten af en 2×2-matrix: udregningen foregår ved, at man multiplicerer matricens to diagonalelementer og trækker produktet af de to andre elementer fra.

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) =$$



$$= ad - cb.$$

Figur 8.1: Determinanten af en 2×2-matrix.

Eksempel 8.1.2

Lad $\mathbb{F} = \mathbb{R}$ samt

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix}$$

ligesom i Eksempel 7.3.2. Udregn determinanten af $\mathbf{A}$.

Svar: Vi har først og fremmest, at

$$\mathbf{A}(1;1) = \begin{bmatrix} 5 & 6 \\ 7 & 9 \end{bmatrix}, \mathbf{A}(2;1) = \begin{bmatrix} 2 & 3 \\ 7 & 9 \end{bmatrix}, \text{ og } \mathbf{A}(3;1) = \begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix}.$$

Derfor får vi ved brug af Definition 8.1.2:

$$\det(\mathbf{A}) = 1 \cdot \det\left(\begin{bmatrix} 5 & 6 \\ 7 & 9 \end{bmatrix}\right) - 4 \cdot \det\left(\begin{bmatrix} 2 & 3 \\ 7 & 9 \end{bmatrix}\right) + 5 \cdot \det\left(\begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix}\right).$$

Med Ligning (8.1) kan vi hurtigt beregne determinanterne af 2×2-matricer. Dermed får vi:

$$\det(\mathbf{A}) = 1 \cdot (45 - 42) - 4 \cdot (18 - 21) + 5 \cdot (12 - 15) = 3 + 12 - 15 = 0.$$

Vi får senere hen nogle flere teknikker til beregning af determinanter af matricer, men for nu vil vi koncentrere os om en særlig klasse af matricer.

Definition 8.1.3

En matrix $\mathbf{A} = \mathbb{F}^{n \times n}$ kaldes en *diagonalmatrix*, hvis der findes $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$, således at

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_{n-1} & 0 \\ 0 & \ldots & 0 & 0 & \lambda_n \end{bmatrix}.$$

Vi har tidligere nævnt, at elementerne $a_{11}, \ldots, a_{nn}$ i en kvadratisk matrix $\mathbf{A} = (a_{ij})_{1 \le i \le n; 1 \le j \le n}$ kaldes diagonalelementer i $\mathbf{A}$. Dette forklarer betegnelsen diagonalmatrix i Definition 8.1.3: en diagonalmatrix er en kvadratisk matrix, hvis elementer alle er nuller bortset fra dem i diagonalen. For eksempel er identitetsmatricen $\mathbf{I}_n$, som introduceret i starten af Afsnit 7.3, en diagonalmatrix, hvis diagonalelementer alle er lig med 1.

Proposition 8.1.1

Lad $\mathbf{A} = \mathbb{F}^{n \times n}$ være en diagonalmatrix med diagonalelementerne $\lambda_1, \ldots, \lambda_n$. Da er

$$\det(\mathbf{A}) = \lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n.$$

Særligt er $\det(\mathbf{I}_n) = 1$.

Bevis. Vi viser dette ved induktion på n. Hvis $n = 1$, så er $\mathbf{A} = \lambda_1$, og Definition 8.1.2 medfører så, at $\det(\mathbf{A}) = \lambda_1$. Antag nu, at $n \geq 2$, og at udsagnet gælder for diagonalmatricer i $\mathbb{F}^{(n-1)\times(n-1)}$. Ved at anvende Definition 8.1.2 får vi, at:

$$\begin{aligned}\det(\mathbf{A}) &= \lambda_1 \cdot \det(\mathbf{A}(1;1)) - 0 \cdot \det(\mathbf{A}(2;1)) + \cdots + (-1)^{n+1} \cdot 0 \cdot \det(\mathbf{A}(n;1)) \\ &= \lambda_1 \cdot \det(\mathbf{A}(1;1)) \\ &= \lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n,\end{aligned}$$

hvor vi i den sidste lighed anvendte induktionshypotesen. Induktionshypotesen gælder, da $\mathbf{A}(1;1)$ er en diagonalmatrix med diagonalelementerne $\lambda_2, \ldots, \lambda_n$. Dette afslutter induktionstrinnet. Vi konkluderer ved induktionsprincippet, at udsagnet er sandt. Det særlige tilfælde med identitetsmatricen følger nu straks, da alle diagonalelementer i så fald er lig med én. □

Vi kan allerede nu opstille en formel for determinanten af matricer i en større klasse, nemlig de øvre trekantmatricer:

Definition 8.1.4

En matrix $\mathbf{A} = \mathbb{F}^{n\times n}$ kaldes en *øvre trekantsmatrix*, hvis der findes $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ og $a_{ij} \in \mathbb{F}$ for $1 \le i < j \le n$, således at

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & a_{12} & a_{13} & \ldots & a_{1n} \\ 0 & \lambda_2 & a_{23} & \ldots & a_{2n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_{n-1} & a_{n-1\,n} \\ 0 & \ldots & 0 & 0 & \lambda_n \end{bmatrix}.$$

I ord: i en øvre trekantsmatrix er alle de elementer, der ikke er nul, placeret over eller i diagonalen. Alle elementer under diagonalen i en øvre trekantsmatrix er nul.

Sætning 8.1.2

Lad $\mathbf{A} = \mathbb{F}^{n\times n}$ være en øvre trekantsmatrix med diagonalelementerne $\lambda_1, \ldots, \lambda_n$. Da er

$$\det(\mathbf{A}) = \lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n.$$

Bevis. Beviset er meget lig beviset for Proposition 8.1.1 og overlades til læseren. □

En tilsvarende type matricer er følgende:

Definition 8.1.5

En matrix $\mathbf{A} = \mathbb{F}^{n\times n}$ kaldes en *nedre trekantsmatrix*, hvis der findes $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ og $a_{ij} \in \mathbb{F}$

for $1 \leq j < i \leq n$, således at

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ a_{21} & \lambda_2 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_{n-1\,1} & \dots & a_{n-1\,n-2} & \lambda_{n-1} & 0 \\ a_{n\,1} & \dots & a_{n\,n-2} & a_{n\,n-1} & \lambda_n \end{bmatrix}.$$

I ord: i en nedre trekantsmatrix er alle de elementer, der ikke er nul, placeret under eller i diagonalen. Alle elementer over diagonalen i en nedre trekantsmatrix er nul. Også her kan vi opskrive en formel for dens determinant. Inden da har vi brug for et lemma, som kan være nyttigt i sig selv.

Lemma 8.1.3

Hvis en kvadratisk matrix i $\mathbb{F}^{n\times n}$ indeholder en nulrække, er dens determinant nul.

Bevis. Dette kan vises ved induktion på n. Detaljerne overlades til læseren. □

Sætning 8.1.4

Lad $\mathbf{A} = \mathbb{F}^{n\times n}$ være en nedre trekantsmatrix med diagonalelementerne $\lambda_1, \dots, \lambda_n$. Da er

$$\det(\mathbf{A}) = \lambda_1 \cdot \dots \cdot \lambda_n.$$

Bevis. Vi viser dette ved induktion på n. Hvis $n = 1$, så er $\mathbf{A} = \lambda_1$, og Definition 8.1.2 medfører så, at $\det(\mathbf{A}) = \lambda_1$. Antag nu, at $n \geq 2$, og at udsagnet gælder for nedre trekantsmatricer i $\mathbb{F}^{(n-1)\times(n-1)}$. Bemærk, at $\mathbf{A}(1;1)$ er en nedre trekantsmatrix med diagonalelementerne $\lambda_2, \dots, \lambda_n$. Derfor medfører induktionshypotesen, at $\det(\mathbf{A}(1;1)) = \lambda_2 \cdot \dots \cdot \lambda_n$. Matricerne $\mathbf{A}(2;1), \dots, \mathbf{A}(n;1)$ har alle en nulrække som første række. Dette skyldes, at det eneste element, der ikke er nul, i første række i $\mathbf{A}$, er at finde på den første elementposition, men denne position er blevet fjernet i matricerne $\mathbf{A}(2;1), \dots, \mathbf{A}(n;1)$. Ved Lemma 8.1.3 har vi derfor, at $\det(\mathbf{A}(2;1)) = 0, \dots, \det(\mathbf{A}(n;1)) = 0$.

Ved hjælp af Definition 8.1.2 ser vi da, at:

$$\begin{aligned} \det(\mathbf{A}) &= \lambda_1 \cdot \det(\mathbf{A}(1;1)) - a_{21} \cdot \det(\mathbf{A}(2;1)) + \dots + (-1)^{n+1} \cdot a_{n1} \cdot \det(\mathbf{A}(n;1)) \\ &= \lambda_1 \cdot \det(\mathbf{A}(1;1)) - a_{21} \cdot 0 + \dots + (-1)^{n+1} \cdot a_{n1} \cdot 0 \\ &= \lambda_1 \cdot \det(\mathbf{A}(1;1)) \\ &= \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n, \end{aligned}$$

hvor vi i den sidste lighed anvendte induktionshypotesen. Dette afslutter induktionstrinnet. Vi konkluderer ved induktionsprincippet, at sætningen er sand. □

8.2 Determinanter og elementære rækkeoperationer

Definition 8.1.2 er ikke altid den hurtigste måde at beregne determinanten af en kvadratisk matrix på. I vores studie af systemer af lineære ligninger definerede vi tre typer af elementære rækkeoperationer, der kunne benyttes til effektiv simplificering af et givet system. Motiveret af dette vil vi nu undersøge effekten af disse tre typer af elementære rækkeoperationer på værdien af en determinant. Den, der er nemmest at håndtere, er den elementære rækkeoperation af typen $R_i \leftarrow c \cdot R_i$. Vi starter med at bevise et generelt resultat.

Sætning 8.2.1

Betragt følgende tre matricer i $\mathbb{F}^{n \times n}$:

$$\mathbf{A} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_{i-1} & - \\ - & \mathbf{a}_i & - \\ - & \mathbf{a}_{i+1} & - \\ & \vdots & \\ - & \mathbf{a}_n & - \end{bmatrix}, \mathbf{B} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_{i-1} & - \\ - & \mathbf{b}_i & - \\ - & \mathbf{a}_{i+1} & - \\ & \vdots & \\ - & \mathbf{a}_n & - \end{bmatrix} \text{ og } \mathbf{C} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_{i-1} & - \\ - & c \cdot \mathbf{a}_i + \mathbf{b}_i & - \\ - & \mathbf{a}_{i+1} & - \\ & \vdots & \\ - & \mathbf{a}_n & - \end{bmatrix},$$

hvor $c \in \mathbb{F}$. Da gælder der, at $\det(\mathbf{C}) = c \cdot \det(\mathbf{A}) + \det(\mathbf{B})$.

Bevis. Vi benytter induktion på n. Hvis $n = 1$, har vi $\mathbf{A} = a$ for en $a \in \mathbb{F}$, $\mathbf{B} = b$ for en $b \in \mathbb{F}$ og $\mathbf{C} = c \cdot a + b$. Ifølge Definition 8.1.2 ser vi så, at $\det(\mathbf{C}) = c \cdot a + b = c \cdot \det(\mathbf{A}) + \det(\mathbf{B})$.

Antag nu, at $n \geq 2$, og at sætningen gælder for $(n-1) \times (n-1)$-matricer. Lad os ved $\sum_{k=1;k\neq i}^{n} (-1)^{k+1} \cdot c_{k1} \cdot \det(\mathbf{C}(k;1))$ betegne den summation, man opnår ved at lade k løbe fra 1 til n, hvor værdien i springes over. Da kan vi skrive

$$\det(\mathbf{C}) = \sum_{k=1;k\neq i}^{n} (-1)^{k+1} \cdot c_{k1} \cdot \det(\mathbf{C}(k;1)) + (-1)^{i+1} \cdot c_{i1} \cdot \det(\mathbf{C}(i;1)).$$

For alle k forskellige fra i medfører induktionshypotesen, at $\det(\mathbf{C}(k;1)) = c \cdot \det(\mathbf{A}(k;1)) + \det(\mathbf{B}(k;1))$. Ydermere er $\mathbf{C}(i;1) = \mathbf{A}(i;1) = \mathbf{B}(i;1)$, da den i'te række er den eneste række, hvor matricerne $\mathbf{A}$, $\mathbf{B}$ og $\mathbf{C}$ adskiller sig fra hinanden. Ved at udnytte, at $c_{i1} = c \cdot a_{i1} + b_{i1}$, og $c_{k1} = a_{k1}$, hvis $k \neq i$, ser vi nu, at

$$\begin{aligned} \det(\mathbf{C}) \quad = \quad & \sum_{k=1;k\neq i}^{n} (-1)^{k+1} \cdot a_{k1} \cdot \det(\mathbf{C}(k;1)) + \\ & (-1)^{i+1} \cdot (a_{i1} + b_{i1}) \cdot \det(\mathbf{C}(i;1)) \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=1;k\neq i}^{n} (-1)^{k+1} \cdot a_{k1} \cdot (c \cdot \det(\mathbf{A}(k;1)) + \det(\mathbf{B}(k;1))) \\
&\quad +(-1)^{i+1} \cdot c \cdot a_{i1} \cdot \det(\mathbf{A}(i;1)) + (-1)^{i+1} \cdot b_{i1} \cdot \det(\mathbf{B}(i;1)) \\
&= \sum_{k=1;k\neq i}^{n} c \cdot (-1)^{k+1} \cdot a_{k1} \cdot \det(\mathbf{A}(k;1)) + (-1)^{i+1} \cdot c \cdot a_{i1} \cdot \det(\mathbf{A}(i;1)) \\
&\quad + \sum_{k=1;k\neq i}^{n} (-1)^{k+1} \cdot b_{k1} \cdot \det(\mathbf{B}(k;1))) + (-1)^{i+1} \cdot b_{i1} \cdot \det(\mathbf{B}(i;1)) \\
&= c \cdot \det(\mathbf{A}) + \det(\mathbf{B}).
\end{aligned}$$

Dette afslutter induktionstrinnet og dermed induktionsbeviset. □

Korollar 8.2.2

Lad $\mathbf{A} \in \mathbb{F}^{n\times n}$ være givet, og antag, at $\mathbf{C}$ er opnået fra $\mathbf{A}$ ved anvendelse af den elementære rækkeoperation $R_i \leftarrow c \cdot R_i$ på $\mathbf{A}$ for et i og et $c \in \mathbb{F}$. Da gælder der, at $\det(\mathbf{C}) = c \cdot \det(\mathbf{A})$.

Bevis. Hvis vi vælger $\mathbf{b}_i = \mathbf{0}$ i Sætning 8.2.1, får vi, at $\det(\mathbf{C}) = c \cdot \det(\mathbf{A}) + \det(\mathbf{B})$, hvor $\mathbf{B}$ er en matrix, hvis i'te række er nulrækken. Sætning 8.1.3 medfører, at $\det(\mathbf{B}) = 0$. Deraf følger korollaret. □

Undersøgelsen af effekterne af de to resterende typer af elementære rækkeoperationer på værdien af determinanten viser sig mere indviklet. I opsummering bliver effekten af disse operationer følgende:

$$\text{Anvendelse af } R_i \leftrightarrow R_j \text{ på en kvadratisk matrix ændrer fortegnet på determinanten.} \tag{8.2}$$

$$\text{Anvendelse af } R_i \leftarrow R_i + c \cdot R_j \text{ på en kvadratisk matrix påvirker ikke determinanten.} \tag{8.3}$$

Eksempel 8.2.1

Lad $\mathbb{F} = \mathbb{R}$ og

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix}$$

som i Eksempel 7.3.2. Udregn determinanten af $\mathbf{A}$ ved hjælp af elementære rækkeoperationer.

Svar: Fra Eksempel 7.3.2 kan vi aflæse, at:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix} \underset{R_2 \leftarrow R_2 - 4 \cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 5 & 7 & 9 \end{bmatrix}$$

$$\underset{R_3 \leftarrow R_3 - 5 \cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -3 & -6 \end{bmatrix} \underset{R_3 \leftarrow R_3 - R_2}{\longrightarrow} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{bmatrix}.$$

Anvendes Ligning (8.3) tre gange, kan vi konkludere, at

$$\det(\mathbf{A}) = \det\left(\begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{bmatrix}\right).$$

Bemærk nu, at matricen på højre side er en øvre trekantsmatrix. Ved anvendelse af Sætning 8.1.2 får vi derfor, at

$$\det(\mathbf{A}) = \det\left(\begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{bmatrix}\right) = 1 \cdot (-3) \cdot 0 = 0.$$

Vi vil bruge resten af dette afsnit på at bevise gyldigheden af Ligningerne (8.2) og (8.3). En læser, der er villig til at acceptere deres gyldighed uden bevis, kan gå direkte videre til Afsnit 8.3. Læsere, der ønsker at se beviserne for Ligningerne (8.2) og (8.3), er velkomne til at læse videre, men det kan være en fordel at have læst Afsnit 8.3 først.

Vi starter med to lemmaer.

Lemma 8.2.3

Antag, at $n \geq 2$, og lad en kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ være givet. Desuden betegnes ved $\mathbf{B} \in \mathbb{F}^{n\times n}$ den matrix, der opnås fra $\mathbf{A}$ ved at bytte rundt på to på hinanden følgende rækker i $\mathbf{A}$. Da gælder der, at $\det(\mathbf{B}) = -\det(\mathbf{A})$.

Bevis. Vi beviser dette ved induktion på n.

Hvis $n = 2$, har vi

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ og } \mathbf{B} = \begin{bmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{bmatrix},$$

hvilket medfører $\det(\mathbf{A}) = a_{11} \cdot a_{22} - a_{21} \cdot a_{12}$ og $\det(\mathbf{B}) = a_{21} \cdot a_{12} - a_{11} \cdot a_{22}$. Derfor har vi, at $\det(\mathbf{B}) = -\det(\mathbf{A})$.

Lad nu $n \geq 3$, og antag, at lemmaet gælder for $n-1$. Lad os betegne de to rækker i $\mathbf{A}$, som har byttet plads, ved R_i og R_{i+1}. Da ser vi, at $\mathbf{A}(i;1) = \mathbf{B}(i+1;1)$, og $\mathbf{A}(i+1;1) = \mathbf{B}(i;1)$. For $k \neq i$ og $k \neq i+1$ kan $\mathbf{B}(k;1)$ desuden opnås fra $\mathbf{A}(k;1)$ ved, at to på hinanden følgende rækker byttes rundt. Ifølge induktionshypotesen har vi derfor for sådanne k, at $\det(\mathbf{B}(k;1)) = -\det(\mathbf{A}(k;1))$. Samles det hele, får vi:

$$\begin{aligned}
\det(\mathbf{B}) &= \sum_{k=1;k\neq i;k\neq i+1}^{n} (-1)^{k+1}\cdot a_{k1}\cdot \det(\mathbf{B}(k;1)) \\
&\quad +(-1)^{i+1}\cdot a_{i+1\,1}\cdot \det(\mathbf{B}(i;1)) + (-1)^{i+2}\cdot a_{i1}\cdot \det(\mathbf{B}(i+1;1)) \\
&= -\sum_{k=1;k\neq i;k\neq i+1}^{n} (-1)^{k+1}\cdot a_{k1}\cdot \det(\mathbf{A}(k;1)) \\
&\quad +(-1)^{i+1}\cdot a_{i+1\,1}\cdot \det(\mathbf{A}(i+1;1)) + (-1)^{i+2}\cdot a_{i1}\cdot \det(\mathbf{A}(i;1)) \\
&= -\sum_{k=1}^{n}(-1)^{k+1}\cdot a_{k1}\cdot \det(\mathbf{A}(k;1)) \\
&= -\det(\mathbf{A}).
\end{aligned}$$

Dette afslutter induktionstrinnet og dermed beviset. □

Lemma 8.2.4

Antag, at $n \geq 2$, og lad en kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ være givet. Antag, at to på hinanden følgende rækker i $\mathbf{A}$ er identiske. Da gælder der, at $\det(\mathbf{A}) = 0$.

Bevis. Dette kan vises ved at følge den samme strategi som i beviset for Lemma 8.2.3. □

Ovenstående lemma er blot et særtilfælde af et mere generelt resultat:

Proposition 8.2.5

Antag, at $n \geq 2$, og lad en kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ være givet. Antag, at to rækker i $\mathbf{A}$ er identiske. Da er $\det(\mathbf{A}) = 0$.

Bevis. Hvis to på hinanden følgende rækker i $\mathbf{A}$ er identiske, medfører Lemma 8.2.4, at $\det(\mathbf{A}) = 0$. Derfor står vi tilbage med tilfældet, hvor to rækker i $\mathbf{A}$ er identiske men ikke på hinanden følgende. Lad os betegne de to givne identiske rækker i $\mathbf{A}$ ved R_i og R_j, hvor $i > j \geq 1$. Vi bytter rundt på rækkerne R_i og R_{i-1}, og flytter dermed rækken R_i højere op i matricen. Effekten på determinanten er et fortegnsskift ifølge Lemma 8.2.3. I den nye matrix er de identiske rækker nu R_j og R_{i-1}. Hvis disse rækker er på hinanden følgende, stopper vi med at bytte rækker, men ellers flytter vi den nederste af de to identiske rækker op, én række ad gangen. Således ender vi på et tidspunkt med en matrix $\mathbf{B}$ med to på hinanden følgende rækker. Jævnfør Lemma 8.2.3 skifter determinanten fortegn, $\det(\mathbf{B}) = \pm\det(\mathbf{A})$, hver gang vi bytter rundt på to på hinanden følgende rækker. Men da vi ifølge Lemma 8.2.4 har $\det(\mathbf{B}) = 0$, konkluderer vi, at $\det(\mathbf{A}) = 0$. □

Vi har nu alle de nødvendige ingredienser til at vise effekten på determinanten af en kvadratisk matrix af at bytte om på to rækker.

Sætning 8.2.6

Lad en kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$ være givet, og betegn ved $\mathbf{B} \in \mathbb{F}^{n \times n}$ en matrix opnået fra $\mathbf{A}$ ved hjælp af en elementær rækkeoperation af formen $R_i \leftrightarrow R_j$ for nogle heltal $i < j$. Da gælder der, at $\det(\mathbf{B}) = -\det(\mathbf{A})$.

Bevis. Vi opskriver

$$\mathbf{A} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_{i-1} & - \\ - & \mathbf{a}_i & - \\ - & \mathbf{a}_{i+1} & - \\ & \vdots & \\ - & \mathbf{a}_{j-1} & - \\ - & \mathbf{a}_j & - \\ - & \mathbf{a}_{j+1} & - \\ & \vdots & \\ - & \mathbf{a}_n & - \end{bmatrix} \text{ og } \mathbf{C} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_{i-1} & - \\ - & \mathbf{a}_i + \mathbf{a}_j & - \\ - & \mathbf{a}_{i+1} & - \\ & \vdots & \\ - & \mathbf{a}_{j-1} & - \\ - & \mathbf{a}_j + \mathbf{a}_i & - \\ - & \mathbf{a}_{j+1} & - \\ & \vdots & \\ - & \mathbf{a}_n & - \end{bmatrix}.$$

Ved at anvende Sætning 8.2.1 på række i i $\mathbf{C}$, ser vi, at

$$\det(\mathbf{C}) = \det\left(\begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_{i-1} & - \\ - & \mathbf{a}_i & - \\ - & \mathbf{a}_{i+1} & - \\ & \vdots & \\ - & \mathbf{a}_{j-1} & - \\ - & \mathbf{a}_j + \mathbf{a}_i & - \\ - & \mathbf{a}_{j+1} & - \\ & \vdots & \\ - & \mathbf{a}_n & - \end{bmatrix}\right) + \det\left(\begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_{i-1} & - \\ - & \mathbf{a}_j & - \\ - & \mathbf{a}_{i+1} & - \\ & \vdots & \\ - & \mathbf{a}_{j-1} & - \\ - & \mathbf{a}_j + \mathbf{a}_i & - \\ - & \mathbf{a}_{j+1} & - \\ & \vdots & \\ - & \mathbf{a}_n & - \end{bmatrix}\right).$$

Nu anvender vi igen Sætning 8.2.1, men denne gang på række j i matricerne i de to determinantudtryk på højre side af denne ligning og bruger derefter Proposition 8.2.5. Så får vi, at

$$\det(\mathbf{C}) = \det(\mathbf{A}) + \det(\mathbf{B}).$$

Men Proposition 8.2.5 medfører, at $\det(\mathbf{C}) = 0$, da rækkerne i og j i $\mathbf{C}$ er identiske. Derfor gælder $0 = \det(\mathbf{A}) + \det(\mathbf{B})$, hvilket medfører, hvad vi ønskede at vise. $\square$

Nu hvor vi kender effekten af de elementære rækkeoperationer $R_i \leftarrow c \cdot R_i$ og $R_i \leftrightarrow R_j$ på determinanten, afslutter vi med et blik på, hvad der sker med determinanten, når vi benytter en elementær rækkeoperation af formen $R_i \leftarrow R_i + c \cdot R_j$.

Sætning 8.2.7

Lad $\mathbf{A} \in \mathbb{F}^{n\times n}$ være givet, og antag, at matricen $\mathbf{B}$ er opnået fra $\mathbf{A}$ ved anvendelse af den elementære rækkeoperation $R_i \leftarrow R_i + c \cdot R_j$ på $\mathbf{A}$ for forskellige rækkeindekser i, j og for et $c \in \mathbb{F}$. Da gælder der, at $\det(\mathbf{B}) = \det(\mathbf{A})$.

Bevis. Dette følger af Sætning 8.2.1 og Proposition 8.2.5. □

8.3 Alternative beskrivelser af determinanten

I vores beskrivelse af determinanten af en kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ spillede den første søjle i $\mathbf{A}$ en særlig rolle. I den rekursive definition gangede vi trods alt elementer fra den første søjle i $\mathbf{A}$ med determinanterne af mindre matricer. Disse mindre matricer blev opnået fra $\mathbf{A}$ ved at fjerne den første søjle og nogle rækker. Af den grund siger man typisk, at man i Definition 8.1.2 beregner determinanten ved at udvikle langs den første søjle. Mere præcist refererer man ofte til dette som *udviklingen* eller *Laplace-udviklingen* af determinanten langs den første søjle.

Man kan nu spørge, om der er nogen grund til, at den første søjle er så speciel. Svaret er: det er den ikke! Det er muligt at udregne determinanter ved udvkling langs andre søjler og faktisk også ved udvikling langs rækker. Mere præcist har vi følgende sætning:

Sætning 8.3.1

Lad $n \geq 2$, og lad $\mathbf{A} \in \mathbb{F}^{n\times n}$ være en kvadratisk matrix. Da gælder der for et ethvert j mellem 1 og n:

$$\det(\mathbf{A}) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(\mathbf{A}(i;j)). \tag{8.4}$$

Yderligere gælder der for ethvert i mellem 1 og n:

$$\det(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(\mathbf{A}(i;j)). \tag{8.5}$$

Bevis. Vi vil ikke bevise denne sætning, men den interesserede læser kan finde nogle bemærkninger i slutningen af dette afsnit, der forklarer de vigtigste ideer bag beviset. □

Bemærk, at for $j = 1$ er Ligning (8.4) blot den formel, der er givet for determinanten i Definition 8.1.2. Ligning (8.4) beskriver Laplace-udviklingen af determinanten langs den

j'te søjle, mens Ligning (8.5) beskriver Laplace-udviklingen af determinanten langs den i'te række. Disse ligninger kan også udtrykkes uden brug af summationssymbolet på følgende måde:

$$\det(\mathbf{A}) = (-1)^{1+j} \cdot a_{1j} \cdot \det(\mathbf{A}(1;j)) + (-1)^{2+j} \cdot a_{2j} \cdot \det(\mathbf{A}(2;j)) + \cdots + (-1)^{n+j} \cdot a_{nj} \cdot \det(\mathbf{A}(n;j)).$$

og

$$\det(\mathbf{A}) = (-1)^{i+1} \cdot a_{i1} \cdot \det(\mathbf{A}(i;1)) + (-1)^{i+2} \cdot a_{i2} \cdot \det(\mathbf{A}(i;2)) + \cdots + (-1)^{i+n} \cdot a_{in} \cdot \det(\mathbf{A}(i;n)).$$

Eksempel 8.3.1

Lad $\mathbb{F} = \mathbb{R}$ og

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix}$$

som i Eksempel 7.3.2. Udregn determinanten af **A** ved anvendelse af Laplace-udvikling langs den første række.

Svar: Bemærk først og fremmest, at

$$\mathbf{A}(1;1) = \begin{bmatrix} 5 & 6 \\ 7 & 9 \end{bmatrix}, \mathbf{A}(1;2) = \begin{bmatrix} 4 & 6 \\ 5 & 9 \end{bmatrix}, \text{ og } \mathbf{A}(1;3) = \begin{bmatrix} 4 & 5 \\ 5 & 7 \end{bmatrix}.$$

Derfor opnår vi ved anvendelse af Laplace-udvikling langs den første række, at

$$\det(\mathbf{A}) = (-1)^{1+1} \cdot 1 \cdot \det\left(\begin{bmatrix} 5 & 6 \\ 7 & 9 \end{bmatrix}\right) + (-1)^{1+2} \cdot 2 \cdot \det\left(\begin{bmatrix} 4 & 6 \\ 5 & 9 \end{bmatrix}\right) + (-1)^{1+3} \cdot 3 \cdot \det\left(\begin{bmatrix} 4 & 5 \\ 5 & 7 \end{bmatrix}\right).$$

Ved brug af Ligning (8.1) kan vi hurtigt beregne determinanterne af 2×2-matricer, og vi får:

$$\det(\mathbf{A}) = 1 \cdot (45 - 42) - 2 \cdot (36 - 30) + 3 \cdot (28 - 25) = 3 - 12 + 9 = 0.$$

Sætning 8.3.1 har en fin konsekvens i forbindelse med transponerede matricer.

Korollar 8.3.2

Lad $\mathbf{A} \in \mathbb{F}^{n \times n}$ være givet. Da gælder der, at $\det(\mathbf{A}) = \det(\mathbf{A}^T)$.

Bevis. Vi bruger induktion på n. Hvis $n = 1$, er $\mathbf{A} = \mathbf{A}^T$, og der gælder klart, at $\det(\mathbf{A}) = \det(\mathbf{A}^T)$. Antag nu at $n \geq 2$, og at korollaret gælder for $n - 1$. Bemærk, at $\mathbf{A}(j;1)^T = \mathbf{A}^T(1;j)$, hvilket sammen med induktionshypotesen lader os udnytte, at $\det(\mathbf{A}^T(1;j)) = \det(\mathbf{A}(j;1)^T) = \det(\mathbf{A}(j;1))$. Ved anvendelse af Laplace-udvikling af determinanten af $\mathbf{A}^T$ langs den første række, få vi

$$\begin{aligned}\det(\mathbf{A}^T) &= \sum_{j=1}^{n}(-1)^{1+j}\cdot(\mathbf{A}^T)_{1j}\cdot\det(\mathbf{A}^T(1;j))\\ &= \sum_{j=1}^{n}(-1)^{1+j}\cdot a_{j1}\cdot\det(\mathbf{A}(j;1))\\ &= \det(\mathbf{A}),\end{aligned}$$

hvor vi i den sidste lighed anvendte Definition 8.1.2. Dette afslutter induktionstrinnet og dermed beviset. □

Afslutningsvis vil vi nævne en særligt vigtig egenskab ved determinanter, nemlig at de opfører sig fint med hensyn til matrixmultiplikation:

Sætning 8.3.3

Lad $\mathbf{A}, \mathbf{B} \in \mathbb{F}^{n\times n}$ være givet. Da er $\det(\mathbf{A}\cdot\mathbf{B}) = \det(\mathbf{A})\cdot\det(\mathbf{B})$.

Den interesserede læser kan finde en skitse af beviset i slutningen af dette afsnit, men det er ikke obligatorisk læsning. Selvom sætningen ser uskyldig ud, har den en række væsentlige konsekvenser, der alle vil vise sig vigtige for os senere. Vi formulerer dem her som en række korollarer.

Korollar 8.3.4

Lad $\mathbf{A} \in \mathbb{F}^{n\times n}$ være givet. Da har $\mathbf{A}$ en invers, hvis og kun hvis $\det(\mathbf{A}) \neq 0$.

Bevis. Hvis $\mathbf{A}$ har en invers $\mathbf{A}^{-1}$, så gælder der, at $\mathbf{A}\cdot\mathbf{A}^{-1} = \mathbf{I}_n$. Ved at anvende Sætning 8.3.3 ser vi, at $\det(\mathbf{A})\cdot\det(\mathbf{A}^{-1}) = \det(\mathbf{I}_n) = 1$. For den sidste lighed benyttede vi os af Proposition 8.1.1. Således gælder der, at $\det(\mathbf{A}) \neq 0$, da produktet $\det(\mathbf{A})\cdot\det(\mathbf{A}^{-1})$ ellers ville være nul.

Antag omvent, at $\det(\mathbf{A}) \neq 0$. Hvis vi transformerer $\mathbf{A}$ ved hjælp af en hvilken som helst sekvens af elementære rækkeoperationer til en matrix $\mathbf{B}$ på reduceret trappeform, medfører Korollar 8.2.2 og Sætningerne 8.2.6 og 8.2.7, at $\det(\mathbf{B}) = d\cdot\det(\mathbf{A})$ for en konstant $d \in \mathbb{F}$ forskellig fra nul. Derfor gælder der, at $\det(\mathbf{B}) \neq 0$. Dette betyder specifikt, at $\mathbf{B}$ ikke indeholder en nulrække, da dens determinant ellers ville være nul ifølge Lemma 8.1.3. Derfor må der gælde, at $\mathbf{B} = \mathbf{I}_n$, hvilket medfører, at $\mathbf{A}$ har rang n. Som set i Ligning (7.10) og Korollar 7.3.5 betyder dette, at $\mathbf{A}$ har en invers. □

Korollar 8.3.5

Lad $\mathbf{A} \in \mathbb{F}^{n\times n}$ være givet. Da er søjlerne i $\mathbf{A}$ lineært uafhængige, hvis og kun hvis $\det(\mathbf{A}) \neq 0$.

Bevis. Dette følger ved at kombinere Sætning 7.1.3, det forrige korollar og Korollar 7.3.5. □

Korollar 8.3.6

Lad $\mathbf{A} \in \mathbb{F}^{n \times n}$ være givet. Da er $\det(\mathbf{A}) \neq 0$, hvis og kun hvis det homogene system af lineære ligninger med koefficientmatricen $\mathbf{A}$ kun har nulvektoren som løsning.

Bevis. Dette følger ved at kombinere Korollarerne 6.4.5, 7.3.5 og 8.3.4. □

Vi vil ikke bevise Sætning 8.3.3 i detaljer, men læseren, der gerne vil vide mere, kan læse resten af dette afsnit og få et godt indtryk af, hvorfor denne sætning samt Sætning 8.3.1 er sande. Resten af dette afsnit kan springes over ved første læsning. Hvis en læser er villig til at acceptere udsagnene i Sætningerne 8.3.1 og 8.3.3 uden bevis, kan de fortsætte til næste kapitel.

Nøglen til at forstå, hvorfor Sætning 8.3.1 er sand, er følgende:

Lemma 8.3.7

Lad $f : \mathbb{F}^{n \times n} \to \mathbb{F}$ være en funktion, der opfylder de følgende to betingelser:

(i) $f(\mathbf{A}) = 0$ for alle kvadratiske matricer $\mathbf{A} \in \mathbb{F}^{n \times n}$, der har to identiske rækker.

(ii) For alle matricer $\mathbf{A}, \mathbf{B}$ og $\mathbf{C}$ som angivet i Sætning 8.2.1, gælder der, at $f(\mathbf{C}) = c \cdot f(\mathbf{A}) + f(\mathbf{B})$.

Da er $f(\mathbf{A}) = \det(\mathbf{A}) \cdot f(\mathbf{I}_n)$ for alle $\mathbf{A} \in \mathbb{F}^{n \times n}$.

Bevis. Vi skitserer kun beviset: De to betingelser, som f opfylder, er nok til at udlede præcis, hvordan værdien af f ændrer sig, når en matrix $\mathbf{A}$ ændres ved hjælp af en elementær rækkeoperation. Faktisk kan mange af beviserne i Afsnit 8.2 genbruges. De to givne betingelser er også nok til at udlede, at $f(\mathbf{A}) = 0$ for enhver $\mathbf{A}$, der har en nulrække. Resultatet er derefter, at f opfører sig nøjagtigt som determinanten under elementære rækkeoperationer, og at både f og determinanten har værdien nul for matricer med en nulrække.

Givet en hvilken som helst kvadratisk matrix $\mathbf{A}$ og en sekvens af elementære rækkeoperationer, der transformerer $\mathbf{A}$ til sin reducerede trappeform, som vi kan kalde $\mathbf{B}$, kan man efter transformationen sammenligne værdierne af f og determinanten under disse elementære rækkeoperationer. Resultatet er, at $f(\mathbf{A}) = d \cdot f(\mathbf{B})$ for en konstant $d \in \mathbb{F}$, samt at også $\det(\mathbf{A}) = d \cdot \det(\mathbf{B})$ for *den samme* konstant d. Hvis $\mathbf{A}$ har rang mindre end n, indeholder dens reducerede trappeform $\mathbf{B}$ en nulrække. Men så er $f(\mathbf{B}) = 0$, og $\det(\mathbf{B}) = 0$. Hvis $\mathbf{A}$ har rang n, så er $\mathbf{B} = \mathbf{I}_n$. Således har vi i dette tilfælde, at $f(\mathbf{A}) = d \cdot f(\mathbf{I}_n)$, mens $\det(\mathbf{A}) = d \cdot \det(\mathbf{I}_n) = d \cdot 1 = d$. I alle tilfælde ser vi, at $f(\mathbf{A}) = \det(\mathbf{A}) \cdot f(\mathbf{I}_n)$. □

Bemærk, at determinanten, som vi definerede den i Definition 8.1.2, opfylder de to betingelser fra Lemma 8.3.7, se Proposition 8.2.5 og Sætning 8.2.1. For at bevise, at Sætning 8.3.1 er gyldig, skal man vise, at funktionen f, man får ved udvikling af en determinant langs en række eller en søjle, altid har de egenskaber, der er nævnt i Lemma 8.3.7, og at $f(\mathbf{I}_n) = 1$. Dette kan langt hen ad vejen gøres på samme måde, som det vi gjorde for determinanten defineret i Definition 8.1.2.

Lad os slutteligt opridse en skitse af beviset for Sætning 8.3.3:

Bevis. For en skitse af beviset af Sætning 8.3.3 betragter vi funktionen $f : \mathbb{F}^{n\times n} \to \mathbb{F}$ defineret ved $f(\mathbf{A}) = \det(\mathbf{A} \cdot \mathbf{B})$ for en vilkårligt valgt $\mathbf{B} \in \mathbb{F}^{n\times n}$. Ved hjælp af Proposition 8.2.5 og Sætning 8.2.1 viser man først, at f opfylder betingelserne i Lemma 8.3.7. Man kan derefter konkludere, at $f(\mathbf{A}) = \det(\mathbf{A}) \cdot f(\mathbf{I}_n)$ for alle $\mathbf{A} \in \mathbb{F}^{n\times n}$. Men så er $\det(\mathbf{A} \cdot \mathbf{B}) = f(\mathbf{A}) = \det(\mathbf{A}) \cdot f(\mathbf{I}_n) = \det(\mathbf{A}) \cdot \det(\mathbf{B})$. I den sidste lighed udnyttede vi, at $\mathbf{I}_n \cdot \mathbf{B} = \mathbf{B}$. $\square$

Kapitel 9

Vektorrum

9.1 Definition af og eksempler på vektorrum

I de foregående kapitler har vi arbejdet med linearkombinationer af vektorer fra $\mathbb{F}^n$, hvor $\mathbb{F}$ er et tallegeme (typisk $\mathbb{F} = \mathbb{R}$, eller $\mathbb{F} = \mathbb{C}$). Vi har set, at elementer i $\mathbb{F}^n$ kan lægges sammen og ganges med skalarer, dvs. ganges med elementer fra $\mathbb{F}$. Det viser sig at være en stor fordel at tage et mere abstrakt synspunkt og beskrive nogle væsentlige egenskaber helt fra start af. Disse egenskaber opstilles som såkaldt aksiomer. Dette gjorde vi på tilsvarende vis, da vi definerede et legeme. Også dér blev flere egenskaber ved de reelle og komplekse tal opstillet som aksiomer for et sådant legeme. For vektorer og skalarer kan vi opstille følgende:

Definition 9.1.1

Et *vektorrum* over et legeme $\mathbb{F}$ er en mængde V af elementer kaldet *vektorer* med to operationer defineret, som opfylder otte aksiomer. Den første operation kaldes addition og betegnes ved symbolet $+$. Den tager som input to elementer $\mathbf{u}, \mathbf{v} \in V$ og returnerer en vektor i V skrevet som $\mathbf{u} + \mathbf{v}$. Den anden operation kaldes skalarmultiplikation og kan betegnes ved symbolet $\cdot$. Den tager som input et element $c \in \mathbb{F}$, i denne sammenhæng typisk kaldet en *skalar*, og en vektor $\mathbf{u} \in V$ og returnerer en vektor i V skrevet som $c \cdot \mathbf{u}$. De otte aksiomer, der skal være opfyldt, er følgende:

(i) $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ for alle $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$

(ii) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ for alle $\mathbf{u}, \mathbf{v} \in V$

(iii) Der eksisterer en vektor $\mathbf{0} \in V$ kaldet *nulvektoren*, således at $\mathbf{u} + \mathbf{0} = \mathbf{u}$ for alle $\mathbf{u} \in V$

(iv) For ethvert $\mathbf{u} \in V$ eksisterer der et element $-\mathbf{u} \in V$ kaldet den additivt inverse af $\mathbf{u}$, således at $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$

(v) $c \cdot (d \cdot \mathbf{u}) = (c \cdot d) \cdot \mathbf{u}$ for alle $\mathbf{u} \in V$ og alle $c, d \in \mathbb{F}$

(vi) $1 \cdot \mathbf{u} = \mathbf{u}$ for alle $\mathbf{u} \in V$

(vii) $c \cdot (\mathbf{u} + \mathbf{v}) = c \cdot \mathbf{u} + c \cdot \mathbf{v}$ for alle $\mathbf{u}, \mathbf{v} \in V$ og alle $c \in \mathbb{F}$

(viii) $(c + d) \cdot \mathbf{u} = c \cdot \mathbf{u} + d \cdot \mathbf{u}$ for alle $\mathbf{u} \in V$ og alle $c, d \in \mathbb{F}$

Bemærk i punkt 5 med formlen $(c \cdot d) \cdot \mathbf{u}$, at den første forekomst af symbolet $\cdot$ (altså i udregningen $c \cdot d$) angiver multiplikation i legemet $\mathbb{F}$, mens den næste $\cdot$ angiver skalarmultiplikation i vektorrummet V. Tilsvarende i punkt 8 med formlen $(c + d) \cdot \mathbf{u} = c \cdot \mathbf{u} + d \cdot \mathbf{u}$ angiver den første forekomst af symbolet $+$ addition i $\mathbb{F}$, mens den næste $+$ angiver addition i V.

Eksempel 9.1.1

Lad os betragte $V = \mathbb{F}^n$ sammen med addition og skalarmultiplikation som defineret tidligere i Ligningerne (7.1) og (7.2). Dette udgør et eksempel på et vektorrum. Vi kan verificere, at vi ganske rigtigt har at gøre med et vektorrum, ved at kontrollere, om de otte aksiomer for vektorrum fra Definition 9.1.1 er opfyldt. Bemærk, at fem af dem allerede blev nævnt i Sætning 7.1.1. Nulvektoren, som er påkrævet af det tredje aksiom, er simpelthen nulvektoren i $\mathbb{F}^n$, og den additivt inverse af en vektor som påkrævet af det fjerde aksiom gives som:

$$-\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} -v_1 \\ \vdots \\ -v_n \end{bmatrix}.$$

Herefter mangler vi kun at vise opfyldelse af kravet i det sjette aksiom, hvilket er nemt gjort, da

$$1 \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} 1 \cdot v_1 \\ \vdots \\ 1 \cdot v_n \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \text{ for alle } \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{F}^n.$$

Vi har hermed bekræftet, at $\mathbb{F}^n$ er et vektorrum over legemet $\mathbb{F}$.

Et vektorrum over legemet $\mathbb{R}$ kaldes typisk et *reelt vektorrum*. På samme måde kaldes et vektorrum over legemet $\mathbb{C}$ typisk et *komplekst vektorrum*. Vi har i de tidligere kapitler faktisk allerede mødt eksempler på vektorrum. Lad os gengive et par stykker.

Eksempel 9.1.2

Betragt mængden $\mathbb{C}$ af komplekse tal. Hvis vi sætter $\mathbb{F} = \mathbb{C}$ og $n = 1$ i Eksempel 9.1.1, konkluderer vi, at vi kan opfatte $\mathbb{C}$ som et vektorrum over sig selv. Vi kan ligeledes også opfatte $\mathbb{C}$ som et vektorrum over de reelle tal $\mathbb{R}$. Som addition $+$ vælger vi blot additionsdefinitionen af komplekse tal. Da vi kan gange to komplekse tal sammen, kan vi selvfølgelig også gange et reelt tal med et komplekst tal. Dette giver os det nødvendige skalarprodukt. At alle otte aksiomer fra Definition 9.1.1 er opfyldt, kan udledes fra

Sætningerne 3.2.2 og 3.2.3.

Eksempel 9.1.3

På samme måde som i Eksempel 9.1.2 kan man betragte mængden af reelle tal $\mathbb{R}$ som et vektorrum over sig selv, men også som et vektorrum over legemet af rationale tal $\mathbb{Q}$ (se Eksemplerne 2.1.4 og 6.1.1 for en beskrivelse af legemet $\mathbb{Q}$).

Eksempel 9.1.4

Betragt mængden $\mathbb{F}^{m\times n}$ af $m \times n$-matricer med elementer i et legeme $\mathbb{F}$. Ved at vælge matrixaddition som defineret i Definition 7.2.3 samt skalarmultiplikation som defineret ved:

$$c \cdot \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} c \cdot a_{11} & \cdots & c \cdot a_{1n} \\ \vdots & & \vdots \\ c \cdot a_{m1} & \cdots & c \cdot a_{mn} \end{bmatrix},$$

angiver de første to punkter i Sætning 7.2.1, at de første to vektorrumsaksiomer er opfyldt. Den $m \times n$-matrix, der kun indeholder nulelementer, fungerer som nulvektor. Alle andre aksiomer kan kontrolleres på tilsvarende vis, hvilket vi overlader til læseren.

Eksempel 9.1.5

Betragt mængden $\mathbb{C}[Z]$ af polynomier i variablen Z med koefficienter i $\mathbb{C}$ som defineret i Definition 4.1.1. På denne mængde har vi som addition $+$ den sædvanlige addition af polynomier. Vi kan gange to polynomier sammen, hvorfor vi selvfølgelig også kan gange et konstant polynomium med et andet polynomium. Dette giver os et skalarprodukt på $\mathbb{C}[Z]$. Vi vil ikke gøre det her, men man kan vise, at alle otte aksiomer fra Definition 9.1.1 er opfyldt. Således er $\mathbb{C}[Z]$ et vektorrum over $\mathbb{C}$.

Eksempel 9.1.6

Betragt mængden F af alle funktioner med definitionsmængden $\mathbb{R}$ og dispositionsmængden $\mathbb{R}$. Hvis $f : \mathbb{R} \to \mathbb{R}$ og $r \in \mathbb{R}$ er givet, kan man definere funktionen $r \cdot f : \mathbb{R} \to \mathbb{R}$ som $(r \cdot f)(a) = r \cdot f(a)$ for alle $a \in \mathbb{R}$. Dette definerer skalarmultiplikation på F. Addition på F defineres på en tilsvarende måde: hvis $f : \mathbb{R} \to \mathbb{R}$ og $g : \mathbb{R} \to \mathbb{R}$ er givet, defineres funktionen $(f + g) : \mathbb{R} \to \mathbb{R}$ som $(f + g)(a) = f(a) + g(a)$ for alle $a \in \mathbb{R}$. Man kan verificere, at dette giver F strukturen af et vektorrum over $\mathbb{R}$. Som nulvektor vælger man nulfunktionen: $\mathbf{0} : \mathbb{R} \to \mathbb{R}$, der opfylder $a \mapsto 0$ for alle $a \in \mathbb{R}$.

I alle de ovenfor nævnte eksempler gælder der, at produktet af skalaren 0 med en hvilken som helst vektor er lig med nulvektoren $\mathbf{0}$. Dog siger ingen af de otte vektorrumsaksiomer, at $0 \cdot \mathbf{u} = \mathbf{0}$ for alle $\mathbf{u} \in V$. Heldigvis er de otte vektorrumsaksiomer velvalgte: man kan udlede en del fra dem, for eksempel at formlen $0 \cdot \mathbf{u} = \mathbf{0}$ faktisk er sand for ethvert vektorrum. Vi beviser dette samt en anden intuitiv formel i følgende lemma:

Lemma 9.1.1

Lad V være et vektorrum. Da gælder der, at

$$0 \cdot \mathbf{u} = \mathbf{0} \text{ for alle } \mathbf{u} \in V, \tag{9.1}$$

og

$$(-1) \cdot \mathbf{u} = -\mathbf{u} \text{ for alle } \mathbf{u} \in V. \tag{9.2}$$

Bevis. Ved at bruge, at $0 = 0 + 0$ og vektorrumsaksiom otte, får vi, at $0 \cdot \mathbf{u} = (0+0) \cdot \mathbf{u} = 0 \cdot \mathbf{u} + 0 \cdot \mathbf{u}$. Ved at lægge $-(0 \cdot \mathbf{u})$ til på begge sider og benytte vektorrumsaksiomerne fire, et og tre, får vi

$$\begin{aligned}
\mathbf{0} &= 0 \cdot \mathbf{u} + (-(0 \cdot \mathbf{u})) \\
&= (0 \cdot \mathbf{u} + 0 \cdot \mathbf{u}) + (-(0 \cdot \mathbf{u})) \\
&= 0 \cdot \mathbf{u} + (0 \cdot \mathbf{u} + (-(0 \cdot \mathbf{u}))) \\
&= 0 \cdot \mathbf{u} + \mathbf{0} \\
&= 0 \cdot \mathbf{u}.
\end{aligned}$$

Dette viser den første del. Den anden del følger tilsvarende. Da $0 = (1 + (-1))$, får vi, at $0 \cdot \mathbf{u} = (1 + (-1)) \cdot \mathbf{u} = 1 \cdot \mathbf{u} + (-1) \cdot \mathbf{u}$. Venstre side af denne ligning er lig med $\mathbf{0}$ ifølge den første del af dette lemma. Benyttes dette sammen med vektorrumsaksiom seks, har vi, at $\mathbf{0} = \mathbf{u} + (-1) \cdot \mathbf{u}$. Derfor er $(-1) \cdot \mathbf{u} = -\mathbf{u}$. □

9.2 Basis for et vektorrum

Meget lig det vi gjorde i Afsnit 7.1 med vektorer i $\mathbb{F}^m$, kan man tale om en *linearkombination* af vektorer i rammerne af generelle vektorrum. Givet et vektorrum V over et legeme $\mathbb{F}$, vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ og skalarer $c_1, \ldots, c_n \in \mathbb{F}$, så kaldes et udtryk af formen

$$c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n$$

en linearkombination af vektorerne $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Ligeledes generaliseres begrebet lineær (u)afhængighed af en endelig sekvens af vektorer fra Definition 7.1.1 nemt til emnet om vektorrum:

Definition 9.2.1

Lad V være et vektorrum over et legeme $\mathbb{F}$. En sekvens af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ kaldes *lineært uafhængig*, hvis og kun hvis ligningen $c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n = \mathbf{0}$ med $c_1, \ldots, c_n \in \mathbb{F}$ kun er sand for $c_1 = \cdots = c_n = 0$.

Hvis sekvensen af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ ikke er lineært uafhængig, siges den at være *lineært afhængig*.

Grundlæggende er den eneste forskel fra Definition 7.1.1, at $\mathbb{F}^m$ her er blevet erstattet med V. Også i rammerne af generelle vektorrum er det almindeligt blot at sige, at vektorerne $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært (u)afhængige snarere end at sige, at sekvensen af vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ er lineært (u)afhængig.

Der opstår dog én komplikation vedrørende lineær uafhængighed af vektorer i generelle vektorrum. I Definition 9.2.1 betragtede vi kun *endeligt mange* vektorer. Det viser sig, at vi nogle gange gerne vil kunne sige, at vektorerne fra en muligvis uendelig mængde er lineært uafhængige. Følgende definition vil tillade os at gøre dette.

Definition 9.2.2

Lad V være et vektorrum over et legeme $\mathbb{F}$. Vektorerne i en mængde S af vektorer kaldes *lineært uafhængige*, hvis og kun hvis enhver endelig sekvens af forskellige vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_n$ fra S er en lineært uafhængig sekvens af vektorer.

Hvis vektorerne i S ikke er lineært uafhængige, siges de at være *lineært afhængige*.

Grundlæggende siger Definition 9.2.2, at selvom antallet af vektorer i en mængde S kan være uendeligt, så betragter vi kun endeligt mange af dem ad gangen, når vi vil afgøre, om de er lineært uafhængige. Ofte vil vi kun arbejde med endelige sekvenser af vektorer, hvor Definition 9.2.1 kan anvendes.

I Eksemplerne 7.1.2 og 7.1.3 har vi allerede givet eksempler på lineært afhængige og lineært uafhængige vektorer i vektorrummet $\mathbb{R}^2$. Lad os se på nogle flere eksempler.

Eksempel 9.2.1

I Eksempel 9.1.2 opfattede vi $\mathbb{C}$ som et vektorrum over $\mathbb{R}$. I dette eksempel ser vi nærmere på lineært afhængige og uafhængige elementer herfra. Lad os som det første betragte elementerne 1 og i. For at afgøre om disse er lineært uafhængige, betragter vi ligningen $c_1 \cdot 1 + c_2 \cdot i = 0$, hvor $c_1, c_2 \in \mathbb{R}$. Vi tillader kun, at c_1 og c_2 kan antage reelle værdier, da vi i dette eksempel opfatter $\mathbb{C}$ som et vektorrum over legemet $\mathbb{R}$. Derfor har vi i Definition 9.2.1 $V = \mathbb{C}$ og $\mathbb{F} = \mathbb{R}$, hvorfor skalarerne skal komme fra $\mathbb{R}$.

Fortsætter vi med ligningen $c_1 \cdot 1 + c_2 \cdot i = 0$, hvor $c_1, c_2 \in \mathbb{R}$, ser vi, at det komplekse tal $c_1 \cdot 1 + c_2 \cdot i$ er på rektangulær form. Da to komplekse tal er ens, hvis og kun hvis de har samme realdel og imaginærdel, medfører ligningen $c_1 \cdot 1 + c_2 \cdot i = 0$, at $c_1 = 0$, og $c_2 = 0$. Vi konkluderer, at de komplekse tal 1 og i er lineært uafhængige over $\mathbb{R}$.

På samme måde kan man vise, at de komplekse tal 2 og $1 + i$ er lineært uafhængige. Vi betragter her ligningen $c_1 \cdot 2 + c_2 \cdot (1 + i) = 0$ for $c_1, c_2 \in \mathbb{R}$. Sammenligninger vi realdelene og imaginærdelene som før, medfører dette, at $2c_1 + c_2 = 0$, og $c_2 = 0$, hvoraf det følger, at $c_1 = c_2 = 0$.

Lad os som et sidste eksempel betragte en sekvens af tre komplekse tal, for eksempel 2, $1 + i$ og $2 + 3i$. Da $-(1/2) \cdot 2 + 3 \cdot (1 + i) + (-1) \cdot (2 + 3i) = 0$, ser vi, at de tre komplekse tal $2, 1 + i$ og $2 + 3i$ er lineært afhængige over $\mathbb{R}$.

Eksempel 9.2.2

I Eksempel 9.1.4 opfattede vi mængden af matricer $\mathbb{F}^{m\times n}$ som et vektorrum over $\mathbb{F}$. For ethvert talpar (i,j), der opfylder $1 \le i \le m$ og $1 \le j \le n$, definerer vi matricen $\mathbf{E}^{(i,j)} \in \mathbb{F}^{m\times n}$ som den matrix, der har nulelementer overalt undtagen på position (i,j), hvor den har et ettal. For eksempel har vi med $m = n = 2$ følgende:

$$\mathbf{E}^{(1,1)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \mathbf{E}^{(1,2)} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \mathbf{E}^{(2,1)} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \text{ og } \mathbf{E}^{(2,2)} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Med $m = n = 2$ har vi, at matricerne $\mathbf{E}^{(1,1)}, \mathbf{E}^{(1,2)}, \mathbf{E}^{(2,1)}, \mathbf{E}^{(2,2)}$ er lineært uafhængige over $\mathbb{F}$. Der gælder nemlig for alle $c_1, c_2, c_3, c_4 \in \mathbb{F}$, at

$$c_1 \cdot \mathbf{E}^{(1,1)} + c_2 \cdot \mathbf{E}^{(1,2)} + c_3 \cdot \mathbf{E}^{(2,1)} + c_4 \cdot \mathbf{E}^{(2,2)} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}.$$

Derfor medfører $c_1 \cdot \mathbf{E}^{(1,1)} + c_2 \cdot \mathbf{E}^{(1,2)} + c_3 \cdot \mathbf{E}^{(2,1)} + c_4 \cdot \mathbf{E}^{(2,2)} = \mathbf{0}$, at $c_1 = c_2 = c_3 = c_4 = 0$.

For generelle værdier af m og n kan man tilsvarende vise, at $m \times n$-matricerne $\mathbf{E}^{(1,1)}, \ldots, \mathbf{E}^{(m,n)}$ er lineært uafhængige over $\mathbb{F}$.

For $m = n = 2$ er et eksempel på en sekvens af lineært afhængige matricer følgende:

$$\begin{bmatrix} -1 & 0 \\ 2 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \text{ og } \begin{bmatrix} 5 & 4 \\ 2 & 0 \end{bmatrix},$$

da

$$1 \cdot \begin{bmatrix} -1 & 0 \\ 2 & 4 \end{bmatrix} - 4 \cdot \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 5 & 4 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Eksempel 9.2.3

Betragt det komplekse vektorrum $\mathbb{C}[Z]$ fra Eksempel 9.1.5. Husk på, at to polynomier, $p_1(Z) = a_0 + a_1 Z + \cdots + a_n Z^n$ af grad n og $p_2(Z) = b_0 + b_1 Z + \cdots + b_m Z^m$ af grad m, er ens, hvis og kun hvis $n = m$, og $a_i = b_i$ for alle i. Dette medfører særligt, at et polynomium $p(Z) = c_0 + c_1 Z + \cdots + c_n Z^n$ er lig med nulpolynomiet, hvis og kun hvis $c_i = 0$ for alle i. Dette viser, at mængden $\{1, Z, Z^2, \ldots\}$ er en mængde af lineært uafhængige polynomier over $\mathbb{C}$.

Alle disse eksempler viser, at begrebet lineær uafhængighed overføres fint til den generelle ramme af vektorrum. Med dette på plads er vi nået til et særdeles vigtigt begreb i teorien om vektorrum.

Definition 9.2.3

Lad V være et vektorrum over et legeme $\mathbb{F}$. En mængde S af vektorer i V kaldes en *basis* for V, hvis følgende to betingelser er opfyldt:

(i) Vektorerne i S er lineært uafhængige.

(ii) Enhver $\mathbf{v} \in V$ kan skrives som en linearkombination af vektorerne i S.

En *ordnet basis* $(\mathbf{v}_1, \mathbf{v}_2, \dots)$ for V er en liste af vektorer i V, således at mængden $\{\mathbf{v}_1, \mathbf{v}_2, \dots\}$ udgør en basis for V.

Det viser sig, at ethvert vektorrum har en basis, og vi vil frit anvende dette faktum. En læser, der har tid og lyst til lidt ekstra materiale om dette, henvises til Afsnit 9.4, men dette er ikke obligatorisk læsning. Hvis et vektorrum har en endelig basis, det vil sige, hvis mængden S, der indeholder basisvektorerne, er endelig, kan vi nummerere elementerne i S og skrive $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. Da er $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ en endelig, ordnet basis for V. Derfor har ethvert vektorrum med en endelig basis en ordnet basis.

Før vi ser på eksempler, opstiller vi her et lemma samt endnu en definition.

Lemma 9.2.1

Lad V være et vektorrum over et legeme $\mathbb{F}$, der har en endelig, ordnet basis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$. Da kan enhver vektor $\mathbf{v} \in V$ på præcis én måde skrives som en linearkombination af basisvektorerne.

Bevis. Det andet punkt i Definition 9.2.3 garanterer, at enhver vektor $\mathbf{v} \in V$ kan skrives som en linearkombination af basisvektorerne, altså $\mathbf{v} = c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n$ for visse $c_1, \dots, c_n \in \mathbb{F}$. Hvad vi skal vise, er, at dette er den eneste måde, hvorpå man kan skrive $\mathbf{v}$ som en linearkombination af basisvektorerne $\mathbf{v}_1, \dots, \mathbf{v}_n$. Antag derfor, at $\mathbf{v} = d_1 \cdot \mathbf{v}_1 + \cdots d_n \cdot \mathbf{v}_n$ for visse $d_1, \dots, d_n \in \mathbb{F}$. Vi ønsker at vise, at $c_1 = d_1, \dots, c_n = d_n$. Først og fremmest har vi

$$c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n = \mathbf{v} = d_1 \cdot \mathbf{v}_1 + \cdots d_n \cdot \mathbf{v}_n.$$

Derfor er

$$c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n - (d_1 \cdot \mathbf{v}_1 + \cdots d_n \cdot \mathbf{v}_n) = \mathbf{0},$$

hvilket igen medfører, at

$$(c_1 - d_1) \cdot \mathbf{v}_1 + \cdots (c_n - d_n) \cdot \mathbf{v}_n = \mathbf{0}.$$

Men da vektorerne $\mathbf{v}_1, \dots, \mathbf{v}_n$ er lineært uafhængige (dette følger af første del af Definition 9.2.3), ser vi, at $c_1 - d_1 = 0, \dots, c_n - d_n = 0$. Men så er $c_1 = d_1, \dots, c_n = d_n$, hvilket var det, vi ønskede at vise. □

Lemma 9.2.1 motiverer for følgende definition:

Definition 9.2.4

Lad V være et vektorrum over et legeme $\mathbb{F}$, der har en endelig, ordnet basis $\beta = (\mathbf{v}_1, \dots, \mathbf{v}_n)$. Hvis vi for $\mathbf{v} \in V$ har

$$\mathbf{v} = c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n,$$

så definerer vi

$$[\mathbf{v}]_\beta = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{F}^n$$

som *koordinatvektoren* for $\mathbf{v}$ med hensyn til den ordnede basis β. Man siger også, at $[\mathbf{v}]_\beta$ er *β-koordinatvektoren* for $\mathbf{v}$.

Funktionen, der sender en vektor fra V til dens β-koordinatvektor, har flere gode egenskaber. Særligt vil de følgende to vise sig nyttige senere hen.

Lemma 9.2.2

Lad V være et vektorrum over et legeme $\mathbb{F}$, der har en endelig, ordnet basis β. Da gælder der, at:

$$[\mathbf{u}+\mathbf{v}]_\beta = [\mathbf{u}]_\beta + [\mathbf{v}]_\beta \text{ for alle } \mathbf{u},\mathbf{v} \in V$$

og

$$[c \cdot \mathbf{v}]_\beta = c \cdot [\mathbf{v}]_\beta \text{ for alle } c \in \mathbb{F} \text{ og } \mathbf{v} \in V.$$

Bevis. Vi beviser kun den første del og overlader beviset for den anden del til læseren. Lad os sige, at den ordnede basis β er givet ved $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Hvis $\mathbf{u} = c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n$, og $\mathbf{v} = d_1 \cdot \mathbf{v}_1 + \cdots d_n \cdot \mathbf{v}_n$, så er $\mathbf{u}+\mathbf{v} = (c_1+d_1)\cdot \mathbf{v}_1 + \cdots (c_n+d_n)\cdot \mathbf{v}_n$. Derfor har vi, at

$$[\mathbf{u}+\mathbf{v}]_\beta = \begin{bmatrix} c_1+d_1 \\ \vdots \\ c_n+d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = [\mathbf{u}]_\beta + [\mathbf{v}]_\beta.$$

□

Nu vil vi benytte dette lemma til at bevise et sætning om lineær uafhængighed af vektorer.

Sætning 9.2.3

Lad V være et vektorrum over et legeme $\mathbb{F}$, der har en endelig, ordnet basis β bestående af n vektorer. Antag, at vi har $\mathbf{u}_1, \ldots, \mathbf{u}_\ell \in V$ og $c_1, \ldots, c_\ell \in \mathbb{F}$. Da gælder der, at:

$$c_1 \cdot \mathbf{u}_1 + \cdots + c_\ell \cdot \mathbf{u}_\ell = \mathbf{0}, \text{ hvis og kun hvis } c_1 \cdot [\mathbf{u}_1]_\beta + \cdots + c_\ell \cdot [\mathbf{u}_\ell]_\beta = \mathbf{0}.$$

Særligt er vektorerne $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$ i V lineært uafhængige, hvis og kun hvis vektorerne $[\mathbf{u}_1]_\beta, \ldots, [\mathbf{u}_\ell]_\beta$ i $\mathbb{F}^n$ er lineært uafhængige.

Bevis. En vektor $\mathbf{v}$ i V er nulvektoren, hvis og kun hvis dens β-koordinatvektor er nulvektoren. Derfor gælder $c_1 \cdot \mathbf{u}_1 + \cdots + c_\ell \cdot \mathbf{u}_\ell = \mathbf{0}$, hvis og kun hvis $[c_1 \cdot \mathbf{u}_1 + \cdots + c_\ell \cdot \mathbf{u}_\ell]_\beta = \mathbf{0}$. Ved at anvende Lemma 9.2.2 gentagne gange kan vi også udlede, at $[c_1 \cdot \mathbf{u}_1 + \cdots + c_\ell \cdot \mathbf{u}_\ell]_\beta =$

$c_1 \cdot [\mathbf{u}_1]_\beta + \cdots + c_\ell \cdot [\mathbf{u}_\ell]_\beta$. Dermed følger den første del af sætningen. Den anden del følger direkte fra den første del. □

Denne sætning reducerer essentielt set spørgsmålet om lineær (u)afhængighed af vektorer i V til et spørgsmål om lineær (u)afhængighed af vektorer i $\mathbb{F}^n$. Til arbejde med $\mathbb{F}^n$ har vi allerede teknikker til rådighed, især Sætning 7.1.3.

Eksempel 9.2.4

Lad $\mathbb{F} = \mathbb{R}$ og $V = \mathbb{R}^2$. Vi påstår, at vektorerne

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ og } \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

udgør en ordnet basis β for $\mathbb{R}^2$. Disse vektorer er lineært uafhængige (læseren opfordres til at kontrollere dette), og enhver vektor i $\mathbb{R}^2$ kan skrives som en linearkombination af $\mathbf{e}_1$ og $\mathbf{e}_2$, da

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = v_1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + v_2 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Det betyder, at vi i dette tilfælde har $[\mathbf{v}]_\beta = \mathbf{v}$.

Lad nu γ være sekvensen af vektorerne

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ og } \mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \in \mathbb{R}^2.$$

Vi har set i Eksempel 7.1.3, at disse to vektorer er lineært uafhængige. Desuden kan man vise, at enhver vektor i $\mathbb{R}^2$ kan skrives som en linearkombination af $\mathbf{u}$ og $\mathbf{v}$. Givet $v_1, v_2 \in \mathbb{R}$ fører ligningen

$$c_1 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

til et system af to lineære ligninger i variablerne c_1 og c_2. Ved at løse dette system kan man vise, at vi for enhver $v_1, v_2 \in \mathbb{R}$, har

$$c_1 = -\frac{v_1}{3} + \frac{2v_2}{3} \quad \text{og} \quad c_2 = \frac{2v_1}{3} - \frac{v_2}{3},$$

således at

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \left(-\frac{v_1}{3} + \frac{2v_2}{3}\right) \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \left(\frac{2v_1}{3} - \frac{v_2}{3}\right) \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Dette betyder, at $\gamma = (\mathbf{u}, \mathbf{v})$ er en ordnet basis for $\mathbb{R}^2$. Desuden ser vi fra ovenstående, at

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}_\gamma = \begin{bmatrix} -v_1/3 + 2v_2/3 \\ 2v_1/3 - v_2/3 \end{bmatrix}.$$

Den første del af Eksempel 9.2.4 kan udvides yderligere: lad os som i Afsnit 7.3 betegne den i'te søjle i identitetsmatricen $\mathbf{I}_n \in \mathbb{F}^{n\times n}$ ved $\mathbf{e}_i$ for $i = 1,\dots,n$. Med andre ord: vektoren $\mathbf{e}_i$ har 1 som sin i'te koordinat og nul alle andre steder. Disse vektorer udgør en ordnet basis $(\mathbf{e}_1,\dots,\mathbf{e}_n)$ for vektorrummet $\mathbb{F}^n$, kaldet en *(ordnet) standardbasis*. Lad os for fuldstændighedens skyld vise, at de udgør en ordnet basis:

Proposition 9.2.4

Vektorerne $\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n$ udgør en ordnet basis for vektorrummet $\mathbb{F}^n$ over $\mathbb{F}$.

Bevis. Ifølge Definition 9.2.3 skal vi kontrollere to ting:

(i) Vektorerne $\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n$ er lineært uafhængige.

(ii) Enhver vektor i $\mathbb{F}^n$ kan skrives som en linearkombination af $\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n$.

Det første punkt følger fra observationen af, at

$$c_1\cdot\mathbf{e}_1 + c_2\cdot\mathbf{e}_2 + \cdots + c_n\cdot\mathbf{e}_n = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Dette betyder, at hvis en linearkombination er lig med nulvektoren i $\mathbb{F}^n$, så er alle skalarer $c_1,\dots,c_n$ nul. Det andet punkt følger, da vi, hvis $\mathbf{v} = (v_1,\dots,v_n) \in \mathbb{F}^n$ er givet, har

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = v_1\cdot\mathbf{e}_1 + v_2\cdot\mathbf{e}_2 + \cdots + v_n\cdot\mathbf{e}_n.$$

□

På samme måde som i Eksempel 9.2.4 gælder der, at hvis β er den ordnede standardbasis for $\mathbb{F}^n$, så er $[\mathbf{v}]_\beta = \mathbf{v}$ for alle $\mathbf{v} \in \mathbb{F}^n$. Bemærk dog, at ligesom i Eksempel 9.2.4 har vektorrummet $\mathbb{F}^n$ mange andre mulige ordnede baser. Lad os fortsætte med nogle eksempler på baser for vektorrum.

Eksempel 9.2.5

Vi fortsætter fra Eksemplerne 9.1.2 og 9.2.1, hvorfra vi ved, at de komplekse tal 1 og i er lineært uafhængige over $\mathbb{R}$. De udgør faktisk en ordnet basis $(1,i)$, som vi betegner ved β, da ethvert komplekst tal er en linearkombination af 1 og i over de reelle tal. Mere specifikt har vi for ethvert $a, b \in \mathbb{R}$, at $a + bi = a \cdot 1 + b \cdot i$. Derfor har vi for $a, b \in \mathbb{R}$, at

$$[a + bi]_\beta = \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2.$$

Dermed er $[a + bi]_\beta$ lig med de rektangulære koordinater for det komplekse tal $a + bi$.

Der er mange flere mulige baser (og dermed ordnede baser) for $\mathbb{C}$, når den betragtes som et vektorrum over $\mathbb{R}$. For eksempel er $(2, 1 + i)$ en mulig ordnet basis. Vi har nemlig allerede vist i Eksempel 9.2.1, at de komplekse tal 2 og $1 + i$ er lineært uafhængige over $\mathbb{R}$. Derfor kan ethvert komplekst tal skrives som en linearkombination med koefficienter i $\mathbb{R}$ af 2 og $1 + i$. For at indse dette kan vi verificere, at ligningen $a + bi = c_1 \cdot 2 + c_2 \cdot (1 + i)$ for et givet komplekst tal $a + bi$, hvor $a, b \in \mathbb{R}$, har en løsning $c_1, c_2 \in \mathbb{R}$. En sammenligning af realdelene og imaginærdelene viser, at $a = 2c_1 + c_2$, og $b = c_2$. Derfor har vi som løsning $c_2 = b$ og $c_1 = (a - c_2)/2 = (a - b)/2$. Betegnes den ordnede basis $(2, 1 + i)$ ved γ, har vi

$$[a + bi]_\gamma = \begin{bmatrix} (a - b)/2 \\ b \end{bmatrix} \in \mathbb{R}^2.$$

Eksempel 9.2.6

Vi fortsætter her fra Eksemplerne 9.1.4 og 9.2.2 og finder en ordnet basis β for vektorrummet $\mathbb{F}^{m \times n}$ over $\mathbb{F}$. En sådan ordnet basis er $(\mathbf{E}^{(1,1)}, \ldots, \mathbf{E}^{(m,n)})$. Vi har allerede set, at matricerne $\mathbf{E}^{(1,1)}, \ldots, \mathbf{E}^{(m,n)}$ er lineært uafhængige, og at enhver matrix $\mathbf{A} = (a_{ij})_{1 \le i \le m; 1 \le j \le n}$ kan skrives som en linearkombination af dem, nemlig som $\mathbf{A} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \mathbf{E}^{(i,j)}$.

Specifikt for $m = n = 2$ udgør matricerne $\mathbf{E}^{(1,1)}, \mathbf{E}^{(1,2)}, \mathbf{E}^{(2,1)}, \mathbf{E}^{(2,2)}$ en ordnet basis $\beta = (\mathbf{E}^{(1,1)}, \mathbf{E}^{(1,2)}, \mathbf{E}^{(2,1)}, \mathbf{E}^{(2,2)})$, og vi har

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}_\beta = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{bmatrix}.$$

Eksempel 9.2.7

I dette eksempel ser vi igen på det komplekse vektorrum $\mathbb{C}[Z]$ fra Eksemplerne 9.1.5 og 9.2.3. Fra disse eksempler ved vi allerede, at mængden $\{1, Z, Z^2, \ldots\}$ er en mængde af lineært uafhængige polynomer over $\mathbb{C}$. Ifølge definitionen af polynomer er ethvert polynomium en linearkombination over $\mathbb{C}$ af et endeligt antal elementer fra denne mængde. Derfor er mængden $\{1, Z, Z^2, \ldots\}$ faktisk en basis for det komplekse vektorrum $\mathbb{C}[Z]$. Dette er et eksempel på et vektorrum, der har en uendelig basis.

Det viser sig, at antallet af vektorer i en basis for et givet vektorrum V over et legeme $\mathbb{F}$ altid er det samme. Senere i dette afsnit vil vi bevise dette i særtilfældet, hvor antallet af vektorer i en basis er endeligt. Generelt kaldes antallet af elementer i en basis for V for *dimensionen* af vektorrummet V. Almindelig notation for dimensionen af et vektorrum V er: $\dim(V)$ eller blot $\dim V$. Hvis man ønsker at præcisere, over hvilket legeme $\mathbb{F}$ vektorrummet er defineret, skriver man $\dim_{\mathbb{F}}(V)$ eller $\dim_{\mathbb{F}} V$. Hvis antallet af vektorer i en basis er endeligt, siges V at have en endelig dimension, og ellers siges V at have en uendelig dimension, hvilket også kan udtrykkes som: $\dim V = \infty$.

Eksempel 9.2.8

Lad os udregne dimensionerne af forskellige eksempler på vektorrum, som vi har mødt indtil nu. Først og fremmest fra Eksempel 9.2.4 ser vi, at $\dim_{\mathbb{R}}(\mathbb{R}^2) = 2$. Mere generelt har man $\dim_{\mathbb{F}}(\mathbb{F}^n) = n$, da en basis for $\mathbb{F}^n$ kan dannes af de n vektorer $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

Et særtilfælde af ovenstående er, når $\mathbb{C}$ betragtes som et vektorrum over sig selv. Da har det dimension én: $\dim_{\mathbb{C}}(\mathbb{C}) = 1$ (en mulig basis kan dannes af det komplekse tal 1). Men hvis $\mathbb{C}$ betragtes som et vektorrum over $\mathbb{R}$, er en basis givet ved $\{1, i\}$, som vi har set i Eksempel 9.2.5. Derfor er $\dim_{\mathbb{R}}(\mathbb{C}) = 2$.

Vektorrummet af $m \times n$-matricer $\mathbb{F}^{m\times n}$ har en basis bestående af de mn matricer $\mathbf{E}^{(i,j)}$ for $1 \leq i \leq m$ og $1 \leq j \leq n$, som vi har set i Eksempel 9.2.6. Derfor er $\dim_{\mathbb{F}}(\mathbb{F}^{m\times n}) = mn$.

Vi har set i Eksempel 9.2.3, at det komplekse vektorrum $\mathbb{C}[Z]$ har en basis med uendeligt mange elementer, nemlig $\{1, Z, Z^2, \ldots\}$. Derfor er $\dim_{\mathbb{C}}(\mathbb{C}[Z]) = \infty$.

Sætning 9.2.5

Hvis V har en endelig basis bestående af n vektorer, har enhver anden mængde af lineært uafhængige vektorer i V højst n elementer.

Bevis. Lad os betegne basisvektorerne ved $\mathbf{v}_1, \ldots, \mathbf{v}_n$ og den resulterende ordnede basis ved β. Vi vil bevise sætningen med et modstridsbevis. Antag derfor, at der findes en mængde med mindst $n+1$ lineært uafhængige vektorer, som vi kalder $\mathbf{w}_1, \ldots, \mathbf{w}_{n+1}$. Da β er en ordnet basis, findes der skalarer $a_{ij} \in \mathbb{F}$, således at

$$\mathbf{w}_j = a_{1j}\mathbf{v}_1 + \cdots + a_{nj}\mathbf{v}_n \text{ for } j = 1, \ldots, n+1.$$

Lad nu $\mathbf{A} = (a_{ij}) \in \mathbb{F}^{n\times(n+1)}$ være matricen med elementer a_{ij}. Bemærk, at den j'te søjle i $\mathbf{A}$ er lig med $[\mathbf{w}_j]_\beta$. Da $\mathbf{A}$ har n rækker, er dens rang $\rho(\mathbf{A})$ højst n. Da $\mathbf{A}$ har $n+1$ søjler, betyder dette, at $\rho(\mathbf{A}) < n+1$. Således ser vi ifølge Korollar 6.4.5, at det homogene system med koefficientmatrix $\mathbf{A}$ har løsninger, der ikke er nulløsningen. Lad $(c_1, \ldots, c_{n+1}) \in \mathbb{F}^{n+1}$ være en sådan løsning forskellig fra nulløsningen. Da har vi

$$c_1 \cdot \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix} + \cdots + c_{n+1} \cdot \begin{bmatrix} a_{1\,n+1} \\ \vdots \\ a_{n\,n+1} \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} c_1 \\ \vdots \\ c_{n+1} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Husk nu, at den j'te søjle i $\mathbf{A}$ er lig med $[\mathbf{w}_j]_\beta$. Dette betyder, at vi har $c_1 \cdot [\mathbf{w}_1]_\beta + c_{n+1} \cdot [\mathbf{w}_{n+1}]_\beta = \mathbf{0}$. Da det følger fra Lemma 9.2.2, at $[c_1 \cdot \mathbf{w}_1 + \cdots + c_{n+1} \cdot \mathbf{w}_{n+1}]_\beta = c_1 \cdot [\mathbf{w}_1]_\beta + c_{n+1} \cdot [\mathbf{w}_{n+1}]_\beta$, kan vi konkludere, at $[c_1 \cdot \mathbf{w}_1 + \cdots + c_{n+1} \cdot \mathbf{w}_{n+1}]_\beta = \mathbf{0}$. Derfor er $c_1 \cdot \mathbf{w}_1 + \cdots + c_{n+1} \cdot \mathbf{w}_{n+1} = \mathbf{0}$. Da $(c_1, \ldots, c_{n+1})$ ikke var nulvektoren, konkluderer vi, at vektorerne $\mathbf{w}_1, \ldots, \mathbf{w}_{n+1}$ ikke er lineært uafhængige alligevel. Denne modstrid viser, at antagelsen om, at der findes en mængde med mindst $n+1$ lineært uafhængige vektorer, var forkert. Derfor er sætningen sand. $\square$

Korollar 9.2.6

Hvis V har en endelig basis bestående af n vektorer, indeholder enhver anden basis for V også præcis n vektorer.

Bevis. Lad S være en basis for V bestående af n vektorer og T en hvilken som helst anden basis. Da vektorerne i T er lineært uafhængige, medfører Sætning 9.2.5, at antallet af vektorer i T højst er n. Lad os betegne ved m antallet af vektorer i T. Hvad vi lige har vist, er, at $m \leq n$. Nu anvender vi Sætning 9.2.5 igen, hvor vi denne gang antager T som basis, og vi konkluderer, at antallet af elementer i S er højst m, det vil sige: $n \leq m$. Ved at kombinere ulighederne $m \leq n$ og $n \leq m$, konkluderer vi, at $n = m$, hvilket er det, vi ønskede at vise. $\square$

Dette korollar retfærdiggør definitionen af dimensionen af et vektorrum V som antallet af basisvektorer i endeligdimensionelle tilfælde: uanset hvilken basis for V du vælger, vil den indeholde nøjagtig samme antal vektorer. Som tidligere nævnt vil basisvektorer typisk være forskellige på tværs af forskellige baser. Faktisk kan vi for endeligdimensionelle vektorrum karakterisere alle de mulige baser således:

Sætning 9.2.7

Lad V være et vektorrum over et legeme $\mathbb{F}$ med dimension n. Da er enhver mængde af n lineært uafhængige vektorer i V en basis for V.

Bevis. Lad os betegne vektorerne i en basis for V ved $\mathbf{v}_1, \ldots, \mathbf{v}_n$, og lad os skrive β for den tilsvarende ordnede basis. Lad $\mathbf{w}_1, \ldots, \mathbf{w}_n$ være n lineært uafhængige vektorer i V. For at vise, at disse udgør en basis, skal vi kun kontrollere punkt 2 i Definition 9.2.3. Det vil sige, vi skal vise, at enhver $\mathbf{v} \in V$ kan skrives som en linearkombination af $\mathbf{w}_1, \ldots, \mathbf{w}_n$. Da β er en basis, kan vi finde skalarer $a_{ij} \in \mathbb{F}$, således at

$$\mathbf{w}_j = a_{1j} \cdot \mathbf{v}_1 + \cdots + a_{nj} \cdot \mathbf{v}_n \text{ for } j = 1, \ldots, n,$$

eller ækvivalent ved brug af summationssymbolet:

$$\mathbf{w}_j = \sum_{i=1}^{n} a_{ij} \cdot \mathbf{v}_i \text{ for } j = 1, \ldots, n. \tag{9.3}$$

Lad nu $\mathbf{A} = (a_{ij}) \in \mathbb{F}^{n \times n}$ være matricen med elementer a_{ij}. Bemærk ligesom i beviset af Sætning 9.2.5, at den j'te søjle i $\mathbf{A}$ er lig med $[\mathbf{w}_j]_\beta$. Vi hævder, at disse søjler er lineært uafhængige vektorer i $\mathbb{F}^n$. For at indse at dette er sandt, antager vi, at $c_1 \cdot [\mathbf{w}_1]_\beta + c_n \cdot [\mathbf{w}_n]_\beta = \mathbf{0}$ for nogle $c_1, \ldots, c_n \in \mathbb{F}$. Da er $[c_1 \cdot \mathbf{w}_1 + \cdots + c_n \cdot \mathbf{w}_n]_\beta = \mathbf{0}$, hvilket medfører, at

$c_1 \cdot \mathbf{w}_1 + \cdots + c_n \cdot \mathbf{w}_n = \mathbf{0}$. Da vektorerne $\mathbf{w}_1, \ldots, \mathbf{w}_n$ er lineært uafhængige, konkluderer vi, at $c_1 = 0, \ldots, c_n = 0$, hvilket er det, vi ønskede at vise for at bevise vores påstand. Nu benytter vi Sætning 7.1.3 og Korollar 7.3.5 for at konkludere, at matricen $\mathbf{A}$ har en invers matrix $\mathbf{A}^{-1}$.

Lad os nu fortsætte med det, vi ønsker at vise: at $\mathbf{v} \in V$ kan skrives som en linearkombination af $\mathbf{w}_1, \ldots, \mathbf{w}_n$. Da $\mathbf{v}$ er en linearkombination af basisvektorerne $\mathbf{v}_1, \ldots, \mathbf{v}_n$, er det tilstrækkeligt at vise, at hver af basisvektorerne selv kan skrives som en linearkombination af $\mathbf{w}_1, \ldots, \mathbf{w}_n$. Lad os skrive $\mathbf{A}^{-1} = (c_{ij})_{1 \le i \le n; 1 \le j \le n}$. Vi hævder, at:

$$\mathbf{v}_j = c_{1j} \cdot \mathbf{w}_1 + \cdots + c_{nj} \cdot \mathbf{w}_n \text{ for } j = 1, \ldots, n.$$

Ækvivalent ved brug af summationssymbolet hævder vi, at:

$$\mathbf{v}_j = \sum_{k=1}^{n} c_{kj} \cdot \mathbf{w}_k \text{ for } k = 1, \ldots, n.$$

For at vise denne påstand anvender vi først Ligning (9.3) til at se, at:

$$\begin{aligned}
\sum_{k=1}^{n} c_{kj} \cdot \mathbf{w}_k &= \sum_{k=1}^{n} c_{kj} \cdot \left(\sum_{i=1}^{n} a_{ik} \cdot \mathbf{v}_i \right) \\
&= \sum_{k=1}^{n} \sum_{i=1}^{n} c_{kj} \cdot a_{ik} \cdot \mathbf{v}_i \\
&= \sum_{k=1}^{n} \sum_{i=1}^{n} a_{ik} \cdot c_{kj} \cdot \mathbf{v}_i \\
&= \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} \cdot c_{kj} \cdot \mathbf{v}_i \\
&= \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik} \cdot c_{kj} \right) \cdot \mathbf{v}_i.
\end{aligned}$$

Bemærk her, at udtrykket $\sum_{k=1}^{n} a_{ik} \cdot c_{kj}$ er det (i, j)'te element i matrixproduktet $\mathbf{A} \cdot \mathbf{A}^{-1}$. Da $\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I}_n$, har vi $\sum_{k=1}^{n} a_{ik} \cdot c_{kj} = 1$, hvis $i = j$, og ellers $\sum_{k=1}^{n} a_{ik} \cdot c_{kj} = 0$. Derfor kan vi konkludere, at $\sum_{k=1}^{n} c_{kj} \cdot \mathbf{w}_k = \mathbf{v}_j$, hvilket er det, vi ønskede at vise. □

9.3 Underrum i et vektorrum

For et givet vektorrum V over et hvilket som helst legeme $\mathbb{F}$ er det muligt, at man har en delmængde W af V, som er lukket under den samme skalarmultiplikation og addition, der er defineret på V. Ordet "lukket" er blot en måde at sige på, at hvis $\mathbf{v} \in W$, og $c \in \mathbb{F}$, så er $c \cdot \mathbf{v} \in W$, og hvis $\mathbf{u}, \mathbf{v} \in W$, så er $\mathbf{u} + \mathbf{v} \in W$. Da V er et vektorrum, har vi altid $c \cdot \mathbf{v} \in V$ og $\mathbf{u} + \mathbf{v} \in V$, men hvis også W er lukket under skalarmultiplikation og addition, ender vektorerne $c \cdot \mathbf{v}$ og $\mathbf{u} + \mathbf{v}$ igen i W. Lad os gennemgå to eksempler på dette.

Eksempel 9.3.1

Lad os overveje det komplekse vektorrum $\mathbb{C}^2$ og delmængden $W = \{(z, 2 \cdot z) \mid z \in \mathbb{C}\}$ heraf. Her vil enhver addition af to elementer fra W give et resultat, der også tilhører W, da $(z, 2 \cdot z) + (w, 2 \cdot w) = (z + w, 2 \cdot (z + w))$ for alle $z, w \in \mathbb{C}$. Desuden giver også multiplikation af et element fra W med en skalar $c \in \mathbb{C}$ et resultat tilhørende W, da $c \cdot (z, 2 \cdot z) = (c \cdot z, 2 \cdot (c \cdot z))$. Faktisk er W et vektorrum ved brug af denne skalarmultiplikation og addition. For eksempel har vi $(0, 0) \in W$, da $(0, 0) = (0, 2 \cdot 0)$. Desuden er $-(z, 2 \cdot z) = ((-z), 2 \cdot (-z))$ for ethvert $z \in \mathbb{C}$, hvilket viser, at hvis $\mathbf{v} \in W$, så er også $-\mathbf{v} \in W$. Læseren opfordres til at verificere de resterende vektorrumsaksiomer også. Bemærk, at $\dim_{\mathbb{C}}(W) = 1$ (en mulig basis for W er givet ved $\{(1, 2)\}$).

Eksempel 9.3.2

Betragt vektorrummet $\mathbb{R}^{2 \times 2}$ af 2×2-matricer med koefficienter i $\mathbb{R}$. Som vi har set tidligere, er dette et reelt vektorrum af dimension fire. Lad nu D være en delmængde af $\mathbb{R}^{2 \times 2}$ bestående af alle diagonalmatricer, det vil sige:

$$D = \left\{ \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \right\}.$$

Denne mængde D er lukket under skalarmultiplikation og matrixaddition, hvilket betyder, at vi for $\mathbf{A}, \mathbf{B} \in D$ og $c \in \mathbb{F}$ har, at $c \cdot \mathbf{A} \in D$, og at $\mathbf{A} + \mathbf{B} \in D$. Lad os kontrollere dette. Hvis $\mathbf{A}$ har diagonalelementerne λ_1 og λ_2, og $\mathbf{B}$ har diagonalelementerne μ_1 og μ_2, så er:

$$c \cdot \mathbf{A} = c \cdot \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} c \cdot \lambda_1 & 0 \\ 0 & c \cdot \lambda_2 \end{bmatrix} \in D,$$

og

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} + \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 + \mu_1 & 0 \\ 0 & \lambda_2 + \mu_2 \end{bmatrix} \in D.$$

Man kan vise, at D faktisk er et reelt vektorrum af dimension to: en mulig ordnet basis er $(\mathbf{E}^{(1,1)}, \mathbf{E}^{(2,2)})$.

Til at samle disse typer af eksempler har vi følgende:

Definition 9.3.1

Lad V være et vektorrum over et legeme $\mathbb{F}$. Et *underrum* af V er en delmængde W af V, som er et vektorrum over $\mathbb{F}$ under samme skalarmultiplikation og vektoraddition, der er defineret på V.

Med andre ord, hvis $W \subseteq V$ er lukket under den skalarmultiplikation og vektoraddition, som V har, "nedarver" W disse operationer. Hvis W med disse operationer opfylder alle vektorrumsaksiomerne fra Definition 9.1.1, kaldes det et underrum i W. Ethvert vektorrum V har mindst to underrum: V kan opfattes som et underrum i sig selv, og der findes også altid underrummet $\{\mathbf{0}\}$, der kun indeholder nulvektoren i V. Generelt har V mange flere underrum. I alle tilfælde vil man altid kunne sige følgende om dimensionen af et underrum:

Lemma 9.3.1

Lad V være et vektorrum over et legeme $\mathbb{F}$ af dimension n og W et underrum i V. Da gælder der, at $\dim W \leq n$.

Bevis. Da V har en basis med n vektorer, og W har en basis med $\dim W$ vektorer, udgør basisvektorerne for W en sekvens af $\dim W$ lineært uafhængige vektorer. Derfor medfører Sætning 9.2.5, at $\dim W \leq n$. □

Da V allerede opfylder alle vektorrumsaksiomer, viser det sig ikke at være nødvendigt at kontrollere dem alle, når man undersøger, om en delmængde W er et underrum. Mere præcist har vi følgende lemma:

Lemma 9.3.2

Lad V være et vektorrum over $\mathbb{F}$ og W en delmængde af V, der ikke er tom. Da er W et underrum i V, hvis følgende er opfyldt:

$$\text{for alle } \mathbf{u}, \mathbf{v} \in W \text{ og alle } c \in \mathbb{F} \text{ gælder der, at } \mathbf{u} + c \cdot \mathbf{v} \in W. \tag{9.4}$$

Bevis. Lad os først vise, at W er lukket under skalarmultiplikation og vektoraddition i V. Da W ikke er tom, indeholder den mindst én vektor, som vi kan kalde $\mathbf{w}$. Ved at vælge $\mathbf{u} = \mathbf{w}$ og $\mathbf{v} = \mathbf{w}$ i Ligning (9.4) kan vi konkludere, at vektoren $\mathbf{w} + (-1) \cdot \mathbf{w}$ også er i W. Ved for eksempel at anvende Ligning (9.2), medfører dette, at $\mathbf{0} \in W$. Nu hvor vi ved dette, kan vi anvende Ligning (9.4) igen, denne gang med $\mathbf{u} = \mathbf{0}$ og $\mathbf{v} \in W$ valgt vilkårligt. Vi kan dermed konkludere, at for vilkårlige $\mathbf{v} \in W$ er også $c \cdot \mathbf{v}$ i W. Dette viser, at W er lukket under skalarmultiplikation. Ved at anvende Ligning (9.4) for vilkårlige $\mathbf{u}, \mathbf{v} \in W$ og $c = 1$, konkluderer vi, at $\mathbf{u} + \mathbf{v}$ er i W. Derfor er W lukket under vektoraddition.

Lad os dernæst vise, at W er et vektorrum ved at overveje de otte vektorrumsaksiomer fra Definition 9.1.1. Da aksiomerne 1, 2, 5, 6, 7 og 8 gælder for alle vektorer i V, må de også gælde for alle vektorer i en delmængde af V. Derfor er det kun aksiomerne 3 og 4, der skal kontrolleres. Aksiom 3 er opfyldt, da vi allerede har vist, at Ligning (9.4) medfører, at $\mathbf{0} \in W$. Hvad angår aksiom 4, har vi, at hvis $\mathbf{v} \in W$, så er $(-1) \cdot \mathbf{v} \in W$, da W er lukket under skalarmultiplikation. Men ifølge Ligning (9.2) er $(-1) \cdot \mathbf{v} = -\mathbf{v}$, så den additivt inverse $-\mathbf{v}$ tilhører W for alle $\mathbf{v}$ i W. □

Eksempel 9.3.3

Anvendes Lemma 9.3.2, er det ikke svært at vise, at delmængderne W og D fra Eksemplerne 9.3.1 og 9.3.2 er underrum. Læseren opfordres til at kontrollere, at betingelsen i Ligning (9.4) er opfyldt for disse eksempler.

Eksempel 9.3.4

Lad $C_\infty(\mathbb{R})$ være mængden af alle uendeligdifferentiable funktioner $f : \mathbb{R} \to \mathbb{R}$. Det er uden for rammerne af denne tekst at definere helt præcist, hvad en uendeligdifferentiabel funktion er, men groft sagt betyder det følgende: hvis grænsen $\lim_{a\to 0}(f(t+a)-f(t))/a$ eksisterer for alle $t \in \mathbb{R}$, kan vi definere den afledte af f, betegnet f', som funktionen $f' : \mathbb{R} \to \mathbb{R}$ med $t \mapsto \lim_{a\to 0}(f(t+a)-f(t))/a$. En uendeligdifferentiabel funktion $f : \mathbb{R} \to \mathbb{R}$ har den egenskab, at man kan blive ved med at differentiere den, så ofte man vil. Specifikt eksisterer ikke kun dens afledte f' men også den afledte af f' (betegnet f'' eller $f^{(2)}$), den afledte af f'' (betegnet f''' eller $f^{(3)}$) og så videre. Mere generelt betegnes for ethvert positivt heltal n den n'te-afledte af f ved $f^{(n)}$. Mere præcist definerer man den n'te-afledte rekursivt som følger:

$$f^{(n)} = \begin{cases} f & \text{hvis } n = 0, \\ (f^{(n-1)})' & \text{hvis } n > 0. \end{cases}$$

Mængden $C_\infty(\mathbb{R})$ er et underrum i vektorrummet F fra Eksempel 9.1.6. Dette vises ved, at man for $f, g \in C_\infty(\mathbb{R})$ og $c \in \mathbb{R}$ viser, at $f + c \cdot g \in C_\infty(\mathbb{R})$. Man kan vise induktivt, at $(f + c\cdot g)^{(n)} = f^{(n)} + c \cdot g^{(n)}$ for enhver $n \in \mathbb{Z}_{\geq 0}$, og derfor er $f + c \cdot g$ uendeligdifferentiabel, hvilket var det, vi skulle vise.

Der findes en særlig måde at konstruere et underrum på, som vi nu vil dykke ned i.

Definition 9.3.2

Lad V være et vektorrum over $\mathbb{F}$ og S en mængde af vektorer fra V. Da betegner *udspændingen* af S, betegnet $\mathrm{Span}(S)$, mængden af alle de mulige linearkombinationer af vektorer fra S. Hvis $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, gælder der specifikt, at

$$\mathrm{Span}(S) = \{c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n \mid c_1, \ldots, c_n \in \mathbb{F}\}.$$

Bemærk, at begrebet udspænding på engelsk kaldes *span*, hvilket forklarer betegnelsen i definitionen. Man vil typisk definere $\mathrm{Span}(\emptyset) = \{\mathbf{0}\}$. Som en konsekvens heraf siges den tomme mængde $\emptyset$ at udgøre en basis for vektorrummet $\{\mathbf{0}\}$. Det viser sig, at mængden $\mathrm{Span}(S)$ for enhver delmængde $S \subseteq V$ faktisk er et underrum i V, hvilket kan vises ved anvendelse af for eksempel Lemma 9.3.2. Hvis W er et givet underrum i et vektorrum V, og $W = \mathrm{Span}(S)$, siger man, at vektorerne i S udspænder W, eller omvendt at W er udspændt af vektorerne i S. Vektorerne i en basis for W udspænder selvfølgelig W, men generelt set er en mængde af vektorer, der udspænder W, ikke nødvendigvis lineært uafhængige.

Eksempel 9.3.5

Betragt det reelle vektorrum $V = \mathbb{R}^2$, og lad

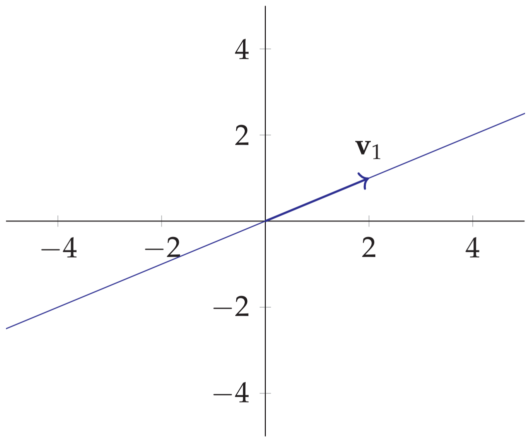
$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Lad os bestemme udspændingen af $\mathbf{v}_1$, hvilket også kan skrives som $\mathrm{Span}(S)$, hvor $S = \{\mathbf{v}_1\}$.

Ved at anvende Definition 9.3.2 får vi:

$$\begin{aligned} \mathrm{Span}(\{\mathbf{v}_1\}) &= \{c_1 \cdot \mathbf{v}_1 \mid c_1 \in \mathbb{R}\} \\ &= \left\{ c_1 \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} \mid c_1 \in \mathbb{R} \right\} \\ &= \left\{ \begin{bmatrix} c_1 \cdot 2 \\ c_1 \cdot 1 \end{bmatrix} \mid c_1 \in \mathbb{R} \right\} \\ &= \left\{ \begin{bmatrix} 2c_1 \\ c_1 \end{bmatrix} \mid c_1 \in \mathbb{R} \right\}. \end{aligned}$$

Grafisk set består udspændingen af $\mathbf{v}_1$ af alle vektorer, der ligger på linjen gennem $\mathbf{v}_1$ (se Figur 9.1, hvor udspændingen er angivet med en blå linje).



Figur 9.1: Udspændingen af en vektor forskellig fra nulvektoren i $\mathbb{R}^2$.

Eksempel 9.3.6

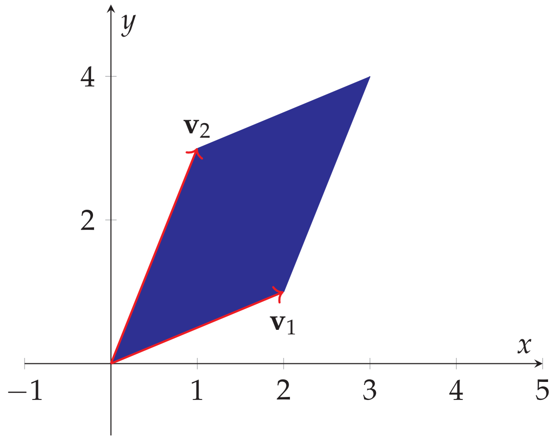
Betragt det reelle vektorrum $V = \mathbb{R}^2$, og lad

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{og} \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

Lad os bestemme udspændingen af vektorerne $\mathbf{v}_1$ og $\mathbf{v}_2$, altså $\mathrm{Span}(S)$ med $S = \{\mathbf{v}_1, \mathbf{v}_2\}$. Ved at benytte Definition 9.3.2 får vi:

$$\begin{aligned} \mathrm{Span}(\{\mathbf{v}_1, \mathbf{v}_2\}) &= \{c_1 \cdot \mathbf{v}_1 + c_2 \cdot \mathbf{v}_2 \mid c_1, c_2 \in \mathbb{R}\} \\ &= \left\{ c_1 \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix} \mid c_1, c_2 \in \mathbb{R} \right\} \\ &= \left\{ \begin{bmatrix} 2c_1 + c_2 \\ c_1 + 3c_2 \end{bmatrix} \mid c_1, c_2 \in \mathbb{R} \right\}. \end{aligned}$$

Grafisk set er situationen, at udspændingen af vektorerne $\mathbf{v}_1$ og $\mathbf{v}_2$ udgør hele rummet $\mathbb{R}^2$. I Figur 9.1 har vi farvelagt området, man opnår, når man plotter alle vektorer af formen $c_1 \cdot \mathbf{v}_1 + c_2 \cdot \mathbf{v}_2$, hvis c_1 og c_2 vælges frit i intervallet $[0, 1]$.



Figur 9.2: Området dækket af linearkombinationen $c_1 \cdot \mathbf{v}_1 + c_2 \cdot \mathbf{v}_2$ for alle $c_1, c_2 \in [0, 1]$.

Eksempel 9.3.7

Betragt det reelle vektorrum $V = \mathbb{R}^3$, og lad

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \text{ og } \mathbf{v}_3 = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}.$$

Spørgsmål: Find en basis for underrummet W udspændt af de tre vektorer $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

Svar:

Et umiddelbart, men desværre forkert, gæt kunne være, at de tre vektorer $\mathbf{v}_1, \mathbf{v}_2$ og $\mathbf{v}_3$ selv udgør en basis. Selvfølgelig kan enhver vektor i W skrives som en linearkombination af $\mathbf{v}_1, \mathbf{v}_2$ og $\mathbf{v}_3$. Dette er en direkte konsekvens af Definition 9.3.2 af udspændingen. Men for at kunne udgøre en basis skal de tre vektorer $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ også være lineært uafhængige, og det viser det sig, at de ikke er. Ved hjælp af Sætning 7.1.3 kan dette undersøges ved, at vi bestemmer den reducerede trappeform af den 3×3-matrix $\mathbf{A}$, hvis søjler er $\mathbf{v}_1, \mathbf{v}_2$ og $\mathbf{v}_3$. Vi undlader detaljerne af denne udregning og opfordrer i stedet læseren til at verificere, at den reducerede trappeform bliver:

$$\begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Dette resultat viser, at de tre vektorer $\mathbf{v}_1, \mathbf{v}_2$ og $\mathbf{v}_3$ er lineært afhængige. Samtidig viser det dog også, at de to første af vektorerne er lineært uafhængige (sammenlign med Eksempel 7.1.4, hvor en lignende tilgang blev anvendt på tre vektorer i $\mathbb{C}^3$). Vi kan konkludere, at $\mathbf{v}_3$ kan udtrykkes som en linearkombination af $\mathbf{v}_1$ og $\mathbf{v}_2$. Dette betyder, at de to vektorer $\mathbf{v}_1$ og

$\mathbf{v}_2$ udspænder præcis det samme underum i $\mathbb{R}^3$ som de tre vektorer $\mathbf{v}_1, \mathbf{v}_2$ og $\mathbf{v}_3$. Derfor er $\{\mathbf{v}_1, \mathbf{v}_2\}$ en basis for W.

Vi har allerede fuldt ud besvaret spørgsmålet, men nogle gange vil man gerne se udpenslet, hvordan $\mathbf{v}_3$ kan udtrykkes som en linearkombination af $\mathbf{v}_1$ og $\mathbf{v}_2$. For at gøre dette skal vi finde en løsning på formen $(c_1, c_2, 1)$ til det homogene system af lineære ligninger, der har koefficientmatrix $\mathbf{A}$. Ved at betragte den reducerede trappeform af $\mathbf{A}$, ser vi, at $(-4, 1, 1)$ er en sådan løsning. Derfor gælder $(-4) \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2 + \mathbf{v}_3 = \mathbf{0}$, hvilket medfører, at $\mathbf{v}_3 = 4 \cdot \mathbf{v}_1 - \mathbf{v}_2$.

Geometrisk set sker der i dette eksempel dét, at vektorerne $\mathbf{v}_1, \mathbf{v}_2$ og $\mathbf{v}_3$ alle tre ligger i samme plan (nemlig i det to-dimensionelle underum W, vi fandt i eksemplet). Mere generelt kan man vise, at når tre vektorer i et reelt vektorrum $\mathbb{R}^n$ alle ligger i samme to-dimensionelle underum i $\mathbb{R}^n$, så er de lineært afhængige.

Som vi så i det tidligere eksempel, betyder det, at et underum er udspændt af bestemte vektorer, ikke nødvendigvis at disse vektorer er lineært uafhængige. Den procedure, vi benyttede os af i Eksempel 9.3.7 for at finde en basis, kan generaliseres. Lad os gøre det i følgende sætning:

Sætning 9.3.3

Lad et underum W af vektorrummet $\mathbb{F}^n$ være udspændt af vektorerne $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$. Antag videre, at den reducerede trappeform af matricen med søjlerne $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$ har pivoter i søjlerne $j_1, \ldots, j_\rho$. Da er $\{\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_\rho}\}$ en basis for W.

Bevis. Lad os betegne matricen med søjlerne $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$ ved $\mathbf{A}$ og den reducerede trappeform af $\mathbf{A}$ ved $\mathbf{B}$. Jævnfør definitionen på den reducerede trappeform af en matrix udgøres søjlerne i $\mathbf{B}$ med søjleindekser $i_1, \ldots, i_\rho$ af de første ρ standardbasisvektorer $\mathbf{e}_1, \ldots, \mathbf{e}_\rho$. Vi ved, at disse er lineært uafhængige. Vi påstår nu, at dette medfører, at søjlerne i $\mathbf{B}$ med søjleindekser $i_1, \ldots, i_\rho$ også er lineært uafhængige. Hvis $c_{j_1} \cdot \mathbf{u}_{j_1} + \cdots + c_{j_\rho} \cdot \mathbf{u}_{j_\rho} = \mathbf{0}$, så er tuplet $(v_1, \ldots, v_\ell) \in \mathbb{F}^\ell$, der defineres ved $v_j = c_j$, hvis $j \in \{j_1, \ldots, j_\rho\}$, og ellers ved $v_j = 0$, en løsning til det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A}$. Men vi ved, at enhver sådan løsning også er en løsning til det homogene system med koefficientmatrix $\mathbf{B}$. Da vi allerede har observeret, at søjlerne i $\mathbf{B}$ med søjleindekser $i_1, \ldots, i_\rho$ er lineært uafhængige, konkluderer vi, at der nødvendigvis gælder, at $c_{j_1} = 0, \ldots, c_{j_\rho} = 0$. Dette viser, at vektorerne $\{\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_\rho}\}$ er lineært uafhængige.

Vælg nu en hvilken som helst søjle $\mathbf{u}_j$ i $\mathbf{A}$, for hvilken $j \notin \{j_1, \ldots, j_\rho\}$. Igen ser vi af definitionen på den reducerede trappeform, at den j'te søjle i $\mathbf{B}$ har nuller som sine sidste $n - \rho$ elementer. Derfor kan den udtrykkes som en linearkombination af $\mathbf{e}_1, \ldots, \mathbf{e}_\rho$, hvilke blot er søjlerne $j_1, \ldots, j_\ell$ i $\mathbf{B}$. Dette betyder, at det homogene system med koefficientmatrix $\mathbf{B}$ har en løsning $(v_1, \ldots, v_\ell)$, således at $v_j = 1$, og $v_k = 0$ for alle $k \notin \{j, j_1, \ldots, j_\rho\}$. Ved at udnytte at dette også er en løsning til det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A}$, får vi nu, at den j'te søjle i $\mathbf{A}$ kan udtrykkes som en linearkombination af søjlerne $j_1, \ldots, j_\rho$. Dette beviser, at udspændingen af $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$ er den samme som udspændingen af $\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_\rho}$.

Kombineres alt det ovenstående, kan vi konkludere, at $\{\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_\rho}\}$ udgør en basis for W. $\square$

Tilbage i Sætning 6.4.4 blev løsningsmængden til et homogent system af lineære ligninger netop beskrevet som udspændingen af $n - \rho$ vektorer. I det tilfælde dannede disse vektorer faktisk en basis for løsningsmængden, da de var lineært uafhængige. Lad os vise dette nu.

Korollar 9.3.4

Lad et homogent system af m lineære ligninger i n variable over et legeme $\mathbb{F}$ være givet. Betegn koefficientmatricen for dette system ved $\mathbf{A}$, og antag, at denne matrix har rang ρ. De $n - \rho$ vektorer $\mathbf{v}_1, \ldots, \mathbf{v}_{n-\rho}$, som er angivet i Sætning 6.4.4, udgør en basis for løsningen til det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A}$.

Bevis. Lad os vise en skitse af beviset: vi benytter samme notation for vektorerne $\mathbf{c}_i$ og matricen $\hat{\mathbf{A}}$ som i Sætning 6.4.4. Ser vi tilbage på, hvordan vektoren $\mathbf{v}_i$ blev defineret i Sætning 6.4.4, ser vi, at $\mathbf{v}_i$ har et 1-tal i koordinatposition j, hvor j opfylder, at $\mathbf{c}_i$ er den j'te søjle i $\hat{\mathbf{A}}$. Ligeledes ser vi, at $\mathbf{v}_i$ har koefficienter lig med 0 derefter, da $\mathbf{c}_i$ kun indeholder nuller efter sin i'te koefficient. Derfor er matricen med søjlerne $\mathbf{v}_1, \ldots, \mathbf{v}_{n-\rho}$ på trappeform. Dette medfører, at den tilsvarende matrix på reduceret trappeform har en pivot i hver søjle. Sætning 9.3.3 medfører da, at $\{\mathbf{v}_1, \ldots, \mathbf{v}_{n-\rho}\}$ udgør en basis. $\square$

Korollar 9.3.5

Lad V være et vektorrum over et legeme $\mathbb{F}$ af endelig dimension n med ordnet basis β, og lad $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$ være vektorer i V. Antag videre, at den reducerede trappeform af matricen med søjlerne $[\mathbf{u}_1]_\beta, \ldots, [\mathbf{u}_\ell]_\beta$ har pivoter i søjlerne $j_1, \ldots, j_\rho$. Da er en basis for $\mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_\ell)$ givet ved $\{\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_\rho}\}$.

Bevis. Vi viser en skitse af beviset: vi har fra Sætning 9.3.3, at en basis for underrummet af $\mathbb{F}^n$, der er genereret af $[\mathbf{u}_1]_\beta, \ldots, [\mathbf{u}_\ell]_\beta$, er givet ved $\{[\mathbf{u}_{j_1}]_\beta, \ldots, [\mathbf{u}_{j_\rho}]_\beta\}$. Nu kan Sætning 9.2.3 anvendes til at vise, at $\{\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_\rho}\}$ udgør en basis for $\mathrm{Span}(\mathbf{u}_1, \ldots, \mathbf{u}_\ell)$. $\square$

9.4 Ekstra: hvorfor har ethvert vektorrum en basis?

Dette afsnit er ikke obligatorisk læsning og kan springes over. Det er tænkt som ekstra materiale for en studerende, der har tid og motivation til det.

I de tidligere afsnit har vi uden videre benyttet det faktum, at ethvert vektorrum V har en basis. For at bevise dette antagede faktum, skal vi studere mængden $\mathcal{I}(V)$, som består af alle delmængder af V, hvis elementer er lineært uafhængige vektorer. For eksempel er $\emptyset \in \mathcal{I}(V)$, da den tomme mængde ikke indeholder nogen vektorer og derfor ikke kan indeholde lineært afhængige vektorer. Hvis $V \neq \{\mathbf{0}\}$, er enhver delmængde af formen $\{\mathbf{v}\}$ i $\mathcal{I}(V)$, så længe $\mathbf{v} \neq \mathbf{0}$. Intuitivt bør en basis B for V være en mængde, der indeholder så mange lineært uafhængige vektorer som muligt. Mere præcist ville denne intuition sige, at $B \in \mathcal{I}(V)$, og at ingen mængde af lineært uafhængige vektorer kan indeholde B som en skarpt mindre delmængde. Denne anden intuitive egenskab kan omformuleres ved at sige, at hvis $C \in \mathcal{I}(V)$, og $B \subseteq C$, så er $B = C$. En sådan B kaldes et maksimalelement i $\mathcal{I}(V)$.

Ovenstående diskussion har kun til formål at give en intuitiv forståelse, men den følgende sætning viser, at der er substans i diskussionen.

Sætning 9.4.1

Lad B være et maksimalelement i $\mathcal{I}(V)$. Da er B en basis for V.

Bevis. Jævnfør definitionen af $\mathcal{I}(V)$ er vektorerne i B lineært uafhængige. Det, vi skal vise, er, at enhver vektor i V kan skrives som en linearkombination af vektorerne i B. Antag, at dette ikke er tilfældet. Da findes der en $\mathbf{v} \in V$, således at enhver linearkombination af vektorer i B er forskellig fra $\mathbf{v}$. Vi hævder, at mængden $B \cup \{\mathbf{v}\}$ i dette tilfælde består af lineært uafhængige vektorer. For at vise dette, antager vi, at

$$c_0 \cdot \mathbf{v} + c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n = \mathbf{0} \tag{9.5}$$

for nogle $c_0, c_1, \ldots, c_n \in \mathbb{F}$ og $\mathbf{v}_1, \ldots, \mathbf{v}_n \in B$. Hvis $c_0 = 0$, ser vi straks, at $c_1 = 0, \ldots, c_n = 0$, da vektorerne i B er lineært uafhængige. Dog kan c_0 ikke være forskellig fra nul, da Ligning (9.5) dermed ville medføre, at $\mathbf{v} = -c_0^{-1} \cdot c_1 \cdot \mathbf{v}_1 - \cdots - c_0^{-1} \cdot c_n \cdot \mathbf{v}_n$, i strid med antagelsen om, at $\mathbf{v}$ ikke kan skrives som en linearkombination af vektorer fra B. Derfor består mængden $B \cup \{\mathbf{v}\}$ rent faktisk af lineært uafhængige vektorer som hævdet. En anden måde at sige dette på er, at $B \cup \{\mathbf{v}\} \in \mathcal{I}(V)$, hvilket igen medfører, at B ikke var et maksimalelement i $\mathcal{I}(V)$, i strid med antagelsen om, at det var. Modstriden viser, at enhver vektor i V kan skrives som en linearkombination af vektorer fra B. Derfor er B en basis. $\square$

Sætningen har følgende konsekvens: for at vise, at ethvert vektorrum V har en basis, er det nok at vise, at mængden $\mathcal{I}(V)$ altid indeholder et maksimalelement. Dette er en direkte konsekvens af et berømt lemma kaldet Zorns lemma. At formulere og bevise Zorns lemma kræver dog værktøjer fra grundlæggende matematik, som ligger uden for rammerne af denne tekst.

Kapitel 10

Lineære afbildninger mellem vektorrum

Lad to vektorrum V_1 og V_2 være givet, begge over det samme legeme $\mathbb{F}$. En lineær afbildning er da en funktion fra V_1 til V_2, der er kompatibel med skalarmultiplikation og vektoraddition. Mere præcist har vi følgende:

Definition 10.0.1

Lad V_1 og V_2 være vektorrum over et legeme $\mathbb{F}$. Da er en *lineær afbildning* fra V_1 til V_2 en funktion $L : V_1 \to V_2$, således at:

(i) $L(\mathbf{u} + \mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v})$ for alle $\mathbf{u}, \mathbf{v} \in V_1$,

(ii) $L(c \cdot \mathbf{u}) = c \cdot L(\mathbf{u})$ for alle $c \in \mathbb{F}$ og $\mathbf{u} \in V_1$.

En lineær afbildning kaldes også en *lineær transformation*. Bemærk, at begrebet afbildning på engelsk kaldes *map*. Bemærk desuden i formlen $L(\mathbf{u} + \mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v})$, at symbolet $+$ i $\mathbf{u} + \mathbf{v}$ betegner vektoraddition i V_1, mens $+$ i $L(\mathbf{u}) + L(\mathbf{v})$ betegner vektoraddition i V_2. Tilsvarende betegner $\cdot$ i $c \cdot \mathbf{u}$ skalarmultiplikation i V_1, mens det i $c \cdot L(\mathbf{u})$ betegner skalarmultiplikation i V_2.

Hvor vi i det forrige kapitel studerede ét vektorrum ad gangen, kan lineære afbildninger forbinde forskellige vektorrum med hinanden. Lineære afbildninger respekterer vektorrum-strukturen: vælges en skalar c lig med 0, og benyttes Ligning (9.1), fås for eksempel

$$L(\mathbf{0}) = \mathbf{0}, \tag{10.1}$$

hvor $\mathbf{0}$ på venstre side af ligningen betegner nulvektoren i V_1, og den på højre side betegner nulvektoren i V_2. Tilsvarende, ved at vælge $c = -1$ og benytte Ligning (9.2), får man

$$L(-\mathbf{u}) = -L(\mathbf{u}). \tag{10.2}$$

Der findes selvfølgelig rigtig mange mulige afbildninger mellem to vektorrum, og generelt vil ikke mange være lineære. Lad os se på nogle eksempler.

Eksempel 10.0.1

Betragt følgende funktioner fra $\mathbb{R}$ til $\mathbb{R}$. Hvilke af dem er lineære afbildninger?

(a) $f : \mathbb{R} \to \mathbb{R}$ defineret ved $x \mapsto x^2$,

(b) $g : \mathbb{R} \to \mathbb{R}$ defineret ved $x \mapsto 2x + 1$,

(c) $h : \mathbb{R} \to \mathbb{R}$ defineret ved $x \mapsto 2x$.

Svar:

(a) $f : \mathbb{R} \to \mathbb{R}$ defineret ved $x \mapsto x^2$ er ikke en lineær afbildning. For eksempel har vi $f(1+1) = f(2) = 4$, hvor vi, hvis f havde været en lineær afbildning, burde have haft $f(1+1) = f(1) + f(1) = 1 + 1 = 2$.

(b) $g : \mathbb{R} \to \mathbb{R}$ defineret ved $x \mapsto 2x + 1$ er heller ikke en lineær afbildning, selvom grafen for denne funktion er en linje. Vi har $g(0) = 1$, men hvis g havde været en lineær afbildning, burde vi have haft $g(0) = 0$ jævnfør Ligning (10.1).

(c) $h : \mathbb{R} \to \mathbb{R}$ defineret ved $x \mapsto 2x$ er en lineær afbildning. For alle $x, y \in \mathbb{R}$ har vi $h(x+y) = 2(x+y) = 2x + 2y = h(x) + h(y)$, og for alle $c \in \mathbb{R}$ og $x \in \mathbb{R}$ har vi $c \cdot h(x) = c2x = 2cx = h(c \cdot x)$.

Mere generelt er lineære afbildninger fra $\mathbb{R}$ til $\mathbb{R}$ netop de funktioner, hvis graf er en ret linje, der går gennem origo. Dette vil altså være funktioner $L : \mathbb{R} \to \mathbb{R}$, der opfylder, at $x \mapsto a \cdot x$ for en konstant $a \in \mathbb{R}$. Årsagen er, at hvis $L : \mathbb{R} \to \mathbb{R}$ er en lineær afbildning, så har vi for alle $x \in \mathbb{R}$, at $L(x) = L(x \cdot 1) = x \cdot L(1)$. I den sidste lighed gjorde vi brug af egenskab to fra Definition 10.0.1. Ved at sætte $a = L(1)$ får vi således $L(x) = a \cdot x$ for alle $x \in \mathbb{R}$.

Vi vil komme til at se, at der er en stærk forbindelse mellem lineære afbildninger og matricer. Af denne grund lægger vi ud med at studere lineære afbildninger, der kommer af matricer, for derefter at vende tilbage til studiet af lineære afbildninger i en mere generel ramme.

10.1 Lineære afbildninger ved anvendelse af matricer

Lad os starte med at definere en stor klasse af lineære afbildninger.

Definition 10.1.1

Lad $\mathbb{F}$ være et legeme og $\mathbf{A} \in \mathbb{F}^{m \times n}$ en matrix. Da defineres funktionen $L_{\mathbf{A}} : \mathbb{F}^n \to \mathbb{F}^m$ ved $L_{\mathbf{A}}(\mathbf{v}) = \mathbf{A} \cdot \mathbf{v}$ for alle $\mathbf{v} \in \mathbb{F}^n$.

Det viser sig, at alle funktioner $L_{\mathbf{A}} : \mathbb{F}^n \to \mathbb{F}^m$ som defineret ovenfor er lineære.

Lemma 10.1.1
Funktionen $L_{\mathbf{A}} : \mathbb{F}^n \to \mathbb{F}^m$ i Definition 10.1.1 er en lineær afbildning.

Bevis. Vi skal kontrollere de to betingelser fra Definition 10.0.1. Først og fremmest har vi

$$\begin{aligned} L_{\mathbf{A}}(\mathbf{u}+\mathbf{v}) &= \mathbf{A}\cdot(\mathbf{u}+\mathbf{v}) \\ &= \mathbf{A}\cdot\mathbf{u}+\mathbf{A}\cdot\mathbf{v} \\ &= L_{\mathbf{A}}(\mathbf{u})+L_{\mathbf{A}}(\mathbf{v}) \text{ for alle } \mathbf{u},\mathbf{v}\in\mathbb{F}^n. \end{aligned}$$

Dernæst har vi

$$L_{\mathbf{A}}(c\cdot\mathbf{u}) = \mathbf{A}\cdot(c\cdot\mathbf{u}) = c\cdot(\mathbf{A}\cdot\mathbf{u}) = c\cdot L_{\mathbf{A}}(\mathbf{u}) \text{ for alle } c\in\mathbb{F} \text{ og } \mathbf{u}\in\mathbb{F}^n.$$

□

I Eksempel 10.0.1 så vi, at funktionen $h : \mathbb{R} \to \mathbb{R}$, $x \mapsto 2x$ var en lineær afbildning. Det er faktisk et interessant særtilfælde af Definition 10.1.1: hvis vi vælger $n = m = 1$, $\mathbb{F} = \mathbb{R}$ og $\mathbf{A} = [2]$ i Definition 10.1.1, får vi funktionen h. I stedet for $\mathbf{A} = [2]$ kunne vi også blot have skrevet $\mathbf{A} = 2$. Når man angiver en 1×1-matrix, er det almindeligt at udelade parenteserne [ ].

Eksempel 10.1.1
Lad $\mathbb{F} = \mathbb{R}$, og vælg

$$\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 0 & 1/2 \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

Yderligerer defineres

$$\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{og} \quad \mathbf{v} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

Spørgsmål: Bestem $L_{\mathbf{A}}(\mathbf{u})$, $L_{\mathbf{A}}(\mathbf{v})$ og $L_{\mathbf{A}}((1/2)\mathbf{u}+\mathbf{v})$ med $\mathbf{A}$, $\mathbf{u}$ og $\mathbf{v}$ givet som ovenfor.

Svar: Vi har

$$L_{\mathbf{A}}(\mathbf{u}) = \mathbf{A}\cdot\mathbf{u} = \begin{bmatrix} -1 & 0 \\ 0 & 1/2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} (-1)\cdot 2 + 0\cdot 1 \\ 0\cdot 2 + (1/2)\cdot 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1/2 \end{bmatrix}$$

og tilsvarende

$$L_{\mathbf{A}}(\mathbf{u}) = \mathbf{A}\cdot\mathbf{v} = \begin{bmatrix} -1 & 0 \\ 0 & 1/2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} (-1)\cdot 1 + 0\cdot 3 \\ 0\cdot 1 + (1/2)\cdot 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 3/2 \end{bmatrix}.$$

Egentlig kan vi udregne $L_{\mathbf{A}}((1/2)\mathbf{u}+\mathbf{v})$ på forskellige måder. Den mest direkte metode er først at beregne vektoren $(1/2)\mathbf{u}+\mathbf{v}$ og derefter beregne $\mathbf{A}\cdot((1/2)\mathbf{u}+\mathbf{v})$. Gøres dette, får vi

$$(1/2)\mathbf{u}+\mathbf{v} = (1/2)\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 7/2 \end{bmatrix}$$

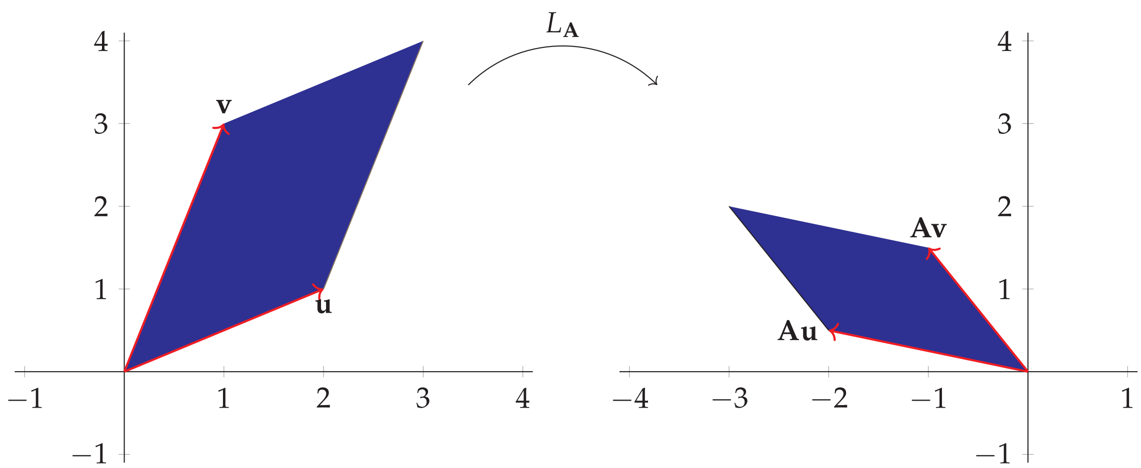
og derfor

$$L_{\mathbf{A}}((1/2)\mathbf{u}+\mathbf{v}) = L_{\mathbf{A}}\left(\begin{bmatrix} 2 \\ 7/2 \end{bmatrix}\right) = \begin{bmatrix} -1 & 0 \\ 0 & 1/2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 7/2 \end{bmatrix} = \begin{bmatrix} -2 \\ 7/4 \end{bmatrix}.$$

En anden metode til beregning af $L_{\mathbf{A}}((1/2)\mathbf{u}+\mathbf{v})$ er at udnytte, at $L_{\mathbf{A}}$ er en lineær afbildning, og at vi allerede har udregnet $L_{\mathbf{A}}(\mathbf{u})$ og $L_{\mathbf{A}}(\mathbf{v})$:

$$\begin{aligned} L_{\mathbf{A}}((1/2)\mathbf{u}+\mathbf{v}) &= L_{\mathbf{A}}((1/2)\mathbf{u}) + L_{\mathbf{A}}(\mathbf{v}) = (1/2)L_{\mathbf{A}}(\mathbf{u}) + L_{\mathbf{A}}(\mathbf{v}) \\ &= (1/2)\begin{bmatrix} -2 \\ 1/2 \end{bmatrix} + \begin{bmatrix} -1 \\ 3/2 \end{bmatrix} = \begin{bmatrix} -2 \\ 7/4 \end{bmatrix}. \end{aligned}$$

Se Figur 10.1 for en illustration af billederne af vektorerne **u** og **v** samt billedet af vektorer, der ligger i et parallelogram, hvor to af siderne er de givne vektorer **u** og **v**.



Figur 10.1: Eksempel på en lineær afbildning givet ved en 2×2-matrix.

Eksempel 10.1.2

Som i forrige eksempel lader vi $\mathbb{F} = \mathbb{R}$ og definerer

$$\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{og} \quad \mathbf{v} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

Denne gang vælger vi matricen

$$\mathbf{A} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

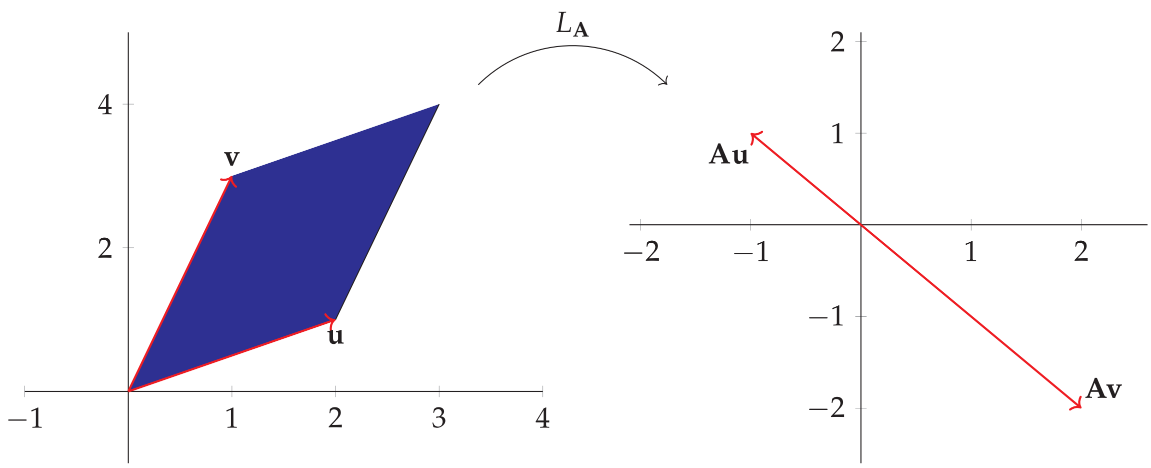
I dette tilfælde har vi

$$L_{\mathbf{A}}(\mathbf{u}) = \mathbf{A}\cdot\mathbf{v} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

og

$$L_{\mathbf{A}}(\mathbf{u}) = \mathbf{A}\cdot\mathbf{v} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}.$$

En illustration er givet i Figur 10.2. Denne gang er billedet af det blå område (parallelogrammet) blot linjen, der forbinder $(-1,1)$ og $(2,-2)$. Det er derfor ikke synligt på figuren, da det er skjult bag tegningen af vektorerne $\mathbf{Au}$ og $\mathbf{Av}$.



Figur 10.2: Endnu et eksempel på en lineær afbildning givet ved en 2×2-matrix.

Eksempel 10.1.3

Lad $\mathbb{F} = \mathbb{R}$, og vælg

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \in \mathbb{R}^{2\times 4}.$$

Den tilsvarende lineære afbildning $L_{\mathbf{A}} : \mathbb{R}^4 \to \mathbb{R}^2$ fungerer som følger:

$$L_{\mathbf{A}}\left(\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}\right) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} v_1 + v_2 + v_3 + v_4 \\ v_1 + 2v_2 + 3v_3 + 4v_4 \end{bmatrix}.$$

Så har vi for eksempel

$$L_{\mathbf{A}}\left(\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad L_{\mathbf{A}}\left(\begin{bmatrix} 0 \\ 2 \\ -1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 5 \end{bmatrix} \quad \text{og} \quad L_{\mathbf{A}}\left(\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Som Eksempel 10.1.3 viser, kan det godt ske, at en lineær afbildning $L_{\mathbf{A}}$ afbilder en vektor til nulvektoren. Mængden af sådanne vektorer har et særligt navn:

Definition 10.1.2

Lad $\mathbb{F}$ være et legeme og $\mathbf{A} \in \mathbb{F}^{m\times n}$ en matrix. Da er *kernen* af matricen $\mathbf{A}$, betegnet ved $\ker \mathbf{A}$, følgende mængde af vektorer:

$$\ker \mathbf{A} = \{\mathbf{v} \in \mathbb{F}^n \mid \mathbf{A}\cdot\mathbf{v} = \mathbf{0}\}.$$

Bemærk, at man på en ækvivalent måde kan definere kernen af en matrix $\mathbf{A}$ som alle vektorer fra $\mathbb{F}^n$, der afbildes til nulvektoren af den lineære afbildning $L_{\mathbf{A}}$. Vi kan også tænke på vektorerne i kernen som netop de vektorer, der er løsninger til det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A}$.

Bemærkning 10.1.1

En bemærkning om terminologien er på sin plads her: nogle forfattere foretrækker at bruge ordet *nulrum* for det, vi kalder kernen af en matrix. På engelsk er den tilsvarende betegnelse *null space*, mens andre mulige engelske betegnelser er *right kernel* eller *right null space*. Årsagen til ordet "right", altså "højre", er, at matricen er ganget med en søjlevektor fra højre. Man kunne også have taget udgangspunkt i mængden af rækkevektorer $\mathbf{u} \in \mathbb{F}^{1\times m}$, således at $\mathbf{u}\cdot\mathbf{A} = \mathbf{0}$. Denne mængde kan på engelsk kaldes *left kernel* eller *left null space*.

En af grundene til, at vi her har introduceret begrebet kerne af en matrix, er, at det faktisk er et underrum. Lad os vise dette i følgende lemma.

Lemma 10.1.2

Lad $\mathbb{F}$ være et legeme og $\mathbf{A} \in \mathbb{F}^{m\times n}$ en matrix. Da er kernen af $\mathbf{A}$ et underrum i $\mathbb{F}^n$.

Bevis. Bemærk først og fremmest, at $\mathbf{0} \in \ker \mathbf{A}$, hvilket viser, at kernen af $\mathbf{A}$ ikke er den tomme mængde. Dette betyder, at hvis vi sætter $W = \ker \mathbf{A}$, så er kravet om, at W ikke er tomt i Lemma 9.3.2, opfyldt.

Lad $\mathbf{u}, \mathbf{v} \in \ker \mathbf{A}$ og $c \in \mathbb{F}$. Da har vi

$$\mathbf{A}\cdot(\mathbf{u} + c\cdot\mathbf{v}) = \mathbf{A}\cdot\mathbf{u} + \mathbf{A}\cdot(c\cdot\mathbf{v}) = \mathbf{A}\cdot\mathbf{u} + c\cdot(\mathbf{A}\cdot\mathbf{v}) = \mathbf{0} + c\cdot\mathbf{0} = \mathbf{0}. \tag{10.3}$$

Her har vi benyttet, at $\mathbf{A}\cdot\mathbf{u} = \mathbf{0}$, og $\mathbf{A}\cdot\mathbf{v} = \mathbf{0}$, da $\mathbf{u}, \mathbf{v} \in \ker \mathbf{A}$. Ligning (10.3) viser, at $\mathbf{u} + c\cdot\mathbf{v} \in \ker \mathbf{A}$. Dermed medfører Lemma 9.3.2, at $\ker \mathbf{A}$ er et underrum i $\mathbb{F}^n$. □

Dimensionen af $\ker \mathbf{A}$ betegnes $\mathrm{null}\mathbf{A}$ grundet den engelske betegnelse *nullity* for dimensionen af kernen.

Eksempel 10.1.4

Lad $\mathbb{F} = \mathbb{R}$, og betragt matricen

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix}.$$

Bestem en basis for $\ker \mathbf{A}$, og beregn dimensionen af kernen af $\mathbf{A}$.

Svar: Kernen af $\mathbf{A}$ består af alle vektorer $\mathbf{v} = (v_1, v_2, v_3, v_4) \in \mathbb{R}^4$, der opfylder $\mathbf{A} \cdot \mathbf{v} = \mathbf{0}$. Vi har for eksempel set i Eksempel 10.1.3, at vektoren $(1, -1, -1, 1)$ afbildes i $(0, 0)$ af den lineære afbildning $L_{\mathbf{A}}$. Derfor er $(1, -1, -1, 1) \in \ker \mathbf{A}$.

Vi kan tænke på vektorerne i kernen som netop de vektorer, der er løsninger til det homogene system af to lineære ligninger med koefficientmatrix $\mathbf{A}$. For at beskrive alle disse løsninger, følger vi den samme procedure som forklaret i Eksempel 6.4.3 og Sætning 6.4.4. Derfor starter vi med at bringe matrix $\mathbf{A}$ på reduceret trappeform:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \underset{R_2 \leftarrow R_2 - R_1}{\longrightarrow} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \underset{R_1 \leftarrow R_1 - R_2}{\longrightarrow} \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \end{bmatrix}.$$

Nu kan vi se, at $\mathbf{v} = (v_1, v_2, v_3, v_4) \in \ker \mathbf{A}$, hvis og kun hvis $v_1 - v_3 - 2v_4 = 0$, og $v_2 + 2v_3 + 3v_4 = 0$. Ligesom i Eksempel 6.4.3 (eller direkte ved hjælp af Sætning 6.4.4) ser vi, at

$$\ker \mathbf{A} = \left\{ t_1 \cdot \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix} \mid t_1, t_2 \in \mathbb{R} \right\}.$$

Vektorerne

$$\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} \quad \text{og} \quad \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix}$$

udspænder dermed $\ker \mathbf{A}$. Faktisk fortæller Korollar 9.3.4, at disse to vektorer danner en basis for $\ker \mathbf{A}$. Derfor er en basis for $\ker \mathbf{A}$ givet ved

$$\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Da denne basis består af to vektorer, konkluderer vi, at dimensionen af kernen af $\mathbf{A}$ er to. Med andre ord: $\mathrm{null}\mathbf{A} = 2$.

Vi har allerede set, at vi kan betragte vektorerne i kernen som netop de vektorer, der er løsninger til det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A}$. Ved anvendelse af Korollar 9.3.4 opnår vi følgende resultat, som ofte kaldes *dimensionssætningen for matricer*.

Sætning 10.1.3

Lad $\mathbb{F}$ være et legeme og $\mathbf{A} \in \mathbb{F}^{m \times n}$ en matrix. Da gælder der, at

$$\rho(\mathbf{A}) + \mathrm{null}(\mathbf{A}) = n,$$

hvor $\rho(\mathbf{A})$ betegner rangen af matricen $\mathbf{A}$, og $\mathrm{null}(\mathbf{A})$ betegner dimensionen af kernen af matricen $\mathbf{A}$.

Bevis. Ved at anvende Korollar 9.3.4 ser vi, at kernens basis indeholder præcist $n - \rho(\mathbf{A})$ vektorer. Derfor er $\mathrm{null}(\mathbf{A}) = \dim \ker(\mathbf{A}) = n - \rho(\mathbf{A})$. Dette medfører, at $\rho(\mathbf{A}) + \mathrm{null}(\mathbf{A}) = n$. □

Bemærk, at dimensionssætningen for matricer på engelsk kaldes *the rank-nullity theorem for matrices*, hvilket afslører sætningens indhold. Vi har set i Lemma 10.1.2, at kernen af en matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ er et underrum i $\mathbb{F}^n$. Med andre ord: $\ker \mathbf{A}$ er et underrum i definitionsmængden for den lineære afbildning $L_\mathbf{A} : \mathbb{F}^n \to \mathbb{F}^m$. Til en matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ kan man også knytte et underrum i $\mathbb{F}^m$, dispositionsmængden for den lineære afbildning $L_\mathbf{A} : \mathbb{F}^n \to \mathbb{F}^m$. Dette gøres i følgende definition.

Definition 10.1.3

Lad $\mathbb{F}$ være et legeme og $\mathbf{A} \in \mathbb{F}^{m \times n}$ en matrix. Da er *søjlerummet* af matricen $\mathbf{A}$, betegnet ved $\mathrm{colsp}\mathbf{A}$, det underrum i $\mathbb{F}^m$, der er udspændt af søjlerne i $\mathbf{A}$. Dimensionen af søjlerummet af en matrix $\mathbf{A}$ kaldes *søjlerangen* af $\mathbf{A}$.

Bemærk, at begreberne søjlerum og søjlerang på engelsk kaldes *column space*, henholdsvis *column rank*.

Lemma 10.1.4

Lad $\mathbb{F}$ være et legeme og $\mathbf{A} \in \mathbb{F}^{m \times n}$ en matrix. Da er $\mathrm{colsp}\mathbf{A}$, søjlerummet af matricen $\mathbf{A}$, netop billedet af den lineære afbildning $L_\mathbf{A} : \mathbb{F}^n \to \mathbb{F}^m$.

Bevis. Et element fra søjlerummet af en matrix $\mathbf{A}$ er per definition en linearkombination af søjlerne i $\mathbf{A}$. På den anden side er et element af billedet af den lineære afbildning $L_\mathbf{A} : \mathbb{F}^n \to \mathbb{F}^m$ på formen $\mathbf{A} \cdot \mathbf{v}$ for en vektor $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{F}^n$. Ved at anvende

Definition 7.2.1 kan vi omskrive dette som en linearkombination af søjlerne i $\mathbf{A}$ som følger:

$$\mathbf{A}\cdot\mathbf{v} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = v_1 \cdot \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots v_n \cdot \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Derfor består billedet af den lineære afbildning $L_{\mathbf{A}}$ netop af alle linearkombinationer af søjlerne i $\mathbf{A}$. Og dette er netop søjlerummet af matricen $\mathbf{A}$. □

Bemærkning 10.1.2

På grund af Lemma 10.1.4 kaldes søjlerummet af en matrix $\mathbf{A}$ også ofte for *billedet* eller *billedrummet* af $\mathbf{A}$.

Vi har tidligere introduceret rangen af en matrix i Definition 6.3.2. Rangen $\rho(\mathbf{A})$ af en matrix $\mathbf{A}$, som defineret i Definition 6.3.2, kaldes nogle gange mere korrekt *rækkerangen* af matricen $\mathbf{A}$, da man kan vise, at dimensionen af vektorrummet udspændt af rækkerne i $\mathbf{A}$ er lig med $\rho(\mathbf{A})$. Det viser sig dog, at for enhver matrix er dens rækkerang og søjlerang ens. Derfor vil vi blot kalde søjlerangen af en matrix $\mathbf{A}$ for rangen af matricen og på simpel vis betegne den ved $\rho(\mathbf{A})$ med samme notation som i Definition 6.3.2.

Det er ikke indlysende fra Definitionerne 6.3.2 og 10.1.3, at rækkerang og søjlerang af en matrix altid er de samme. En læser, der er villig til at acceptere, at dette er tilfældet, kan springe resten af dette afsnit over, men for den interesserede læser giver vi herunder et kort bevis af, hvorfor rækkerang og søjlerang altid er lig med hinanden.

Sætning 10.1.5

Lad $\mathbb{F}$ være et legeme og $\mathbf{A} \in \mathbb{F}^{m\times n}$ en matrix. Da er rækkerangen og søjlerangen af matricen $\mathbf{A}$ ens.

Bevis. Rækkerangen $\rho(\mathbf{A})$ af en matrix $\mathbf{A}$ er per definition lig med antallet af pivoter i $\mathbf{A}$'s reducerede trappeform. Sætning 9.3.3 medfører dog, at antallet af vektorer i en basis for søjlerummet af $\mathbf{A}$ også er lig med dette antal pivoter. Derfor er dimensionen af søjlerummet af $\mathbf{A}$ også lig med $\rho(\mathbf{A})$. □

Eksempel 10.1.5

I dette eksempel ønsker vi at beregne rangen af matricen $\mathbf{A}$ fra Eksempel 10.1.1 samt dimensionen af dens kerne. Vi har altså at gøre med

$$\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 0 & 1/2 \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

For at bestemme kernen af $\mathbf{A}$ kan man i princippet følge samme procedure som i Eksempel 10.1.4, men da vi her har at gøre med en kvadratisk matrix, vælger vi en lidt anderledes tilgang og anvender determinanter.

Først og fremmest er $\det \mathbf{A} = (-1) \cdot 1/2 = -1/2$. Specifikt er $\det \mathbf{A} \neq 0$. Derfor medfører Korollar 8.3.6, at kernen af matricen $\mathbf{A}$ kun indeholder nulvektoren, det vil sige $\ker \mathbf{A} = \{\mathbf{0}\}$. Derfor er $\text{null}(\mathbf{A}) = \dim\{\mathbf{0}\} = 0$ i dette eksempel.

Dimensionssætningen (Sætning 10.1.3) medfører så, at matricen $\mathbf{A}$ har rang $\rho(\mathbf{A}) = 2 - 0 = 2$. Vi kunne også have beregnet rangen ved hjælp af Korollar 8.3.5: da $\det \mathbf{A} \neq 0$, medfører denne korollar, at de to søjler i $\mathbf{A}$ er lineært uafhængige (man kunne også verificere dette direkte). Derfor har $\mathbf{A}$ søjlerang to, og så medfører Sætning 10.1.5, at $\rho(\mathbf{A}) = 2$.

Eksempel 10.1.6

I dette eksempel ønsker vi at beregne rangen af matricen $\mathbf{A}$ fra Eksempel 10.1.2 samt dimensionen af dens kerne, så vi har at gøre med

$$\mathbf{A} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

Bemærk, at $\det \mathbf{A} = 0$ denne gang. Derfor er søjlerne i $\mathbf{A}$ ikke lineært uafhængige ifølge Korollar 8.3.5. Faktisk er summen af søjlerne i $\mathbf{A}$ lig nulvektoren, hvilket bekræfter, at de er lineært afhængige. Vi konkluderer, at

$$\text{colsp}\mathbf{A} = \text{Span}\left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\} = \text{Span}\left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}.$$

Da vektoren $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ ikke er nulvektoren, konkluderer vi, at dimensionen af $\text{colsp}\mathbf{A}$ er én. Specifikt har vi $\rho(\mathbf{A}) = 1$ (som i det forrige eksempel benyttede vi Sætning 10.1.5). Dimensionssætningen medfører, at $\text{null}(\mathbf{A}) = 2 - 1 = 1$.

Bemærk, at Figur 10.2 intuitivt antyder, at billedet af den lineære afbildning $L_{\mathbf{A}} : \mathbb{R}^2 \to \mathbb{R}^2$ er endimensionelt. Ifølge Lemma 10.1.4 er billedet af $L_{\mathbf{A}}$ det samme som søjlerummet af $\mathbf{A}$, så intuitionen stemmer.

10.2 Lineære afbildninger mellem generelle vektorrum

Forrige afsnit fokuserede på lineære afbildninger, der baseres på matricer, men Definition 10.0.1 tillader langt mere generelle lineære afbildninger. Det viser sig, at begreberne kerne og billede også giver mening i den generelle ramme. Lad os først gennemgå nogle flere eksempler.

Eksempel 10.2.1

Lad $\mathbb{F} = \mathbb{C}$, og betragt det komplekse vektorrum $\mathbb{C}[Z]$ (se Eksempel 9.1.5). Husk på, at $\mathbb{C}[Z]$ betegner mængden af alle polynomier med koefficienter i $\mathbb{C}$. Betragt nu afbildningen

$D : \mathbb{C}[Z] \to \mathbb{C}[Z]$ defineret ved $D(a_0 + a_1 Z + a_2 Z^2 + \cdots + a_n Z^n) = a_1 + 2a_2 Z + \cdots + n a_n Z^{n-1}$. Med andre ord sender afbildningen D et polynomium $p(Z)$ til dets afledte $p(Z)'$. Man kan vise, at D er en lineær afbildning.

Eksempel 10.2.2

Lad $V_1 = \mathbb{F}^{n \times n}$ og $V_2 = \mathbb{F}$, og lad en kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$ være givet. Da er *sporet*, betegnet ved $\mathrm{Tr}(\mathbf{A})$, defineret som summen af diagonalelementerne. Med andre ord:

$$\mathrm{Tr}\left(\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}\right) = a_{11} + \cdots + a_{nn}.$$

Betegnelsen Tr er en forkortelse for det engelske begreb for sporet af en matrix: *trace*.

Spørgsmål: Er afbildningen $\mathrm{Tr} : \mathbb{F}^{n \times n} \to \mathbb{F}$ defineret ved $\mathbf{A} \mapsto \mathrm{Tr}(\mathbf{A})$ en lineær afbildning?

Svar: For at finde ud af, om sporafbildningen $\mathrm{Tr} : \mathbb{F}^{n \times n} \to \mathbb{F}$ som defineret ovenfor, er lineær, tjekker vi, om alle betingelser i Definition 10.0.1 er opfyldt. Ved at anvende notationen fra Definition 10.0.1 har vi $V_1 = \mathbb{F}^{n \times n}$ og $V_2 = \mathbb{F}$. Vi bør først tjekke, at disse er vektorrum over et legeme $\mathbb{F}$. Det viser sig, at begge er vektorrum over $\mathbb{F}$, hvilket for V_1 ses i Eksempel 9.1.4 med $m = n$, og hvilket for V_2 ses i Eksempel 9.1.1 med $n = 1$.

Nu skal vi tjekke, om Tr opfylder de to betingelser fra Definition 10.0.1. Lad os vælge vilkårlige $c \in \mathbb{F}$ og $\mathbf{u}, \mathbf{v} \in \mathbb{F}^{n \times n}$. Derfor kan vi skrive

$$\mathbf{u} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \quad \text{og} \quad \mathbf{v} = \begin{bmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nn} \end{bmatrix}.$$

Da gælder

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} + \begin{bmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{n1} + b_{n1} & \dots & a_{nn} + b_{nn} \end{bmatrix}$$

og

$$c \cdot \mathbf{u} = c \cdot \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} c \cdot a_{11} & \dots & c \cdot a_{1n} \\ \vdots & & \vdots \\ c \cdot a_{n1} & \dots & c \cdot a_{nn} \end{bmatrix}.$$

Derfor er

$$\mathrm{Tr}(\mathbf{u} + \mathbf{v}) = a_{11} + b_{11} + \cdots a_{nn} + b_{nn} = a_{11} + \cdots a_{nn} + b_{11} + \cdots b_{nn} = \mathrm{Tr}(\mathbf{u}) + \mathrm{Tr}(\mathbf{v})$$

og

$$\mathrm{Tr}(c \cdot \mathbf{u}) = c \cdot a_{11} + \cdots c \cdot a_{nn} = c \cdot (a_{11} + \cdots a_{nn}) = c \cdot \mathrm{Tr}(\mathbf{u}).$$

Vi konkluderer, at $\mathrm{Tr} : \mathbb{F}^{n \times n} \to \mathbb{F}$ defineret ved $\mathbf{A} \mapsto \mathrm{Tr}(\mathbf{A})$ er en lineær afbildning.

Eksempel 10.2.3

Lad $\mathbb{F} = \mathbb{R}$, og betragt afbildningen $m_5 : \mathbb{R}^2 \to \mathbb{R}^2$ defineret ved $m_5(v_1, v_2) = (5v_1, 5v_2)$. Med andre ord, effekten af afbildning m_5 på en vektor er, at vektoren ganges med skalaren 5. Visuelt betyder dette, at retningen af en vektor ikke ændres, men at den bliver fem gange længere. Man kan vise, at dette er en lineær afbildning mellem reelle vektorrum.

Mere generelt kan man vise, at hvis $\mathbb{F}$ er et legeme, og $c \in \mathbb{F}$ er en skalar, så er afbildningen $m_c : \mathbb{F}^n \to \mathbb{F}^n$ defineret ved $m_c(\mathbf{u}) = c \cdot \mathbf{u}$ en lineær afbildning mellem vektorrum.

Eksempel 10.2.4

Betragt for $\alpha \in \mathbb{R}$ afbildningen $R_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ defineret ved $R_\alpha(v_1, v_2) = (\cos(\alpha) \cdot v_1 - \sin(\alpha) \cdot v_2, \sin(\alpha) \cdot v_1 + \cos(\alpha) \cdot v_2)$. Geometrisk er effekten af R_α på $(v_1, v_2) \in \mathbb{R}^2$ en rotation mod uret om $(0, 0)$ med vinklen α. For eksempel, hvis $\alpha = \pi/2$, så er $R_{\pi/2}(v_1, v_2) = (-v_2, v_1)$. Man kan vise, at R_α er en lineær afbildning.

Eksempel 10.2.5

Vi vælger $\mathbb{F} = \mathbb{C}$. Lad V_1 være mængden af polynomier i $\mathbb{C}[Z]$ af grad højst tre, og lad V_2 være mængden af polynomier i $\mathbb{C}[Z]$ af grad højst fire. Både V_1 og V_2 er vektorrum over $\mathbb{C}$. En mulig basis for V_1 er givet ved mængden $\{1, Z, Z^2, Z^3\}$, mens en basis for V_2 er $\{1, Z, Z^2, Z^3, Z^4\}$. Derfor er $\dim V_1 = 4$ og $\dim V_2 = 5$. Definér nu afbildningen $L : V_1 \to V_2$ ved $p(Z) \mapsto (i + 2Z) \cdot p(Z)$. Bemærk, at vi for ethvert $p(Z) \in V_1$ har $(i + 2Z) \cdot p(Z) \in V_2$ ved brug af Ligning (4.1). Man kan vise, at L er en lineær afbildning. Hvis $p_1(Z), p_2(Z) \in V_1$, og $c \in \mathbb{C}$, så gælder der nemlig, at

$$\begin{aligned} L(p_1(Z) + p_2(Z)) &= (i + 2Z) \cdot (p_1(Z) + p_2(Z)) \\ &= (i + 2Z) \cdot p_1(Z) + (i + 2Z) \cdot p_2(Z) \\ &= L(p_1(Z)) + L(p_2(Z)) \end{aligned}$$

og

$$L(c \cdot p_1(Z)) = (i + 2Z) \cdot c \cdot p_1(Z) = c \cdot (i + 2Z) \cdot p_1(Z) = c \cdot L(p_1(Z)).$$

Eksempel 10.2.6

Som et afsluttende eksempel på en lineær afbildning betragter vi afbildningen $\mathrm{ev} : \mathbb{C}[Z] \to \mathbb{C}^2$ defineret ved $p(Z) \mapsto (p(0), p(1))$. For eksempel er $\mathrm{ev}(Z^2 + Z + 1) = (0^2 + 0 + 1, 1^2 + 1 + 1) = (1, 3)$. Man kan vise, at ev er en lineær afbildning.

Vi afslutter dette afsnit med nogle generelle egenskaber ved lineære afbildninger. Først ser vi på sammensætningen af to lineære afbildninger. Det vil være en fordel, hvis læseren først genopfrisker Afsnit 2.2, hvor vi definerede sammensætningen af to funktioner.

Sætning 10.2.1

Lad $\mathbb{F}$ være et legeme og V_1, V_2, V_3 vektorrum over $\mathbb{F}$. Antag yderligere, at $L_1 : V_1 \to V_2$, og $L_2 : V_2 \to V_3$ er lineære afbildninger. Da er sammensætningen $L_2 \circ L_1 : V_1 \to V_3$ også en lineær afbildning.

Bevis. Lad os vælge vilkårlige $\mathbf{u}, \mathbf{v} \in V_1$ og $c \in \mathbb{F}$. Ved at udnytte lineariteten af L_1 og L_2 samt definitionen af sammensætningen af to funktioner får vi

$$\begin{aligned}(L_2 \circ L_1)(\mathbf{u}+\mathbf{v}) &= L_2(L_1(\mathbf{u}+\mathbf{v})) \\ &= L_2(L_1(\mathbf{u}) + L_1(\mathbf{v})) \\ &= L_2(L_1(\mathbf{u})) + L_2(L_1(\mathbf{v})) \\ &= (L_2 \circ L_1)(\mathbf{u}) + (L_2 \circ L_1)(\mathbf{v})\end{aligned}$$

og

$$(L_2 \circ L_1)(c \cdot \mathbf{u}) = L_2(L_1(c \cdot \mathbf{u})) = L_2(c \cdot L_1(\mathbf{u})) = c \cdot L_2(L_1(\mathbf{u})) = c \cdot (L_2 \circ L_1)(\mathbf{u}).$$

Ifølge Definition 10.0.1 er afbildningen $L_2 \circ L_1 : V_1 \to V_3$ dermed en lineær afbildning. □

Da enhver funktion $f : A \to B$ har et billede, nemlig mængden $\text{image}(f) = \{f(a) \mid a \in A\}$ som behandlet i Afsnit 2.2, har en lineær afbildning $L : V_1 \to V_2$ også et billede. I lyset af Lemma 10.1.4 generaliserer dette ideen om søjlerummet af en matrix til rammerne af generelle lineære afbildninger. Man kan vise, at billedet af en lineær afbildning $L : V_1 \to V_2$ er et underrum i V_2. Begrebet kerne kan også generaliseres.

Definition 10.2.1

Lad $\mathbb{F}$ være et legeme, V_1 og V_2 vektorrum over $\mathbb{F}$ og $L : V_1 \to V_2$ en lineær afbildning. Da er *kernen* af afbildningen L:

$$\ker L = \{\mathbf{v} \in V_1 \mid L(\mathbf{v}) = \mathbf{0}\}.$$

På samme måde som for kernen af en matrix kan man vise, at kernen af en lineær afbildning $L : V_1 \to V_2$ er et underrum i V_1.

Eksempel 10.2.7

Lad os genbesøge Eksempel 10.2.1. Vi betragtede den lineære afbildning $D : \mathbb{C}[Z] \to \mathbb{C}[Z]$, der sender et polynomium $p(Z)$ til dets afledte. De eneste polynomier, hvis afledte er 0, er konstante polynomier, det vil sige polynomier på formen $p(Z) = a_0$. Derfor er $\ker D = \{a_0 \mid a_0 \in \mathbb{C}\} = \mathbb{C}$. Bemærk, at $\{1\}$ er en basis for $\ker D$, så vi kan konkludere, at $\dim \ker D = 1$.

Da denne teori senere vil blive anvendt i emnet om differentialligninger, tager vi her endnu et eksempel, der involverer afledte.

Eksempel 10.2.8

Lad $C_\infty(\mathbb{R})$ være vektorrummet af alle uendeligdifferentiable funktioner fra $\mathbb{R}$ til $\mathbb{R}$, som behandlet i Eksempel 9.3.4. Lad os betragte afbildningen $L : C_\infty(\mathbb{R}) \to C_\infty(\mathbb{R})$, hvor

$f \mapsto f' - f$. Som sædvanlig betegner f' den afledte af funktionen f. Da $f \in C_\infty(\mathbb{R})$, er også f' uendeligdifferentiabel, således at $f' \in C_\infty(\mathbb{R})$. Man kan vise, at L er en lineær afbildning. Ved at bruge Definition 10.2.1 ser vi, at $\ker L = \{f \in C_\infty(\mathbb{R}) \mid f' - f = 0\}$. Med andre ord: kernen af L består af de funktioner $f \in C_\infty(\mathbb{R})$, der opfylder, at den afledte af f er den samme som f selv. Med endnu andre ord: kernen af L består af alle funktioner i $C_\infty(\mathbb{R})$, som er løsninger til differentialligningen $f' = f$. Et eksempel på en funktion, der opfylder denne differentialligning, er eksponentialfunktionen $\exp : \mathbb{R} \to \mathbb{R}$ defineret ved $t \mapsto e^t$. Også alle skaleringer $f = c \cdot \exp$ med $c \in \mathbb{R}$ er løsninger til differentialligningen $f' = f$. Det er muligt at vise, at der ikke er flere løsninger i $C_\infty(\mathbb{R})$ end disse. Derfor er $\ker L$ et endimensionelt underrum i $C_\infty(\mathbb{R})$ med basis $\{\exp\}$.

Bemærkning 10.2.1

Eksponentialfunktionen blev også diskuteret i Eksempel 2.3.2, men dens dispositionsmængde blev dér defineret til $\mathbb{R}_{\geq 0}$. Strengt taget er eksponentialfunktionen fra Eksempel 2.3.2 derfor ikke den samme funktion som eksponentialfunktionen i dette eksempel. Da begge funktioner afbilder ethvert $x \in \mathbb{R}$ til nøjagtig den samme værdi, nemlig e^x, er det dog en smule overflødigt at anvende forskellige notationssystemer for disse funktioner. Af denne grund vil vi gennemgående betegne begge funktioner ved exp.

Eksempel 10.2.9

Som et sidste eksempel på kernen af en lineær afbildning tager vi et kig på afbildningen ev fra Eksempel 10.2.6. Afbildningen $\mathrm{ev} : \mathbb{C}[Z] \to \mathbb{C}^2$ blev dér defineret ved $p(Z) \mapsto (p(0), p(1))$. Derfor har vi

$$\ker \mathrm{ev} = \{p(Z) \in \mathbb{C}[Z] \mid (p(0), p(1)) = (0, 0)\} = \{p(Z) \in \mathbb{C}[Z] \mid p(0) = 0 \ \wedge \ p(1) = 0\}.$$

Det er muligt at beskrive kernen af ev mere specifikt. Lad os starte med at beskrive mængden af polynomier $p(Z)$, der opfylder $p(0) = 0$, altså alle de polynomier $p(Z)$, der har værdien 0 som en rod. Ved at benytte Lemma 4.6.2 konkluderer vi, at

$$\{p(Z) \in \mathbb{C}[Z] \mid p(0) = 0\} = \{Z \cdot q(Z) \mid q(Z) \in \mathbb{C}[Z]\}.$$

Hvis både $p(0) = 0$, og $p(1) = 0$, så har vi, at $p(Z) = Z \cdot q(Z)$ for et $q(Z) \in \mathbb{C}[Z]$, og $p(1) = 0$. Men dette er ækvivalent med at sige, at $p(Z) = Z \cdot q(Z)$ for et $q(Z) \in \mathbb{C}[Z]$, og $q(1) = 0$. Anvendes Lemma 4.6.2 igen men nu for $q(Z)$ og roden 1 ser vi, at $q(Z) = (Z - 1) \cdot s(Z)$ for et $s(Z) \in \mathbb{C}[Z]$. Derfor får vi, at $p(Z) \in \ker \mathrm{ev}$, hvis og kun hvis $p(Z) = Z \cdot (Z - 1) \cdot s(Z)$ for et $s(Z) \in \mathbb{C}[Z]$. Vi konkluderer derfor, at

$$\ker \mathrm{ev} = \{p(Z) \in \mathbb{C}[Z] \mid p(0) = 0 \ \wedge \ p(1) = 0\} = \{Z \cdot (Z - 1) \cdot s(Z) \mid s(Z) \in \mathbb{C}[Z]\}.$$

10.3 Lineære afbildninger mellem endeligdimensionelle vektorrum

Lad os antage, at vi er givet et endeligdimensionelt vektorrum V over et legeme $\mathbb{F}$, hvor vi for eksempel har $\dim V = n$. Da kan vi vælge en ordnet basis for V, som vi kan kalde $\beta = (\mathbf{v}_1, \dots, \mathbf{v}_n)$, hvor $\mathbf{v}_1, \dots, \mathbf{v}_n$ er lineært uafhængige vektorer i V. Som vi har set i Definition 9.2.4, kan vi for alle $\mathbf{v} \in V$ producere en unik koordinatvektor $[\mathbf{v}]_\beta \in \mathbb{F}^n$. Dette betyder, at vi kan definere en funktion $\phi_\beta : V \to \mathbb{F}^n$ ved $\mathbf{v} \mapsto [\mathbf{v}]_\beta$. Ved at kombinere Lemma 9.2.2 og Definition 10.0.1 konkluderer vi straks, at funktionen ϕ_β er en lineær afbildning. For en givet vektor $(c_1, \dots, c_n) \in \mathbb{F}^n$ er det nemt at opskrive en vektor i V, der har $(c_1, \dots, c_n)$ som sin koordinatvektor (med hensyn til basis β). Denne vektor er nemlig netop vektoren $\mathbf{v} = c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n$, og denne er ifølge Lemma 9.2.1 den eneste vektor, der har koordinaterne $(c_1, \dots, c_n)$! Det, vi rent faktisk har opnået her, er den inverse funktion af ϕ_β. Lad os formulere disse udsagn i et lemma og give en fuldstændig bevis.

Lemma 10.3.1

Lad $\mathbb{F}$ være et legeme, V et vektorrum over $\mathbb{F}$ med dimension n og $\beta = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ en ordnet basis for V. Da er funktionen $\phi_\beta : V \to \mathbb{F}^n$ defineret ved $\mathbf{v} \mapsto [\mathbf{v}]_\beta$ en lineær afbildning. Desuden er funktionen $\psi_\beta : \mathbb{F}^n \to V$ defineret ved $(c_1, \dots, c_n) \mapsto c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n$ den inverse af ϕ_β og også en lineær afbildning.

Bevis. Vi har allerede vist i diskussionen forinden dette lemma, at $\phi_\beta : V \to \mathbb{F}^n$ er en lineær afbildning mellem vektorrum over $\mathbb{F}$. Lad os nu ved $\psi_\beta : \mathbb{F}^n \to V$ betegne den afbildning, der er defineret ved $(c_1, \dots, c_n) \mapsto c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n$. Vi skal først sikre os, at ψ_β er den inverse funktion ϕ_β^{-1}, hvilket vil betyde, at $\psi_\beta \circ \phi_\beta(\mathbf{v}) = \mathbf{v}$ for alle $\mathbf{v} \in V$, samt at $\phi_\beta \circ \psi_\beta(c_1, \dots, c_n) = (c_1, \dots, c_n)$ for alle $(c_1, \dots, c_n) \in \mathbb{F}^n$. Vi ser, at dette er opfyldt, da

$$(\psi_\beta \circ \phi_\beta)(\mathbf{v}) = \psi_\beta([\mathbf{v}]_\beta) = \mathbf{v}$$

og

$$(\phi_\beta \circ \psi_\beta)(c_1, \dots, c_n) = \phi_\beta(c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n) = (c_1, \dots, c_n).$$

Der er nu kun tilbage at kontrollere, om ψ_β er en lineær afbildning, hvilket vi overlader til læseren. □

Årsagen til, at de lineære afbildninger ϕ_β og ψ_β er så nyttige, er, at de kan anvendes til at beskrive en generel lineær afbildning mere eksplicit. Lad os antage, at vi er givet en lineær afbildning $L : V_1 \to V_2$ som i Definition 10.0.1, og at vi ved, at både V_1 og V_2 er endeligdimensionelle vektorrum, for eksempel med $\dim V_1 = n$ og $\dim V_2 = m$. Vi kan så vælge en ordnet basis for V_1, lad os sige $\beta = (\mathbf{v}_1, \dots, \mathbf{v}_n)$, hvor $\mathbf{v}_1, \dots, \mathbf{v}_n$ er lineært uafhængige vektorer i V_1. På samme måde kan vi vælge en ordnet basis for V_2, for eksempel $\gamma = (\mathbf{w}_1, \dots, \mathbf{w}_m)$, hvor $\mathbf{w}_1, \dots, \mathbf{w}_m \in V_2$ er lineært uafhængige vektorer i V_2. Pointen er nu, at vi i stedet for at arbejde med den abstrakte lineære afbildning $L : V_1 \to V_2$ kan vælge at arbejde med funktionen $\phi_\gamma \circ L \circ \psi_\beta : \mathbb{F}^n \to \mathbb{F}^m$. Effekten er, at de abstrakte vektorrum V_1 og

V_2 er blevet erstattet af de mere håndgribelige vektorrum $\mathbb{F}^n$ og $\mathbb{F}^m$. Ved at benytte Sætning 10.2.1 i kombination med Lemma 10.3.1 kan vi konkludere, at funktionen $\phi_\gamma \circ L \circ \psi_\beta$ rent faktisk er en lineær afbildning mellem vektorrum over $\mathbb{F}$, da det er sammensætningen af lineære afbildninger.

Vi har i Afsnit 10.1 set, at enhver matrix $\mathbf{A} \in \mathbb{F}^{m\times n}$ giver anledning til en lineær afbildning $L_\mathbf{A} : \mathbb{F}^n \to \mathbb{F}^m$, ved at der defineres $\mathbf{v} \mapsto \mathbf{A} \cdot \mathbf{v}$. Faktisk er enhver lineær afbildning fra $\mathbb{F}^n$ til $\mathbb{F}^m$ af denne form. Lad os vise dette nu:

Lemma 10.3.2

Lad $\mathbb{F}$ være et legeme og $\tilde{L} : \mathbb{F}^n \to \mathbb{F}^m$ en lineær afbildning. Da findes der præcis én matrix $\mathbf{A} \in \mathbb{F}^{m\times n}$, således at $\tilde{L} = L_\mathbf{A}$. Betegnes standardbasisvektorerne for $\mathbb{F}^n$ ved $\mathbf{e}_1, \dots, \mathbf{e}_n$, så udgøres søjlerne i matricen $\mathbf{A}$ af $\tilde{L}(\mathbf{e}_1), \dots, \tilde{L}(\mathbf{e}_n)$.

Bevis. Hvis $\mathbf{v} = (c_1, \dots, c_n) \in \mathbb{F}^n$, så er $\mathbf{v} = c_1 \cdot \mathbf{e}_1 + \cdots + c_n \cdot \mathbf{e}_n$, da den i'te standardbasisvektor for $\mathbb{F}^n$ har et ettal som koordinat i og nuller derudover. Da $\tilde{L}$ er en lineær afbildning, har vi $\tilde{L}(\mathbf{v}) = \tilde{L}(c_1 \cdot \mathbf{e}_1 + \cdots + c_n \cdot \mathbf{e}_n) = c_1 \cdot \tilde{L}(\mathbf{e}_1) + \cdots + c_n \cdot \tilde{L}(\mathbf{e}_n)$. Derfor opfylder matricen $\mathbf{A}$ med søjlerne $\tilde{L}(\mathbf{e}_1), \dots, \tilde{L}(\mathbf{e}_n)$, at $\mathbf{A} \cdot \mathbf{v} = c_1 \cdot \tilde{L}(\mathbf{e}_1) + \cdots + c_n \cdot \tilde{L}(\mathbf{e}_n) = \tilde{L}(\mathbf{v})$. Dette viser, at $\tilde{L} = L_\mathbf{A}$.

Det, der er tilbage at vise, er, at matricen $\mathbf{A}$ er unik. Antag, at der findes en anden matrix $\mathbf{B} \in \mathbb{F}^{m\times n}$, således at $\tilde{L} = L_\mathbf{B}$. Vi vil vise, at $\mathbf{A} = \mathbf{B}$. Hvis $\mathbf{A} \neq \mathbf{B}$, vil det være muligt at finde en søjle, lad os kalde den søjle i, for hvilken matricerne $\mathbf{A}$ og $\mathbf{B}$ er forskellige. Bemærk, at den i'te søjle i $\mathbf{A}$ er lig med $\tilde{L}(\mathbf{e}_i)$ grundet konstruktionen af matricen $\mathbf{A}$. På den anden side er $\tilde{L}(\mathbf{e}_i) = L_\mathbf{B}(\mathbf{e}_i) = \mathbf{B} \cdot \mathbf{e}_i$, hvilket netop er den i'te søjle i $\mathbf{B}$. Tilsyneladende er den i'te søjle i $\mathbf{A}$ og $\mathbf{B}$ begge lig med $L(\mathbf{e}_i)$ og altså alligevel ikke forskellige. Denne modstrid viser, at antagelsen $\mathbf{A} \neq \mathbf{B}$ ikke kan være gyldig, og derfor at $\mathbf{A} = \mathbf{B}$. □

Vi vil for en givet lineær afbildning $L : V_1 \to V_2$ anvende dette lemma på den tilknyttede lineære afbildning $\tilde{L} = \phi_\gamma \circ L \circ \psi_\beta : \mathbb{F}^n \to \mathbb{F}^m$. Lad os, før vi fortsætter med den generelle teori, først gennemgå et eksempel.

Eksempel 10.3.1

Vi genbesøger Eksempel 10.2.5. I det eksempel var V_1 vektorrummet bestående af polynomier i $\mathbb{C}[Z]$ af højst tredje grad og V_2 vektorrummet af polynomier i $\mathbb{C}[Z]$ af højst fjerde grad. Derfor kan vi som ordnet basis for V_1 vælge $\beta = (1, Z, Z^2, Z^3)$, mens en mulig ordnet basis for V_2 er givet ved $\gamma = (1, Z, Z^2, Z^3, Z^4)$. Den lineære afbildning $L : V_1 \to V_2$ beskrevet i Eksempel 10.2.5 afbildede et polynomium $p(Z)$ til $(i + 2Z) \cdot p(Z)$.

Lad os starte med at forklare, hvad den lineære afbildning $\phi_\gamma : V_2 \to \mathbb{F}^5$ er i dette tilfælde. Et element i V_2 er et polynomium af højst fjerde grad. Derfor er $\mathbf{v} \in V_2$ et polynomium på formen $a_0 + a_1 Z + \cdots + a_4 Z^4$ med $a_0, a_1, \dots, a_4 \in \mathbb{C}$, hvilket allerede er skrevet som en linearkombination af vektorerne i den ordnede basis $(1, Z, \dots, Z^4)$. Derfor er

$\phi_\gamma(a_0 + a_1 Z + \cdots + a_4 Z^4) = (a_0, a_1, \ldots, a_4)$, i vektornotation:

$$\phi_\gamma(a_0 + a_1 Z + \cdots + a_4 Z^4) = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_4 \end{bmatrix}.$$

Ligeledes har vi

$$\psi_\beta\left(\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix}\right) = b_0 + b_1 Z + b_2 Z^2 + b_3 Z^3.$$

Vi kan beskrive den lineære afbildning L ved at undersøge, hvad der sker med basisvektorerne i den valgte ordnede basis β, når de passerer igennem afbildningen L. Det gøres nemmest ved at udtrykke resultatet som en linearkombination af basisvektorerne i den valgte ordnede basis γ. Vi får:

$$L(1) = (i + 2Z) \cdot 1 = i + 2Z, \quad L(Z) = (i + 2Z) \cdot Z = iZ + 2Z^2,$$
$$L(Z^2) = (i + 2Z) \cdot Z^2 = iZ^2 + 2Z^3, \quad L(Z^3) = (i + 2Z) \cdot Z^3 = iZ^3 + 2Z^4.$$

Lad os nu bestemme matricen $\mathbf{A}$ beskrevet i Lemma 10.3.2. Vi skal udregne $\tilde{L}(\mathbf{e}_i)$ for $i = 1, \ldots, 4$, hvor $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ er den ordnede standardbasis i $\mathbb{F}^4$, og $\tilde{L} = \phi_\gamma \circ L \circ \psi_\beta : \mathbb{F}^4 \to \mathbb{F}^5$. Da får vi:

$$\begin{aligned} \tilde{L}(\mathbf{e}_1) = (\phi_\gamma \circ L \circ \psi_\beta)(\mathbf{e}_1) &= (\phi_\gamma \circ L \circ \psi_\beta)\left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}\right) \\ &= (\phi_\gamma \circ L)(1) \\ &= \phi_\gamma(i + 2Z) \\ &= \begin{bmatrix} i \\ 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} \end{aligned}$$

og tilsvarende

$$\tilde{L}(\mathbf{e}_2) = \begin{bmatrix} 0 \\ i \\ 2 \\ 0 \\ 0 \end{bmatrix}, \quad \tilde{L}(\mathbf{e}_3) = \begin{bmatrix} 0 \\ 0 \\ i \\ 2 \\ 0 \end{bmatrix}, \quad \text{og} \quad \tilde{L}(\mathbf{e}_4) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ i \\ 2 \end{bmatrix}.$$

Ved anvendelse af Lemma 10.3.2 har vi $\tilde{L} = \phi_\gamma \circ L \circ \psi_\beta = L_{\mathbf{A}}$, hvor

$$\mathbf{A} = \begin{bmatrix} i & 0 & 0 & 0 \\ 2 & i & 0 & 0 \\ 0 & 2 & i & 0 \\ 0 & 0 & 2 & i \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

Definition 10.3.1

Lad $\mathbb{F}$ være et legeme og $L : V_1 \to V_2$ en lineær afbildning mellem to endeligdimensionelle vektorrum med $\dim V_1 = n$ og $\dim V_2 = m$. Lad desuden β være en ordnet basis for V_1 og γ en ordnet basis for V_2. Da betegner vi ved ${}_\gamma[L]_\beta \in \mathbb{F}^{m\times n}$ matricen beskrevet i Lemma 10.3.2, når den anvendes på den lineære afbildning $\tilde{L} = \phi_\gamma \circ L \circ \psi_\beta : \mathbb{F}^n \to \mathbb{F}^m$. Matricen ${}_\gamma[L]_\beta$ kaldes *matrixrepræsentationen* af eller alternativt *afbildningsmatricen* for L med hensyn til de ordnede baser β og γ.

Bemærk, at begrebet afbildningsmatrix på engelsk kaldes *mapping matrix*. For at undgå unødvendige beregninger vil vi beskrive afbildningsmatricen ${}_\gamma[L]_\beta$ mere direkte:

Lemma 10.3.3

Lad $\mathbb{F}$ være et legeme og $L : V_1 \to V_2$ en lineær afbildning mellem to endeligdimensionelle vektorrum med $\dim V_1 = n$ og $\dim V_2 = m$. Lad desuden $\beta = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ være en ordnet basis for V_1 og γ en ordnet basis for V_2. Da har afbildningsmatricen for L med hensyn til de ordnede baser β og γ billedvektorerne $[L(\mathbf{v}_1)]_\gamma, \ldots, [L(\mathbf{v}_n)]_\gamma$ som søjler. Det vil sige:

$${}_\gamma[L]_\beta = [[L(\mathbf{v}_1)]_\gamma \cdots [L(\mathbf{v}_n)]_\gamma].$$

Bevis. Ved at kombinere Definition 10.3.1 og Lemma 10.3.2 ser vi, at ${}_\gamma[L]_\beta$ har søjlerne $\tilde{L}(\mathbf{e}_1), \ldots, \tilde{L}(\mathbf{e}_n)$, hvor $\mathbf{e}_1, \ldots, \mathbf{e}_n$ er standardbasisvektorerne for $\mathbb{F}^n$, og $\tilde{L} = \phi_\gamma \circ L \circ \psi_\beta$. Bemærk nu, at vi for alle i mellem 1 og n har $\psi_\beta(\mathbf{e}_i) = \mathbf{v}_i$ ved brug af definitionen af ψ_β givet i Lemma 10.3.1. Yderligere er $\phi_\gamma(\mathbf{w}) = [\mathbf{w}]_\gamma$ for alle $\mathbf{w} \in V_2$ ved definitionen af afbildningen ϕ_β. Derfor har vi for alle i mellem 1 og n, at

$$\tilde{L}(\mathbf{e}_i) = (\phi_\gamma \circ L \circ \psi_\beta)(\mathbf{e}_i) = (\phi_\gamma \circ L)(\mathbf{v}_i) = \phi_\gamma(L(\mathbf{v}_i)) = [L(\mathbf{v}_i)]_\gamma.$$

$\square$

I Eksempel 10.3.1 bestemte vi matrixrepræsentationen af en lineær afbildning (matricen betegnet ved $\mathbf{A}$ i eksemplet). Lad os gennemgå et par eksempler mere.

Eksempel 10.3.2

Dette eksempel er en fortsættelse af Eksempel 10.2.4. Dér betragtede vi for $\alpha \in \mathbb{R}$ den lineære afbildning $R_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ defineret ved $R_\alpha(v_1, v_2) = (\cos(\alpha) \cdot v_1 - \sin(\alpha) \cdot v_2, \sin(\alpha) \cdot v_1 + \cos(\alpha) \cdot v_2)$. Ved at vælge den ordnede standardbasis $\beta = \gamma = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right)$ for $\mathbb{R}^2$ for både definitionsmængden og dispositionsmængden for den lineære afbildning R_α, opnår vi, at

$$_\gamma[R_\alpha]_\beta = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}. \tag{10.4}$$

Pointen med at repræsentere en lineær afbildning L med matricen $_\gamma[L]_\beta$, er, at strukturen af den oprindelige lineære afbildning er "indkodet" i denne matrix. Den følgende sætning gør dette mere præcist.

Sætning 10.3.4

Lad $\mathbb{F}$ være et legeme og V_1, V_2 og V_3 tre endeligdimensionelle vektorrum over $\mathbb{F}$. Lad desuden β, γ og δ være ordnede baser for henholdsvis V_1, V_2 og V_3. Da gælder der, at

(i) $[L(\mathbf{v})]_\gamma = {}_\gamma[L]_\beta \cdot [\mathbf{v}]_\beta$ for enhver lineær afbildning $L : V_1 \to V_2$ og enhver $\mathbf{v} \in V_1$.

(ii) $_\delta[M \circ L]_\beta = {}_\delta[M]_\gamma \cdot {}_\gamma[L]_\beta$ for alle lineære afbildninger $L : V_1 \to V_2$ og $M : V_2 \to V_3$.

Bevis. Vi beviser først det første punkt. Lad os skrive $\mathbf{A} = {}_\gamma[L]_\beta$ for nemheds skyld. Vi har set, at $\phi_\gamma \circ L \circ \psi_\beta = L_\mathbf{A}$ ved brug af notationen fra Lemma 10.3.1. Derfor er $\phi_\gamma \circ L = L_\mathbf{A} \circ (\psi_\beta)^{-1} = L_\mathbf{A} \circ \phi_\beta$. Men så får vi for enhver $\mathbf{v} \in V_1$, at $(\phi_\gamma \circ L)(\mathbf{v}) = (L_\mathbf{A} \circ \phi_\beta)(\mathbf{v})$. En reduktion af venstre- og højresiden giver

$$(\phi_\gamma \circ L)(\mathbf{v}) = \phi_\gamma(L(\mathbf{v})) = [L(\mathbf{v})]_\gamma$$

og

$$(L_\mathbf{A} \circ \phi_\beta)(\mathbf{v}) = L_\mathbf{A}(\phi_\beta(\mathbf{v})) = L_\mathbf{A}([\mathbf{v}]_\beta) = {}_\gamma[L]_\beta \cdot [\mathbf{v}]_\beta.$$

Derfor er $[L(\mathbf{v})]_\gamma = {}_\gamma[L]_\beta \cdot [\mathbf{v}]_\beta$, hvilket er, hvad vi skulle vise.

Beviset for det andet punkt er lignende. Vi skriver $\mathbf{A} = {}_\gamma[L]_\beta$ og $\mathbf{B} = {}_\delta[M]_\gamma$ for nemheds skyld. Vi har $L_\mathbf{A} = \phi_\gamma \circ L \circ \psi_\beta$ og $L_\mathbf{B} = \phi_\delta \circ M \circ \psi_\gamma$, hvilket medfører, at $L_\mathbf{B} \circ L_\mathbf{A} = \phi_\delta \circ M \circ \psi_\gamma \circ \phi_\gamma \circ L \circ \psi_\beta$. Ved at udnytte, at ψ_γ og ϕ_γ er hinandens inverse jævnfør Lemma 10.3.1, får vi, at $L_\mathbf{B} \circ L_\mathbf{A} = \phi_\delta \circ M \circ L \circ \psi_\beta$. Vi har på den ene, at $L_\mathbf{B} \circ L_\mathbf{A} = L_{\mathbf{B} \cdot \mathbf{A}}$, og på den anden side, at $\phi_\delta \circ M \circ L \circ \psi_\beta = L_\mathbf{C}$, hvor $\mathbf{C} = {}_\delta[M \circ L]_\beta$. Dette medfører, at $_\delta[M \circ L]_\beta = \mathbf{B} \cdot \mathbf{A} = {}_\delta[M]_\gamma \cdot {}_\gamma[L]_\beta$, hvilket er, hvad vi ønskede at vise. □

Det første punkt i denne sætning fortæller os simpelthen, at matricen $_\gamma[L]_\beta$ indeholder al den information, vi har brug for for at kunne beskrive den lineære afbildning L: at udregne $L(\mathbf{v})$ og dernæst at udregne koordinatvektoren for resultatet med hensyn til den ordnede basis γ for V_2 er nøjagtig det samme som at gange matricen ${}_\gamma[L]_\beta$ med koordinatvektoren for $\mathbf{v}$ med hensyn til den ordnede basis β for V_1. Det andet punkt siger, at sammensætning af lineære afbildninger opfører sig pænt med hensyn til matrixrepræsentationer. Lad os se på et eksempel på dette.

Eksempel 10.3.3

Vi fortsætter med Eksempel 10.3.2. Vi har set, at hvis vi vælger β og γ til at være standardbasen for $\mathbb{R}^2$, så er ${}_\gamma[R_\alpha]_\beta$ som angivet i Ligning (10.4). Husk på, at afbildningen $R_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ geometrisk kan forstås som en rotation om origo med vinklen α mod uret. For eksempel svarer $R_{\pi/2}$ til en rotation på $\pi/2$ radianer (90 grader). Dette betyder, at $R_{\pi/2} \circ R_{\pi/2} = R_\pi$, altså en rotation på samlet π radianer (180 grader), hvorfor vi har, at $R_\pi(v_1, v_2) = (-v_1, -v_2)$. Lad os tjekke det andet punkt i Sætning 10.3.4 for $V_1 = V_2 = V_3 = \mathbb{R}^2$, $\beta = \gamma = \delta = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$ og $L = M = R_{\pi/2}$. Så på den ene side har vi

$$ {}_\delta[R_{\pi/2}]_\gamma = {}_\gamma[R_{\pi/2}]_\beta = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} $$

og derfor

$$ {}_\delta[R_{\pi/2}]_\gamma \cdot {}_\gamma[R_{\pi/2}]_\beta = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}. $$

På den anden side ser vi ved anvendelse af Ligning (10.4) med $\alpha = \pi$, at

$$ {}_\delta[R_{\pi/2} \circ R_{\pi/2}]_\beta = {}_\delta[R_\pi]_\beta = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}. $$

Vi konkluderer, at vi ganske rigtigt har ${}_\delta[R_{\pi/2} \circ R_{\pi/2}]_\beta = {}_\delta[R_{\pi/2}]_\gamma \cdot {}_\gamma[R_{\pi/2}]_\beta$, nøjagtigt som det bør være.

Ved at udføre samme udregninger for $M = R_{\alpha_1}$ og $L = R_{\alpha_2}$ og udnytte, at $R_{\alpha_1} \circ R_{\alpha_2} = R_{\alpha_1+\alpha_2}$, opnår man, at

$$ \begin{bmatrix} \cos(\alpha_1) & -\sin(\alpha_1) \\ \sin(\alpha_1) & \cos(\alpha_1) \end{bmatrix} \cdot \begin{bmatrix} \cos(\alpha_2) & -\sin(\alpha_2) \\ \sin(\alpha_2) & \cos(\alpha_2) \end{bmatrix} = \begin{bmatrix} \cos(\alpha_1+\alpha_2) & -\sin(\alpha_1+\alpha_2) \\ \sin(\alpha_1+\alpha_2) & \cos(\alpha_1+\alpha_2) \end{bmatrix}. $$

Denne identitet medfører faktisk additionsformlerne for cosinus og sinus, som vi benyttede i beviset af Lemma 3.4.1:

$$ \cos(\alpha_1 + \alpha_2) = \cos(\alpha_1)\cos(\alpha_2) - \sin(\alpha_1)\sin(\alpha_2) $$

og

$$ \sin(\alpha_1 + \alpha_2) = \sin(\alpha_1)\cos(\alpha_2) + \cos(\alpha_1)\sin(\alpha_2). $$

Eksempel 10.3.4

Lad $\mathbb{F} = \mathbb{R}$ og $V_1 = \mathbb{R}^2$, $V_2 = \mathbb{R}^2$, og lad

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}.$$

Betegn ved $L_\mathbf{A} : \mathbb{R}^2 \to \mathbb{R}^2$ den lineære afbildning defineret ved $\mathbf{v} \mapsto \mathbf{A} \cdot \mathbf{v}$. Vi har for eksempel

$$L_\mathbf{A}\left(\begin{bmatrix} -1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

og

$$L_\mathbf{A}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

Spørgsmål:

(a) Vælg de ordnede standardbaser $\beta = \gamma = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$ for V_1 og V_2, og bestem ${}_\gamma[L_\mathbf{A}]_\beta$.

(b) Vælg de ordnede standardbaser $\beta = \gamma = \left(\begin{bmatrix} -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$ for V_1 og V_2, og bestem ${}_\gamma[L_\mathbf{A}]_\beta$.

Svar:

(a) Da γ er valgt som standardbasen, og

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = v_1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + v_2 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

har vi, at $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}_\gamma = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ for alle $v_1, v_2 \in \mathbb{R}$. Ved brug af Lemma 10.3.3 får vi, at

$${}_\gamma[L_\mathbf{A}]_\beta = [[L_\mathbf{A}(\mathbf{e}_1)]_\gamma \; [L_\mathbf{A}(\mathbf{e}_2)]_\gamma] = [L_\mathbf{A}(\mathbf{e}_1) \; L_\mathbf{A}(\mathbf{e}_2)] = [\mathbf{A} \cdot \mathbf{e}_1 \; \mathbf{A} \cdot \mathbf{e}_2] = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} = \mathbf{A}.$$

Man kan faktisk på en lignende måde se, at man for ethvert legeme $\mathbb{F}$ og enhver matrix $\mathbf{A} \in \mathbb{F}^{m \times n}$ har ${}_\gamma[L_\mathbf{A}]_\beta = \mathbf{A}$, hvis β og γ er de ordnede standardbaser for $\mathbb{F}^m$, henholdsvis $\mathbb{F}^n$.

(b) Nu vælger vi de ordnede baser $\beta = \gamma = \left(\begin{bmatrix} -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$ for V_1 og V_2. Ved brug af Lemma 10.3.3 har vi, at

$${}_\gamma[L_\mathbf{A}]_\beta \;=\; \left[[L_\mathbf{A}\left(\begin{bmatrix} -1 \\ 2 \end{bmatrix}\right)]_\gamma \; [L_\mathbf{A}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)]_\gamma \right] = \left[[\mathbf{A} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix}]_\gamma \; [\mathbf{A} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}]_\gamma \right]$$

$$= \left[\begin{bmatrix} 1 \\ -2 \end{bmatrix}_\gamma \begin{bmatrix} 2 \\ 2 \end{bmatrix}_\gamma \right].$$

For at bestemme $[\mathbf{w}]_\gamma$ for en $\mathbf{w} \in \mathbb{R}^2$ skal man generelt løse et lineært ligningssystem. Mere præcist, hvis vi skriver $\mathbf{w} = (w_1, w_2)$, er målet at finde nogle $c_1, c_2 \in \mathbb{R}^2$, således at

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = c_1 \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix} + c_2 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Derfor skal vi løse systemet af lineære ligninger i de ubekendte c_1 og c_2 givet ved:

$$\begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

Dette kan i princippet gøres ved anvendelse af teorien fra Kapitel 6, alternativt ved at multiplicere systemet på begge sider af lighedstegnet med matricen.

$$\begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}^{-1} = \frac{1}{3} \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}.$$

Vi er dog heldige i dette tilfælde, da vi direkte kan se, at

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} = (-1) \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad \text{og derfor} \quad \begin{bmatrix} 1 \\ -2 \end{bmatrix}_\gamma = \begin{bmatrix} -1 \\ 0 \end{bmatrix},$$

samt

$$\begin{bmatrix} 2 \\ 2 \end{bmatrix} = 2 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \text{hvilket medfører} \quad \begin{bmatrix} 2 \\ 2 \end{bmatrix}_\gamma = \begin{bmatrix} 0 \\ 2 \end{bmatrix}.$$

Vi konkluderer, at

$${}_\gamma[L_\mathbf{A}]_\beta = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Resultatet er en overraskende pæn matrix, nemlig en diagonalmatrix (se Definition 8.1.3).

Som det sidste i dette afsnit betragter vi matricer på formen ${}_\gamma[L]_\beta$ i et scenarie, hvor L er identitetsafbildningen fra et vektorrum V til sig selv: $\mathrm{id}_V : V \to V$, $\mathbf{v} \mapsto \mathbf{v}$. Her er β og γ to, muligvis forskellige, ordnede baser for V. Fra den første del af Sætning 10.3.4 ser vi, at

$${}_\gamma[\mathrm{id}_V]_\beta \cdot [\mathbf{v}]_\beta = [\mathbf{v}]_\gamma \quad \text{for alle } \mathbf{v} \in V. \tag{10.5}$$

I ord siger Ligning (10.5), at hvis man ganger matricen ${}_\gamma[\mathrm{id}_V]_\beta$ med β-koordinatvektoren af en vektor $\mathbf{v}$ i V, er resultatet γ-koordinatvektoren af $\mathbf{v}$. Af denne grund kaldes matricen ${}_\gamma[\mathrm{id}_V]_\beta$ for en *koordinatskiftematrix*, alternativt en *basisskiftematrix*. Bemærk, at disse begreber på engelsk kaldes *change-of-coordinates matrix*, henholdsvis *change-of-basis matrix*.

Eksempel 10.3.5

Lad $V = \{p(Z) \in \mathbb{C}[Z] \mid \deg p(Z) \leq 3\}$. $\beta = (1, Z, Z^2, Z^3)$ og $\gamma = (Z^3, Z^2, Z, 1)$ er to ordnede baser for V.

Spørgsmål: Bestem den tilsvarende koordinatskiftematrix ${}_\gamma[\mathrm{id}_V]_\beta$.

Svar: Lemma 10.3.3 fortæller, at vi skal bestemme $[1]_\gamma$, $[Z]_\gamma$, $[Z^2]_\gamma$ og $[Z^3]_\gamma$ for at løse opgaven. Da den eneste forskel mellem β og γ er rækkefølgen på basisvektorerne, er dette ikke så svært at gøre. For eksempel er

$$[1]_\gamma = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

da 1 er den fjerde basisvektor i γ. Ved at fortsætte på samme måde for de andre basisvektorer opnår man den ønskede koordinatskiftematrix:

$${}_\gamma[\mathrm{id}_V]_\beta = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Vi afslutter dette afsnit med nogle fakta om koordinatskiftematricer, der vil vise sig nyttige senere hen.

Lemma 10.3.5

Lad $\mathbb{F}$ være et legeme, V et vektorrum over $\mathbb{F}$ af endelig dimension n og β, γ samt δ ordnede baser i V. Da gælder

(i) ${}_\delta[\mathrm{id}_V]_\gamma \cdot {}_\gamma[\mathrm{id}_V]_\beta = {}_\delta[\mathrm{id}_V]_\beta$,

(ii) ${}_\beta[\mathrm{id}_V]_\beta = \mathbf{I}_n$, hvor $\mathbf{I}_n$ betegner $n \times n$-identitetsmatricen, og

(iii) $({}_\gamma[\mathrm{id}_V]_\beta)^{-1} = {}_\beta[\mathrm{id}_V]_\gamma$.

Bevis. Det første punkt følger direkte fra punkt to i Sætning 10.3.4. Det andet punkt gælder klart, da koordinaterne af en vektor med hensyn til β ikke ændrer sig, hvis den ordnede basis β ikke ændres. Bemærk for det tredje punkt, at vi ifølge den første og anden del af sætningen har ${}_\gamma[\mathrm{id}_V]_\beta \cdot {}_\beta[\mathrm{id}_V]_\gamma = {}_\gamma[\mathrm{id}_V]_\gamma = \mathbf{I}_n$ og på samme måde ${}_\beta[\mathrm{id}_V]_\gamma \cdot {}_\gamma[\mathrm{id}_V]_\beta = {}_\beta[\mathrm{id}_V]_\gamma = \mathbf{I}_n$. Derfor er $({}_\gamma[\mathrm{id}_V]_\beta)^{-1} = {}_\beta[\mathrm{id}_V]_\gamma$. □

10.4 Anvendelser af matrixrepræsentationen af en lineær afbildning

Nu hvor vi har muligheden for at repræsentere lineære afbildninger mellem endeligdimensionelle vektorrum med en matrix, vil vi benytte dette til at beskrive mere detaljeret, hvordan man bestemmer kernen og billedet af en lineær afbildning. Vi starter med en mere generel beskrivelse af løsninger til ligninger, der involverer en lineær afbildning.

Sætning 10.4.1

Lad $\mathbb{F}$ være et legeme og $L : V_1 \to V_2$ en lineær afbildning mellem vektorrum over $\mathbb{F}$. Lad også en vektor $\mathbf{w} \in V_2$ være givet, og definér $S = \{\mathbf{v} \in V_1 \mid L(\mathbf{v}) = \mathbf{w}\}$. Da vil netop én af de følgende to scenarier forekomme:

(i) $S = \emptyset$. Dette er tilfældet, hvis og kun hvis $\mathbf{w} \notin \operatorname{image} L$.

(ii) $S = \{\mathbf{v}_p + \mathbf{v} \mid \mathbf{v} \in \ker L\}$, hvor $\mathbf{v}_p \in V_1$ er en vektor, således at $L(\mathbf{v}_p) = \mathbf{w}$.

Bevis. Hvis $S = \emptyset$, så har ligningen $L(\mathbf{v}) = \mathbf{w}$ ingen løsninger. Dette er ækvivalent med udsagnet om, at ingen vektor $\mathbf{v} \in V_1$ afbildes til $\mathbf{w}$. Dette er igen det samme som at sige, at $\mathbf{w}$ ikke er i billedet af L.

Hvis $S \neq \emptyset$, kan vi konkludere, at der findes en vektor $\mathbf{v}_p \in V_1$, således at $L(\mathbf{v}_p) = \mathbf{w}$. Hvis $\tilde{\mathbf{v}}$ er en vektor, således at $L(\tilde{\mathbf{v}}) = \mathbf{w}$, så ser vi ved at bruge lineariteten af L, at $L(\tilde{\mathbf{v}} - \mathbf{v}_p) = \mathbf{w} - \mathbf{w} = \mathbf{0}$. Derfor er $\tilde{\mathbf{v}} - \mathbf{v}_p \in \ker L$. Da $\tilde{\mathbf{v}} = \mathbf{v}_p + (\tilde{\mathbf{v}} - \mathbf{v}_p)$, og, som vi allerede har set, $\tilde{\mathbf{v}} - \mathbf{v}_p \in \ker L$, viser dette, at $S \subseteq \{\mathbf{v}_p + \mathbf{v} \mid \mathbf{v} \in \ker L\}$. Omvendt, hvis en vektor er på formen $\mathbf{v}_p + \mathbf{v}$ for et $\mathbf{v} \in \ker L$, så er $L(\mathbf{v}_p + \mathbf{v}) = L(\mathbf{v}_p) + L(\mathbf{v}) = \mathbf{w} + \mathbf{0} = \mathbf{w}$. Dette viser, at $\{\mathbf{v}_p + \mathbf{v} \mid \mathbf{v} \in \ker L\} \subseteq S$. Ved at kombinere begge resultater, kan vi konkludere, at $S = \{\mathbf{v}_p + \mathbf{v} \mid \mathbf{v} \in \ker L\}$. □

Derfor er strukturen af løsningsmængden til en ligning på formen $L(\mathbf{v}) = \mathbf{w}$ fuldstændig bestemt. Vektoren $\mathbf{v}_p$, hvis en sådan eksisterer, kaldes en *partikulær løsning*. Bemærk, i hvor høj grad dette ligner Sætning 6.1.2. Det er ikke en tilfældighed. Løsningsmængden til et system af lineære ligninger med totalmatrix $[\mathbf{A}|\mathbf{b}]$ er trods alt den samme som løsningsmængden til ligningen $L_\mathbf{A}(\mathbf{v}) = \mathbf{b}$. Desuden er $\ker L_\mathbf{A}$ lig med løsningsmængden til det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A}$. Derfor er Sætning 6.1.2 sådan set blot et særtilfælde af Sætning 10.4.1.

Hvis både V_1 og V_2 er endeligdimensionelle vektorrum, kan vi beregningsmæssigt løse en ligning på formen $L(\mathbf{v}) = \mathbf{w}$ ved at løse et passende system af lineære ligninger. Vi formulerer dette mere præcist i den følgende sætning.

Sætning 10.4.2

Lad $\mathbb{F}$ være et legeme og $L : V_1 \to V_2$ en lineær afbildning mellem endeligdimensionelle vektorrum over $\mathbb{F}$. Lad $\beta = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ være en ordnet basis for V_1 og $\gamma = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$ en ordnet basis for V_2. Så gælder der, at

$$\{\mathbf{v} \in V_1 \mid L(\mathbf{v}) = \mathbf{w}\} = \{c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n \mid \mathbf{c} = (c_1, \ldots, c_n) \text{ opfylder } {}_\gamma[L]_\beta \cdot \mathbf{c} = [\mathbf{w}]_\gamma\}.$$

Særligt er

$$\ker L = \{c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n \mid (c_1, \ldots, c_n) \in \ker {}_\gamma[L]_\beta\}.$$

Bevis. Anvendes Lemma 10.3.1 på vektorrummet V_1 med den givne ordnede basis β, ser vi, at de lineære afbildninger $\phi_\beta : V_1 \to \mathbb{F}^n$ defineret ved $\mathbf{v} \mapsto [\mathbf{v}]_\beta$ og $\psi_\beta : \mathbb{F}^n \to V_1$ defineret ved $(c_1, \ldots, c_n) \mapsto c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n$ er hinandens inverse.

Antag, at $L(\mathbf{v}) = \mathbf{w}$. Da er $[L(\mathbf{v})]_\gamma = [\mathbf{w}]_\gamma$, hvilket ved hjælp af det første punkt i Sætning 10.3.4 medfører, at ${}_\gamma[L]_\beta[\mathbf{v}]_\beta = [\mathbf{w}]_\gamma$. Hvis vi skriver $(c_1, \ldots, c_n) = [\mathbf{v}]_\beta$, så er $\mathbf{v} = c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n$. Dette viser, at

$$\{\mathbf{v} \in V_1 \mid L(\mathbf{v}) = \mathbf{w}\} \subseteq \{c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n \mid \mathbf{c} = (c_1, \ldots, c_n) \text{ opfylder } {}_\gamma[L]_\beta \cdot \mathbf{c} = [\mathbf{w}]_\gamma\}.$$

Antag omvent, at $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{F}^n$ opfylder ${}_\gamma[L]_\beta \cdot \mathbf{c} = [\mathbf{w}]_\gamma$. Vektoren $\mathbf{v} = c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n$ har egenskaben, at $[\mathbf{v}]_\beta = \mathbf{c}$. Derfor er ${}_\gamma[L]_\beta \cdot [\mathbf{v}]_\beta = [\mathbf{w}]_\gamma$. Ved at benytte Sætning 10.3.4 igen, ser vi, at $[L(\mathbf{v})]_\gamma = [\mathbf{w}]_\gamma$. Men så er $L(\mathbf{v}) = \mathbf{w}$. Dette viser, at

$$\{\mathbf{v} \in V_1 \mid L(\mathbf{v}) = \mathbf{w}\} \supseteq \{c_1 \cdot \mathbf{v}_1 + \cdots c_n \cdot \mathbf{v}_n \mid \mathbf{c} = (c_1, \ldots, c_n) \text{ opfylder } {}_\gamma[L]_\beta \cdot \mathbf{c} = [\mathbf{w}]_\gamma\}.$$

Ved at kombinere ovenstående to resultater følger den første del af sætningen.

Ved at vælge $\mathbf{w} = \mathbf{0}$ følger udsagnet om $\ker L$. $\square$

Pointen med denne sætning er, at det for at bestemme alle løsninger til ligningen $L(\mathbf{v}) = \mathbf{w}$ er nok at bestemme alle løsninger til ligningen ${}_\gamma[L]_\beta \cdot \mathbf{c} = [\mathbf{w}]_\gamma$. Sidstnævnte ligning er et system af lineære ligninger med totalmatrix $[{}_\gamma[L]_\beta | [\mathbf{w}]_\gamma]$, som vi kan løse ved hjælp af teknikkerne fra Kapitel 6. Det faktum, at kernen af en lineær afbildning kan bestemmes ved hjælp af matrixrepræsentationen af den pågældende afbildning, har følgende pæne konsekvens, der kendes som *dimensionssætningen for lineære afbildninger*. Bemærk, at dette begreb på engelsk kaldes *the rank-nullity theorem for linear maps*.

Korollar 10.4.3

Lad $\mathbb{F}$ være et legeme og $L : V_1 \to V_2$ en lineær afbildning mellem endeligdimensionelle vektorrum over $\mathbb{F}$. Da gælder der, at

$$\dim(\ker L) + \dim(\operatorname{image} L) = \dim V_1.$$

Bevis. Hvis $\{\mathbf{v}_1, \ldots, \mathbf{v}_d\}$ er en basis for $\ker L$, så er $\{[\mathbf{v}_1]_\beta, \ldots, [\mathbf{v}_d]_\beta\}$ en basis for $\ker {}_\gamma[L]_\beta$ jævnfør Sætning 9.2.3. Defor er $\dim \ker L = \dim \ker {}_\gamma[L]_\beta$. Derudover er $\dim \operatorname{image} L = \dim \operatorname{image}_\gamma[L]_\beta$ jævnfør Korollar 9.3.5. Dermed følger resultatet fra dimensionssætningen for matricer (se Sætning 10.1.3). □

Eksempel 10.4.1

Dette eksempel er en variation af Eksempel 10.2.9. I det eksempel betragtede vi afbildningen $\mathrm{ev} : \mathbb{C}[Z] \to \mathbb{C}^2$ defineret ved $p(Z) \mapsto (p(0), p(1))$ og bestemte dens kerne. Lad nu $V_1 \subseteq \mathbb{C}[Z]$ være underrummet af $\mathbb{C}[Z]$ bestående af alle polynomier af grad højst tre. Så er $\beta = (1, Z, Z^2, Z^3)$ en ordnet basis for V_1. For $\mathbb{C}^2$ vælger vi den ordnede standardbasis $(\mathbf{e}_1, \mathbf{e}_2)$. Lad os nu betragte den lineære afbildning $L : V_1 \to \mathbb{C}^2$ defineret ved $L(p(Z)) = (p(0), p(1))$. Med andre ord: vi afgrænser definitionsmængden for ev til V_1 men ændrer ellers ikke noget.

Spørgsmål: Hvad er kernen af den lineære afbildning L beskrevet ovenfor? Hvad er alle løsninger til ligningen $L(p(Z)) = (5, 8)$?

Svar:

Vi kan bestemme $\ker L$ på flere måder, men lad os her benytte Sætning 10.4.2. For at bestemme $\ker L$ bestemmer vi først kernen af ${}_\gamma[L]_\beta$. Vi har $L(1) = (1,1) = 1 \cdot \mathbf{e}_1 + 1 \cdot \mathbf{e}_2$, $L(Z) = (0,1) = 1 \cdot \mathbf{e}_2$, $L(Z^2) = (0,1) = 1 \cdot \mathbf{e}_2$ og $L(Z^3) = (0,1) = 1 \cdot \mathbf{e}_2$. Derfor er

$${}_\gamma[L]_\beta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Den reducerede trappeform af denne matrix kan i dette tilfælde hurtigt bestemmes ved, at den første række trækkes fra den anden række. Vi opnår matricen:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}.$$

Fra dette resultat får vi ved hjælp af Sætning 6.4.4 og Korollar 9.3.4, at en basis for $\ker_\gamma[L]_\beta$ er givet ved mængden

$$\left\{ \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Dermed er en basis for $\ker L$ givet ved mængden $\{-Z + Z^2, -Z + Z^3\}$, og vi har så, at

$$\ker L = \{t_1 \cdot (-Z + Z^2) + t_2 \cdot (-Z + Z^3) \mid t_1, t_2 \in \mathbb{C}\}.$$

Vi besvarer det afsluttende spørgsmål om løsningerne til ligningen $L(p(Z)) = (5, 8)$ ved at benytte Sætning 10.4.1. Det eneste, vi mangler, er en partikulær løsning. Vi kan igen omdanne ligningen til et system af lineære ligninger, hvilket giver anledning til et system af inhomogene lineære ligninger med totalmatrix

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 5 \\ 1 & 1 & 1 & 1 & 8 \end{array}\right],$$

som har reduceret trappeform

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 5 \\ 0 & 1 & 1 & 1 & 3 \end{array}\right].$$

En partikulær løsning (c_1, c_2, c_3, c_4) skal opfylde $c_1 = 5$ og $c_2 + c_3 + c_4 = 3$. Derfor er $(5, 3, 0, 0)$ en partikulær løsning, som svarer til polynomiet $f(Z) = 5 + 3Z$. Ved hjælp af Sætning 10.4.1 konkluderer vi, at alle løsninger til ligningen $L(p(Z)) = (5, 8)$ udgør mængden

$$\{5 + 3Z + t_1 \cdot (-Z + Z^2) + t_2 \cdot (-Z + Z^3) \mid t_1, t_2 \in \mathbb{C}\}.$$

En lille sidebemærkning: en anden måde at bestemme $\ker L$ på er at udnytte, at vi allerede har bestemt kernen af $\mathrm{ev} : \mathbb{C}[Z] \to \mathbb{C}^2$ i Eksempel 10.2.9. Da har vi nemlig, at

$$\begin{aligned} \ker L &= \ker \mathrm{ev} \cap V_1 \\ &= \{Z \cdot (Z-1) \cdot s(Z) \mid s(Z) \in \mathbb{C}[Z]\} \cap V_1 \\ &= \{Z \cdot (Z-1) \cdot s(Z) \mid s(Z) \in \mathbb{C}[Z], \deg s(Z) \leq 1\}. \end{aligned}$$

Her benyttede vi, at $Z \cdot (Z-1) \cdot s(Z) \in V_1$ gælder, hvis $\deg(Z \cdot (Z-1) \cdot s(Z)) \leq 3$. Da $\deg(Z \cdot (Z-1) \cdot s(Z)) = 1 + 1 + \deg s(Z)$, ser vi, at $Z \cdot (Z-1) \cdot s(Z) \in V_1$, når $\deg s(Z) \leq 1$. Vi lader det være op til læseren at kontrollere, at denne måde at bestemme $\ker L$ på giver samme resultat som før.

Kapitel 11

Egenværdiproblemet og diagonalisering

Lad os tage et kig på Eksempel 10.3.4 igen. Her arbejdede vi med matricen

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

og den tilknyttede lineære afbildning $L_\mathbf{A} : \mathbb{R}^2 \to \mathbb{R}^2$. Vi så, at hvis vi valgte den samme ordnede basis $\beta = \left(\begin{bmatrix} -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$ for både definitionsmængden og dispositionsmængden for $L_\mathbf{A}$, så var den resulterende afbildningsmatrix ${}_\beta[L_\mathbf{A}]_\beta$ for $L_\mathbf{A}$ særligt pæn:

$${}_\beta[L_\mathbf{A}]_\beta = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}.$$

I dette kapitel undersøger vi, i hvilket omfang dette kan gøres for en vilkårlig kvadratisk matrix.

11.1 Egenvektorer og egenværdier

Vi lægger ud med at studere lineære afbildninger af typen $L : V \to V$. Forskellen fra vores tidligere studier af lineære afbildninger er, at vi nu antager, at definitionsmængden for L er den samme som dispositionsmængden for L, nemlig vektorrummet V.

Definition 11.1.1

Lad $\mathbb{F}$ være et legeme, V et vektorrum over $\mathbb{F}$ og $L : V \to V$ en lineær afbildning. Lad $\mathbf{v} \in V$

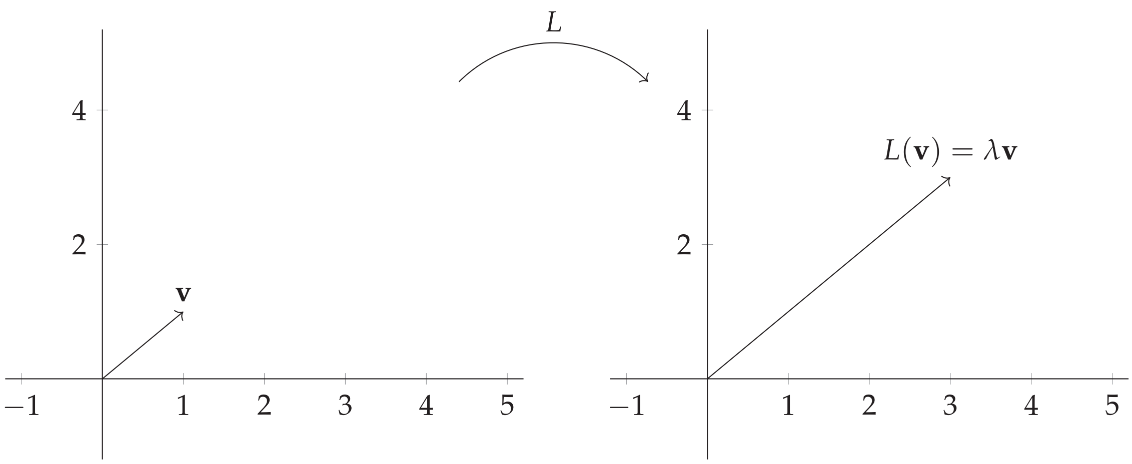
være en vektor forskellig fra nulvektoren og $\lambda \in \mathbb{F}$ en skalar, således at

$$L(\mathbf{v}) = \lambda \cdot \mathbf{v}.$$

Da kaldes vektoren $\mathbf{v}$ en *egenvektor* for den lineære afbildning L med *egenværdi* λ.

Bemærk, at begreberne egenvektor og egenværdi på engelsk kaldes de tyskklingende *eigenvector*, henholdsvis *eigenvalue*.

Definitionen angiver, at en egenvektor per definition altid er en vektor forskellig fra nulvektoren. Grunden til dette er, at man typisk ønsker at udelade uinteressante løsninger til ligningen $L(\mathbf{v}) = \lambda \cdot \mathbf{v}$. Vælger man $\mathbf{v} = \mathbf{0}$ og en hvilken som helst $\lambda \in \mathbb{F}$, vil der nemlig altid gælde, at $L(\mathbf{v}) = \lambda \cdot \mathbf{v}$, da $L(\mathbf{0}) = \mathbf{0}$, og $\lambda \cdot \mathbf{0} = \mathbf{0}$. Bemærk også, at en egenværdi altid er et element fra legemet $\mathbb{F}$, over hvilket V er et vektorrum. Intuitivt set er en egenvektor for en lineær operator L en vektor, der skaleres, når L opererer på den. Vi kan opfatte $\lambda \cdot \mathbf{v}$ som en skalering af vektoren $\mathbf{v}$ med faktor λ. Se Figur 11.1 for en illustration.



Figur 11.1: En egenvektor for en lineær afbildning L.

For matricer kan man også tale om egenvektorer og egenværdier:

Definition 11.1.2

Lad $\mathbb{F}$ være et legeme, n et positivt heltal og $\mathbf{A} \in \mathbb{F}^{n\times n}$ en matrix. Lad $\mathbf{v} \in \mathbb{F}^n$ være en vektor forskellig fra nulvektoren og $\lambda \in \mathbb{F}$ en skalar, således at

$$\mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}.$$

Da kaldes vektoren $\mathbf{v}$ en *egenvektor* for matricen $\mathbf{A}$ med *egenværdi* λ.

Bemærk, at denne definition antager, at matricen $\mathbf{A}$ er en kvadratisk matrix. Som vi har set tidligere, giver en matrix $\mathbf{A} \in \mathbb{F}^{m\times n}$ anledning til en lineær afbildning $L_{\mathbf{A}} : \mathbb{F}^n \to \mathbb{F}^m$. Hvis $m = n$, ser vi derfor, at en kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ giver anledning til en lineær afbildning $L_{\mathbf{A}} : \mathbb{F}^n \to \mathbb{F}^n$. Bemærk, at $\mathbf{v} \in \mathbb{F}^n$ er en egenvektor for en kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$, hvis og kun hvis $\mathbf{v} \in \mathbb{F}^n$ er en egenvektor for den lineære afbildning $L_{\mathbf{A}} : \mathbb{F}^n \to \mathbb{F}^n$. I den forstand er Definition 11.1.2 blot et særtilfælde af Definition 11.1.1. Matricer opfylder desuden, at hvis legemet er valgt til at være $\mathbb{F}$, så er matricens egenværdier per definition elementer i legemet $\mathbb{F}$.

Lad os gennemgå nogle eksempler.

Eksempel 11.1.1

Lad $\mathbb{F} = \mathbb{R}$, og betragt matricen

$$\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

Spørgsmål: Bestem alle egenværdier for matricen $\mathbf{A}$ samt en tilhørende egenvektor for hver af dem.

Svar: Antag, at $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2 \setminus \{(0,0)\}$ er en egenvektor med egenværdi λ. Ligningen $\mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}$ svarer til de to ligninger $-v_1 = \lambda v_1$ og $2v_2 = \lambda v_2$. Disse to ligninger kan omskrives til $(-1-\lambda)v_1 = 0$ og $(2-\lambda)v_2 = 0$, som igen kan skrives på matrixform som følger:

$$\begin{bmatrix} -1-\lambda & 0 \\ 0 & 2-\lambda \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{11.1}$$

Nu skelner vi mellem tre tilfælde.

I det første tilfælde antager vi, at $-1-\lambda \neq 0$, og $2-\lambda \neq 0$. Med andre ord: vi antager, at $\lambda \neq -1$, og $\lambda \neq 2$. I dette tilfælde er de diagonalelementer i matricen, der optræder i Ligning (11.1), begge forskellige fra nul. Derfor er den eneste løsning til Ligning (11.1), at $(v_1, v_2) = (0,0)$. Men egenvektorer må per definition ikke være lig med nulvektoren, så vi konkluderer, at der i dette tilfælde ikke findes nogen egenvektorer og derfor ej heller nogen egenværdier.

I tilfælde to antager vi, at $\lambda = -1$. I dette tilfælde har Ligning (11.1) løsninger på formen $(v_1, 0)$, hvor $v_1 \in \mathbb{R}$ kan vælges frit. Derfor er $\lambda = -1$ en egenværdi for den givne matrix $\mathbf{A}$. Som egenvektor kan vi vælge enhver vektor på formen $\begin{bmatrix} v_1 \\ 0 \end{bmatrix}$, så længe $v_1 \neq 0$. For eksempel er $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ en egenvektor for den givne matrix $\mathbf{A}$ med egenværdi -1.

Slutteligt antager vi som det tredje og sidste tilfælde, at $\lambda = 2$. I dette tilfælde har Ligning (11.1) løsninger på formen $(0, v_2)$, hvor $v_2 \in \mathbb{R}$ kan vælges frit. Derfor er $\lambda = 2$ en egenværdi for den givne matrix $\mathbf{A}$. Som egenvektor kan vi vælge enhver vektor på formen $\begin{bmatrix} 0 \\ v_2 \end{bmatrix}$, så længe $v_2 \neq 0$. For eksempel er $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ en egenvektor for den givne matrix $\mathbf{A}$ med egenværdi 2.

Også hvis V er et uendeligdimensionelt vektorrum, giver definitionen af egenvektorer og egenværdier mening. Vi ser på et eksempel af denne type.

Eksempel 11.1.2

Lad os betragte den lineære afbildning $D : \mathbb{C}[Z] \to \mathbb{C}[Z]$ defineret i Eksempel 10.2.1. Vi arbejder specifikt over legemet $\mathbb{C}$, da vi i Eksempel 10.2.1 opfattede $\mathbb{C}[Z]$ som et komplekst vektorrum. Afbildningen D blev defineret på den måde, at den sender et polynomium til sin afledte.

Spørgsmål: Hvad er egenværdierne for D? Find også en tilsvarende egenvektor for hver egenværdi.

Svar: Vi leder efter polynomier $p(Z)$ forskellige fra nulpolynomiet i $\mathbb{C}[Z]$ og skalarer $\lambda \in \mathbb{C}$, således at $D(p(Z)) = \lambda \cdot p(Z)$. Lad $p(Z) = a_0 + a_1 Z + a_2 Z^2 + \cdots + a_n Z^n$ være et polynomium forskelligt fra nulpolynomiet. Da $D(a_0 + a_1 Z + a_2 Z^2 + \cdots + a_n Z^n) = a_1 + 2a_2 Z + \cdots + n a_n Z^{n-1}$, vil graden af polynomiet $D(p(Z))$ typisk være én mindre end graden af polynomiet $p(Z)$ selv. Den eneste undtagelse er, hvis $p(Z) = a_0$, i hvilket tilfælde $D(p(Z)) = 0$. Derfor kan $D(p(Z)) = \lambda \cdot p(Z)$ kun opfyldes for konstante polynomier. Hvis $p(Z)$ er et konstant polynomium, så er $p(Z) = a_0$, og $D(a_0) = 0 = 0 \cdot a_0$. Dette viser, at 0 er den eneste egenværdi, som den lineære afbildning D har. Ethvert polynomium $p(Z) = a_0$ med $a_0 \in \mathbb{C} \setminus \{0\}$ er en egenvektor for D med egenværdi 0. Grunden til at nulpolynomiet ikke er en egenvektor for D, er, at egenvektorer per definition skal være forskellige fra nulvektoren. Som dette eksempel viser, kan en egenværdi dog godt være nul.

De tidligere to eksempler kunne antyde, at en lineær afbildning altid har mindst én egenvektor, men det er ikke tilfældet. Lad os se på et eksempel, der viser det.

Eksempel 11.1.3

Rotationsafbildningen $R_{\pi/2} : \mathbb{R}^2 \to \mathbb{R}^2$ fra Eksempel 10.2.4 blev defineret ved $R_{\pi/2}(v_1, v_2) = (-v_2, v_1)$.

Spørgsmål: Har den lineære afbildning $R_{\pi/2} : \mathbb{R}^2 \to \mathbb{R}^2$ nogen egenvektorer?

Svar: Lad os først give et intuitivt svar og derefter et, der anvender definitionerne mere direkte. Hvad den lineære afbildning $R_{\pi/2}$ gør geometrisk, er at tage en vektor som input og returnere vektoren roteret med $\pi/2$ radianer mod uret som output. Hvis en vektor forskellig fra nulvektoren er en egenvektor, vil en rotation med $\pi/2$ radianer skulle resultere i en skalering af inputvektoren. Dette er intuitivt ikke muligt, så vi vil forvente, at den lineære afbildning $R_{\pi/2}$ slet ikke har nogen egenvektorer.

Lad os bevise dette ved hjælp af definitionerne. Hvis afbildningen $R_{\pi/2} : \mathbb{R}^2 \to \mathbb{R}^2$ har egenvektorer, findes der $(v_1, v_2) \in \mathbb{R}^2 \setminus \{(0,0)\}$ og $\lambda \in \mathbb{R}$, således at $R_{\pi/2}(v_1, v_2) = \lambda \cdot (v_1, v_2)$. Tilsvarende er $(-v_2, v_1) = (\lambda \cdot v_1, \lambda \cdot v_2)$, hvilket igen kan omskrives til de to ligninger $-\lambda v_1 - v_2 = 0$ og $v_1 - \lambda v_2 = 0$. Formuleret på matrixform får vi matrixligningen

$$\begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Ved at lægge række to λ gange til række et får vi ligningen $-(1+\lambda^2)v_2 = 0$. Dette medfører, at $v_2 = 0$, da $\lambda^2 + 1$ ikke er nul for noget $\lambda \in \mathbb{R}$. Benyttes den anden række ser vi, at også $v_1 = 0$. Vi konkluderer, at $(v_1, v_2) = (0, 0)$ er eneste løsning. Men en egenvektor må ikke være nulvektoren, så den lineære afbildning $R_{\pi/2} : \mathbb{R}^2 \to \mathbb{R}^2$ har derfor ingen egenvektorer.

Fremgangsmåden til at bestemme de mulige egenvektorer og egenværdier i de tidligere eksempler føltes en smule ad hoc. Heldigvis findes der en procedure, der altid virker, hvis V har en endelig dimension. Vi vil forklare denne procedure nu og lægger ud med egenværdier for en kvadratisk matrix.

Sætning 11.1.1

Lad $\mathbb{F}$ være et legeme og $\mathbf{A} \in \mathbb{F}^{n\times n}$ en kvadratisk matrix. Da er $\lambda \in \mathbb{F}$ en egenværdi for $\mathbf{A}$, hvis og kun hvis $\det(\mathbf{A} - \lambda \cdot \mathbf{I}_n) = 0$, hvor $\mathbf{I}_n$ betegner $n \times n$-identitetsmatricen.

Bevis. Hvis $\lambda \in \mathbb{F}$ er en egenværdi for matricen $\mathbf{A}$, så findes der en vektor forskellig fra nulvektoren $\mathbf{v} \in \mathbb{F}^n$, således at $\mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}$. Da $\lambda \cdot \mathbf{v} = \lambda \cdot (\mathbf{I}_n \cdot \mathbf{v}) = (\lambda \cdot \mathbf{I}_n) \cdot \mathbf{v}$, ser vi, at ligningen $\mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}$ kan omskrives til $\mathbf{A} \cdot \mathbf{v} = (\lambda \cdot \mathbf{I}_n)\mathbf{v}$, som igen kan omskrives til $(\mathbf{A} - \lambda \cdot \mathbf{I}_n) \cdot \mathbf{v} = \mathbf{0}$. Dette viser, at det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A} - \lambda \cdot \mathbf{I}_n$ har en løsning, der ikke er nulløsningen. Ved at bruge Korollar 8.3.6 på den kvadratiske matrix $\mathbf{A} - \lambda \cdot \mathbf{I}_n$, konkluderer vi, at $\det(\mathbf{A} - \lambda \cdot \mathbf{I}_n) = 0$.

Omvendt, hvis $\det(\mathbf{A} - \lambda \cdot \mathbf{I}_n) = 0$, medfører Korollar 8.3.6, at det homogene system af lineære ligninger med koefficientmatrix $\mathbf{A} - \lambda \cdot \mathbf{I}_n$ har en løsning, der ikke er nulløsningen. Enhver sådan løsning $\mathbf{v} \in \mathbb{F}^n$, der ikke er nulløsningen, opfylder så $(\mathbf{A} - \lambda \cdot \mathbf{I}_n) \cdot \mathbf{v} = \mathbf{0}$. Dette kan omskrives til $\mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}$. Derfor er $\mathbf{v}$ en egenvektor for $\mathbf{A}$ med egenværdi λ. □

For en given kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ er udtrykket $\det(\mathbf{A} - Z \cdot \mathbf{I}_n)$ et polynomium i $\mathbb{F}[Z]$ af grad n. Dette polynomium kaldes det *karakteristiske polynomium* for $\mathbf{A}$. Vi vil betegne det ved $p_{\mathbf{A}}(Z)$. Rødderne i dette polynomium i legemet $\mathbb{F}$ er netop alle egenværdierne for matricen $\mathbf{A}$.

Eksempel 11.1.4

Sætning 11.1.1 gør det muligt at beskrive alle egenværdier for en kvadratisk matrix. For eksempel har vi for matricen

$$\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} \in \mathbb{R}^{2\times 2},$$

som vi betragtede i Eksempel 11.1.1, at

$$p_{\mathbf{A}}(Z) = \det(\mathbf{A} - Z \cdot \mathbf{I}_2) = \det\left(\begin{bmatrix} -1-Z & 0 \\ 0 & 2-Z \end{bmatrix}\right) = (-1-Z) \cdot (2-Z) = (Z+1) \cdot (Z-2).$$

Derfor er rødderne i det karakteristiske polynomium $p_{\mathbf{A}}(Z)$ netop -1 og 2. Dette betyder, at egenværdierne for matricen $\mathbf{A}$ er -1 og 2. Ser vi tilbage på Eksempel 11.1.1, kan vi gøre svaret, der blev givet dér, lidt kortere, da det første tilfælde, vi arbejdede på, ikke længere er nødvendigt. Dér behandlede vi nemlig i det første tilfælde alle λ, hvor $\lambda \neq -1$, og $\lambda \neq 2$, men vi ved nu, at der ikke findes nogen egenvektorer med sådanne egenværdier λ.

Eksempel 11.1.5

Som et andet eksempel vil vi betragte matricen

$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

Vi så i Eksempel 10.3.3, at denne matrix repræsenterer den lineære afbildning $R_{\pi/2} : \mathbb{R}^2 \to \mathbb{R}^2$, når den ordnede standardbasis vælges for $\mathbb{R}^2$. I det tilfælde er

$$p_{\mathbf{A}}(Z) = \det(\mathbf{A} - Z \cdot \mathbf{I}_2) = \det\left(\begin{bmatrix} -Z & -1 \\ 1 & -Z \end{bmatrix}\right) = Z^2 + 1.$$

Da vi arbejder over de reelle tal $\mathbb{R}$, og polynomiet $Z^2 + 1$ ingen rødder har i $\mathbb{R}$, konkluderer vi, at matricen $\mathbf{A}$, når den betragtes over $\mathbb{R}$, ikke har nogen egenværdier og derfor heller ikke har nogen egenvektorer.

Nu hvor vi ved, hvordan man finder egenværdier for en kvadratisk matrix, er det naturligt at spørge, hvordan man finder egenvektorer. Vi vil vende tilbage til det spørgsmål i det næste afsnit. I resten af dette afsnit vil vi forklare, hvordan man finder egenværdier for en vilkårlig lineær afbildning $L : V \to V$, hvis V er et endeligdimensionelt vektorrum.

Sætning 11.1.2

Lad $\mathbb{F}$ være et legeme, V et vektorrum over $\mathbb{F}$ af dimension n og $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ en ordnet basis for V. Da er $\lambda \in \mathbb{F}$ en egenværdi for en lineær afbildning $L : V \to V$, hvis og kun hvis $\det({}_\beta[L]_\beta - \lambda \cdot \mathbf{I}_n) = 0$.

Bevis. Hvis $\lambda \in \mathbb{F}$ er en egenværdi for den lineære afbildning $L : V \to V$, så findes der en vektor $\mathbf{v} \in V$ forskellig fra nulvektoren, således at $L(\mathbf{v}) = \lambda \cdot \mathbf{v}$. Derfor har vi jævnfør det første punkt i Sætning 10.3.4, at ${}_\beta[L]_\beta \cdot [\mathbf{v}]_\beta = [L(\mathbf{v})]_\beta = [\lambda \cdot \mathbf{v}]_\beta$. Ved at anvende Lemma 9.2.2 har vi $[\lambda \cdot \mathbf{v}]_\beta = \lambda \cdot [\mathbf{v}]_\beta$. Ved at kombinere disse to ligninger, får vi, at ${}_\beta[L]_\beta \cdot [(\mathbf{v})]_\beta = \lambda \cdot [\mathbf{v}]_\beta$. Derfor er $[\mathbf{v}]_\beta$ en egenvektor for matricen ${}_\beta[L]_\beta$ med egenværdi λ.

Antag omvendt, at $\det({}_\beta[L]_\beta - \lambda \cdot \mathbf{I}_n) = 0$ for en $\lambda \in \mathbb{F}$. Så er λ ifølge Sætning 11.1.1 en egenværdi for matricen ${}_\beta[L]_\beta$. Derfor findes der en vektor forskellig fra nulvektoren $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{F}^n$, som er en egenvektor for matricen ${}_\beta[L]_\beta$ med egenværdi λ. Definerer vi nu $\mathbf{v} = c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n \in V$, så er $\mathbf{c} = [\mathbf{v}]_\beta$. Derfor har vi ${}_\beta[L]_\beta \cdot [(\mathbf{v})]_\beta = \lambda \cdot [\mathbf{v}]_\beta$, hvilket medfører $[L(\mathbf{v})]_\beta = \lambda \cdot [\mathbf{v}]_\beta = [\lambda \cdot \mathbf{v}]_\beta$. Dette medfører, at $L(\mathbf{v}) = \lambda \cdot \mathbf{v}$. Derfor er λ en egenværdi for den lineære afbildning $L : V \to V$. $\square$

Denne sætning viser, at hvis V er et endeligdimensionelt vektorrum, så kan vi reducere udregningen af egenværdier for en lineær afbildning $L : V \to V$ til udregningen af egenværdierne for den kvadratiske matrix ${}_\beta[L]_\beta$, der repræsenterer den lineære afbildning. Her betyder det ikke noget, hvilken ordnet basis for V man vælger. Til fremtidig brug vil vi dog alligevel undersøge effekten på matricen, der repræsenterer L, af at vælge en anden ordnet basis. Her vil koordinatskiftematricerne introduceret i Ligning (10.5) spille en vigtig rolle.

Lemma 11.1.3

Lad $\mathbb{F}$ være et legeme, V et vektorrum over $\mathbb{F}$ af dimension n og $L : V \to V$ en lineær afbildning. Lad derudover β og γ være to ordnede baser for V, og betegn ved $\mathrm{id}_V : V \to V$ identitetsafbildningen $\mathbf{v} \mapsto \mathbf{v}$. Da gælder der, at

$$ {}_\gamma[L]_\gamma = ({}_\beta[\mathrm{id}_V]_\gamma)^{-1} \cdot {}_\beta[L]_\beta \cdot {}_\beta[\mathrm{id}_V]_\gamma. $$

Bevis. Vi ved fra Lemma 10.3.5, at $({}_\beta[\mathrm{id}_V]_\gamma)^{-1} = {}_\gamma[\mathrm{id}_V]_\beta$. Derfor er

$$ \begin{aligned} ({}_\beta[\mathrm{id}_V]_\gamma)^{-1} \cdot {}_\beta[L]_\beta \cdot {}_\beta[\mathrm{id}_V]_\gamma &= {}_\gamma[\mathrm{id}_V]_\beta \cdot {}_\beta[L]_\beta \cdot {}_\beta[\mathrm{id}_V]_\gamma \\ &= {}_\gamma[\mathrm{id}_V]_\beta \cdot {}_\beta[L \circ \mathrm{id}_V]_\gamma \\ &= {}_\gamma[\mathrm{id}_V \circ L \circ \mathrm{id}_V]_\gamma \\ &= {}_\gamma[L]_\gamma. \end{aligned} $$

I den anden og tredje lighed benyttede vi første punkt fra Sætning 10.3.4. □

To kvadratiske matricer $\mathbf{A} \in \mathbb{F}^{n\times n}$ og $\mathbf{B} \in \mathbb{F}^{n\times n}$ kaldes *similære*, hvis der findes en invertibel matrix $\mathbf{Q} \in \mathbb{F}^{n\times n}$, således at $\mathbf{A} = \mathbf{Q}^{-1} \cdot \mathbf{B} \cdot \mathbf{Q}$. Derfor kan Lemma 11.1.3 omskrives med ord som følger: effekten af at vælge en anden ordnet basis for V er, at matricen, der repræsenterer L, erstattes af en similær matrix. Det viser sig, at dette lemma også forklarer, hvorfor det ikke betyder noget, hvilken ordnet basis man vælger, når man udregner egenværdierne for en lineær afbildning. Faktisk har vi følgende:

Sætning 11.1.4

Lad $\mathbb{F}$ være et legeme, V et vektorrum over $\mathbb{F}$ af dimension n og $L : V \to V$ en lineær afbildning. Lad yderligere β og γ være to ordnede baser for V. Da er de karakteristiske polynomier for ${}_\beta[L]_\beta$ og ${}_\gamma[L]_\gamma$ identiske.

Bevis. Lad os for nemheds skyld skrive $\mathbf{Q} = {}_\beta[\mathrm{id}_V]_\gamma$. Ved at anvende Lemma 11.1.3 ser vi, at:

$$ \begin{aligned} p_{{}_\gamma[L]_\gamma}(Z) &= \det({}_\gamma[L]_\gamma - Z \cdot \mathbf{I}_n) \\ &= \det(\mathbf{Q}^{-1} \cdot {}_\beta[L]_\beta \cdot \mathbf{Q} - Z \cdot \mathbf{I}_n) \\ &= \det(\mathbf{Q}^{-1} \cdot {}_\beta[L]_\beta \cdot \mathbf{Q} - Z \cdot \mathbf{Q}^{-1} \cdot \mathbf{Q}) \\ &= \det(\mathbf{Q}^{-1} \cdot {}_\beta[L]_\beta \cdot \mathbf{Q} - Z \cdot \mathbf{Q}^{-1} \cdot \mathbf{I}_n \cdot \mathbf{Q}) \\ &= \det(\mathbf{Q}^{-1} \cdot ({}_\beta[L]_\beta - Z \cdot \mathbf{I}_n) \cdot \mathbf{Q}). \end{aligned} $$

Herefter kommer Sætning 8.3.3 til hjælp. Ved brug af denne sætning kan vi nemlig fortsætte som følger:

$$\begin{aligned}
p_{{}_\gamma[L]_\gamma}(Z) &= \det(\mathbf{Q}^{-1} \cdot ({}_\beta[L]_\beta - Z \cdot \mathbf{I}_n) \cdot \mathbf{Q}) \\
&= \det(\mathbf{Q}^{-1}) \cdot \det({}_\beta[L]_\beta - Z \cdot \mathbf{I}_n) \cdot \det(\mathbf{Q}) \\
&= \det(\mathbf{Q}^{-1}) \cdot \det(\mathbf{Q}) \cdot \det({}_\beta[L]_\beta - Z \cdot \mathbf{I}_n) \\
&= \det(\mathbf{Q})^{-1} \cdot \det(\mathbf{Q}) \cdot \det({}_\beta[L]_\beta - Z \cdot \mathbf{I}_n) \\
&= 1 \cdot \det({}_\beta[L]_\beta - Z \cdot \mathbf{I}_n) \\
&= \det({}_\beta[L]_\beta - Z \cdot \mathbf{I}_n) \\
&= p_{{}_\beta[L]_\beta}(Z).
\end{aligned}$$

Dette er netop, hvad vi skulle vise. □

Korollar 11.1.5

Med samme notation som før er $\det{}_\beta[L]_\beta = \det{}_\gamma[L]_\gamma$.

Bevis. Dette følger ved at indsætte $Z = 0$ i de karakteristiske polynomier $p_{{}_\beta[L]_\beta}(Z)$ og $p_{{}_\gamma[L]_\gamma}(Z)$. □

Vi kan nu definere det karakteristiske polynomium for en lineær afbildning $L : V \to V$, så længe V er et endeligdimensionelt vektorrum.

Definition 11.1.3

Lad $\mathbb{F}$ være et legeme, V et vektorrum over $\mathbb{F}$ af endelig dimension n og $L : V \to V$ en lineær afbildning. Da defineres *det karakteristiske polynomium* som polynomiet $p_L(Z) = \det({}_\beta[L]_\beta - Z \cdot \mathbf{I}_n) \in \mathbb{F}[Z]$, hvor β er en ordnet basis for V.

Grunden til, at denne definition giver mening, er, at valget af den ordnede basis β ifølge Sætning 11.1.4 ikke betyder noget: et andet valg vil ikke ændre det tilsvarende karakteristiske polynomium. På en lignende måde kan man, baseret på Korollar 11.1.5, definere determinanten af en sådan lineær afbildning: $\det L = \det{}_\beta[L]_\beta$.

Eksempel 11.1.6

Som et eksempel vil vi betragte en lineær afbildning, der ligner den lineære afbildning $D : \mathbb{C}[Z] \to \mathbb{C}[Z]$ fra Eksempel 10.2.1. Dog er $\mathbb{C}[Z]$ et uendeligdimensionelt vektorrum, så vi justerer først definitionsmængden og dispositionsmængden for afbildningen en smule. Lader vi V være det komplekse vektorrum af polynomier af grad højst tre, kan vi definere $\tilde{D} : V \to V$ som $p(Z) \mapsto p(Z)'$.

Spørgsmål: Hvad er det karakteristiske polynomium $p_{\tilde{D}}(\lambda)$ for den lineære afbildning $\tilde{D}$?

Svar: Lad os vælge den ordnede basis $\beta = (1, Z, Z^2, Z^3)$ for V. Da $\tilde{D}(1) = 0, \tilde{D}(Z) =$

$1, \tilde{D}(Z^2) = 2Z$, og $\tilde{D}(Z^3) = 3Z^2$, ser vi, at

$$_\beta[\tilde{D}]_\beta = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \text{og derfor} \quad _\beta[\tilde{D}]_\beta - Z \cdot \mathbf{I}_4 = \begin{bmatrix} -Z & 1 & 0 & 0 \\ 0 & -Z & 2 & 0 \\ 0 & 0 & -Z & 3 \\ 0 & 0 & 0 & -Z \end{bmatrix}.$$

Vi ser, at $_\beta[\tilde{D}]_\beta - Z \cdot \mathbf{I}_4$ er en øvre trekantsmatrix (se Definition 8.1.4). Dette betyder, at dens determinant simpelthen er produktet af elementerne i dens diagonal, se Sætning 8.1.2. Derfor er $\tilde{D}$'s karakteristiske polynomium $p_{\tilde{D}}(Z) = (-Z)^4 = Z^4$.

11.2 Egenrum

Indtil videre har vi hovedsageligt fokuseret på, hvordan man finder en matrix', henholdsvis en lineær afbildnings, egenværdier. I dette afsnit vil vi fokusere på at finde alle de mulige egenvektorer, der har en given egenværdi.

Sætning 11.2.1
Lad $\mathbb{F}$ være et legeme, V et vektorrum over $\mathbb{F}$ af endelig dimension n og $L : V \to V$ en lineær afbildning. Antag, at $\lambda \in \mathbb{F}$ er en egenværdi for L. Da er mængden

$$E_\lambda = \{\mathbf{v} \in V \mid L(\mathbf{v}) = \lambda \cdot \mathbf{v}\}$$

et underrum i V.

Bevis. Lad $\mathbf{u}, \mathbf{v} \in E_\lambda$ og $c \in \mathbb{F}$. Ifølge Lemma 9.3.2 kan vi konkludere, at E_λ er et underrum i V, hvis vi kan vise, at $\mathbf{u} + c \cdot \mathbf{v} \in E_\lambda$. Bemærk nu, at

$$L(\mathbf{u} + c \cdot \mathbf{v}) = L(\mathbf{u}) + c \cdot L(\mathbf{v}) = \lambda \cdot \mathbf{u} + c \cdot \lambda \cdot \mathbf{v} = \lambda \cdot (\mathbf{u} + c \cdot \mathbf{v}).$$

Derfor er $\mathbf{u} + c \cdot \mathbf{v} \in E_\lambda$, hvilket var det, vi skulle vise. □

For kvadratiske matricer har denne sætning en direkte konsekvens.

Korollar 11.2.2
Lad $\mathbb{F}$ være et legeme og $\mathbf{A} \in \mathbb{F}^{n \times n}$ en kvadratisk matrix. Antag, at $\lambda \in \mathbb{F}$ er en egenværdi for $\mathbf{A}$. Da er mængden $E_\lambda = \{\mathbf{v} \in V \mid \mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}\}$ et underrum i $\mathbb{F}^n$.

Bevis. Dette følger af Sætning 11.2.1 anvendt på den lineære afbildning $L_\mathbf{A} : \mathbb{F}^n \to \mathbb{F}^n$, $\mathbf{v} \mapsto \mathbf{A} \cdot \mathbf{v}$. □

For en given lineær afbildning $L : V \to V$ i et endeligdimensionelt vektorrum V, hvor λ er en egenværdi for L, kaldes underrummet E_λ for *egenrummet svarende til egenværdien* λ for den lineære afbildning L. Ligeledes, for en given kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ kaldes underrummet E_λ for egenrummet svarende til egenværdien λ for matricen $\mathbf{A}$. Bemærk, at det tilsvarende engelske begreb er *eigenspace*.

Nu hvor vi ved, at mængden af alle egenvektorer tilhørende en given egenværdi λ sammen med nulvektoren danner et underrum E_λ, kan vi beskrive alle de egenvektorer, der eksisterer for en given egenværdi, ved at angive en basis for dette underrum E_λ. Heldigvis viser dette sig at være endnu en anvendelse af teorien om systemer af lineære ligninger. Først og fremmest har vi:

Lemma 11.2.3

Lad $L : V \to V$ være en lineær afbildning mellem vektorrum over et legeme $\mathbb{F}$, og antag, at $\dim V = n$. Antag videre, at $\lambda \in \mathbb{F}$ er en egenværdi til L. Da gælder der, at $E_\lambda = \ker(L - \lambda \cdot \mathrm{id}_V)$. Tilsvarende hvis $\mathbf{A} \in \mathbb{F}^{n\times n}$ er en matrix, og $\lambda \in \mathbb{F}$ er en egenværdi for $\mathbf{A}$, da er $E_\lambda = \ker(\mathbf{A} - \lambda \cdot \mathbf{I}_n)$.

Bevis. Per definition har vi $\mathbf{v} \in E_\lambda$, hvis og kun hvis $L(\mathbf{v}) = \lambda \cdot \mathbf{v}$. Bemærk, at $L(\mathbf{v}) = \lambda \cdot \mathbf{v}$, hvis og kun hvis $(L - \lambda \cdot \mathrm{id}_n)(\mathbf{v}) = \mathbf{0}$, hvilket igen er ækvivalent med at sige, at $\mathbf{v} \in \ker(L - \lambda \cdot \mathbf{I}_n)$. Anden del af lemmaet, der involverer matricen $\mathbf{A}$, kan bevises på samme måde. $\square$

Som vi har set tidligere, svarer dét at bestemme vektorer i kernen af en matrix $\mathbf{B}$ med dét at finde løsninger til et homogent system af lineære ligninger med koefficientmatrix $\mathbf{B}$. Desuden ved vi allerede, hvordan man bestemmer en basis for løsningsrummet til et homogent system af lineære ligninger ved hjælp af Korollar 9.3.4 og Sætning 6.4.4. Derfor behøver vi ikke at udvikle nye værktøjer for at kunne bestemme en basis for egenrummet E_λ for en matrix. Heller ikke når vi beskæftiger os med den tilsvarende problemstilling for lineære afbildninger, behøver vi nye værktøjer: Sætning 10.4.2 medfører, at vi kan bestemme kernen af en lineær afbildning $L : V \to V$ ved at bestemme kernen af en matrix ${}_\beta[L]_\beta$, der repræsenterer den lineære afbildning, hvor β er en ordnet basis for V. Vi er altså allerede i stand til at finde egenvektorer i alle tilfælde. Lad os illustrere dette i to eksempler.

Eksempel 11.2.1

Lad os først betragte matricen

$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in \mathbb{C}^{2\times 2}.$$

Vi har stødt på denne matrix før i Eksempel 11.1.5, men bemærk en væsentlig forskel denne gang: i dette eksempel arbejder vi over de komplekse tal $\mathbb{C}$, da matricen er defineret som et element i $\mathbb{C}^{2\times 2}$ fremfor som et element i $\mathbb{R}^{2\times 2}$. Som i Eksempel 11.1.5 har vi, at

$$p_{\mathbf{A}}(Z) = \det(\mathbf{A} - Z \cdot \mathbf{I}_2) = \det\left(\begin{bmatrix} -Z & -1 \\ 1 & -Z \end{bmatrix}\right) = Z^2 + 1.$$

Da vi arbejder over legemet $\mathbb{C}$, har polynomiet $Z^2 + 1$ to rødder, nemlig i og $-i$.

Spørgsmål: Find en basis for egenrummet E_i.

Svar: Vi ved fra Lemma 11.2.3, at $E_i = \ker(\mathbf{A} - i \cdot \mathbf{I}_2)$. Vi har

$$\mathbf{A} - i \cdot \mathbf{I}_2 = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix}.$$

For at bestemme kernen af denne matrix, bringer vi den på reduceret trappeform:

$$\begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \underset{R_2 \leftarrow R_2 - i \cdot R_1}{\longrightarrow} \begin{bmatrix} -i & -1 \\ 0 & 0 \end{bmatrix} \underset{R_1 \leftarrow i \cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix}.$$

Dette viser, at $\mathbf{v} = (v_1, v_2) \in \ker(\mathbf{A} - i \cdot \mathbf{I}_2)$, hvis og kun hvis $v_1 = i \cdot v_2$. Derfor har vi:

$$E_i = \ker(\mathbf{A} - i \cdot \mathbf{I}_2) = \left\{ c \cdot \begin{bmatrix} i \\ 1 \end{bmatrix} \mid c \in \mathbb{C} \right\}.$$

En basis for E_i er derfor givet ved

$$\left\{ \begin{bmatrix} i \\ 1 \end{bmatrix} \right\}.$$

Dette besvarer spørgsmålet. På en lignende måde kan man vise, at en basis for E_{-i} er givet ved

$$\left\{ \begin{bmatrix} -i \\ 1 \end{bmatrix} \right\}.$$

Eksempel 11.2.2

Lad os genbesøge den lineære afbildning $\tilde{D} : V \to V$ introduceret i Eksempel 11.1.6. I det eksempel var V det komplekse vektorrum af polynomier af højst tredje grad, og $\tilde{D} : V \to V$ blev defineret ved $p(Z) \mapsto p(Z)'$. Vi har allerede set i Eksempel 11.1.6, at $p_{\tilde{D}}(Z) = Z^4$. Derfor har $\tilde{D}$ kun én egenværdi, nemlig 0.

Spørgsmål: Find en basis for egenrummet E_0.

Svar: Vi ved fra Lemma 11.2.3, at $E_0 = \ker(\tilde{D} - 0 \cdot \mathrm{id}_V) = \ker \tilde{D}$. For at finde en basis for $\ker \tilde{D}$ bestemmer vi først kernen af den matrix ${}_\beta[\tilde{D}]_\beta \in \mathbb{C}^{4\times 4}$, der repræsenterer $\tilde{D}$. Lad os vælge den ordnede basis $\beta = (1, Z, Z^2, Z^3)$ for V. Vi har allerede set i Eksempel 11.1.6, at vi i dette tilfælde har:

$$
{}_\beta[\tilde{D}]_\beta = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.
$$

Denne matrix er allerede på trappeform, og vi kan direkte aflæse, at $(v_1, v_2, v_3, v_4) \in \ker {}_\beta[\tilde{D}]_\beta$, hvis og kun hvis $v_2 = 0$, og $v_3 = 0$, og $v_4 = 0$. Derfor,

$$
\ker {}_\beta[\tilde{D}]_\beta = \left\{ c \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \mid c \in \mathbb{C} \right\}.
$$

Vi ser, at en basis for $\ker {}_\beta[\tilde{D}]_\beta$ er givet ved

$$
\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\}.
$$

Basisvektoren $(1, 0, 0, 0)$ svarer til polynomiet $1 \cdot 1 + 0 \cdot Z + 0 \cdot Z^2 + 0 \cdot Z^3 = 1$. Derfor ser vi ved hjælp af Sætning 10.4.2, at

$$
E_0 = \ker \tilde{D} = \{c \cdot 1 \mid c \in \mathbb{C}\} = \mathbb{C},
$$

og at en basis for E_0 er givet ved $\{1\}$.

Lad os afslutte dette afsnit med en teoretisk overvejelse om egenvektorer, der vil blive meget vigtig senere. Vi starter med en definition.

Definition 11.2.1

Lad $\mathbb{F}$ være et legeme, V et endeligdimensionelt vektorrum og $L : V \to V$ en lineær afbildning. Antag, at $\lambda \in \mathbb{F}$ er en egenværdi for L. Da defineres *algebraisk multiplicitet* $\mathrm{am}(\lambda)$ af egenværdien λ som multipliciteten af λ som rod i det karakteristiske polynomium $p_L(Z)$ for L. Yderligere defineres *geometrisk multiplicitet* $\mathrm{gm}(\lambda)$ af egenværdien λ som dimensionen af E_λ.

Ligeledes defineres for en egenværdi $\lambda \in \mathbb{F}$ for en kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$ $\mathrm{am}(\lambda)$ som multipliciteten af λ som rod i det karakteristiske polynomium $p_{\mathbf{A}}(Z)$ for $\mathbf{A}$, og $\mathrm{gm}(\lambda) = \dim E_\lambda$.

Eksempel 11.2.3

I Eksempel 11.2.1 er egenværdien i en rod med multiplicitet 1 i det karakteristiske polynomium

$p_{\mathbf{A}}(Z) = Z^2 + 1$. Derfor er $\mathrm{am}(i) = 1$. I det eksempel så vi også, at E_i er et vektorrum med dimension én. Derfor er $\mathrm{gm}(i) = 1$.

Eksempel 11.2.4

I Eksempel 11.2.2 er egenværdien 0 en rod med multiplicitet 4 i det karakteristiske polynomium $p_{\tilde{D}}(Z) = Z^4$. Derfor er $\mathrm{am}(0) = 4$. I det eksempel så vi også, at E_0 er et vektorrum med dimension én. Derfor er $\mathrm{gm}(0) = 1$ i dette tilfælde.

Som det sidste eksempel viser, behøver den algebraiske og den geometriske multiplicitet af en egenværdi ikke at være ens. Vi har dog følgende sætning, der siger, at $1 \le \mathrm{gm}(\lambda) \le \mathrm{am}(\lambda)$. En læser, der er villig til at acceptere dette udsagn, kan fortsætte til næste afsnit.

Sætning 11.2.4

Lad $\mathbb{F}$ være et legeme. Lad $\lambda \in \mathbb{F}$ være en egenværdi for en lineær afbildning $L : V \to V$ med $\dim V = n < \infty$ eller en egenværdi for en kvadratisk matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$. Da gælder der, at $1 \le \mathrm{gm}(\lambda) \le \mathrm{am}(\lambda) \le n$.

Bevis. Hvis λ er en egenværdi, findes der per definition mindst én tilhørende egenvektor. Derfor er $\mathrm{gm}(\lambda) = \dim E_\lambda \ge 1$.

Lad os for en egenværdi λ skrive $s = \mathrm{gm}(\lambda)$ for nemheds skyld. Vi vil bevise sætningen i et tilfælde, hvor λ er en egenværdi for en lineær afbildning $L : V \to V$ alene, da tilfældet med en matrix $\mathbf{A}$ følger ved at overveje den lineære afbildning $L_{\mathbf{A}} : \mathbb{F}^n \to \mathbb{F}^n$. Da $\dim E_\lambda = \mathrm{gm}(\lambda) = s$, indeholder enhver basis for E_λ præcis s vektorer. Lad os vælge en sådan basis, for eksempel $\{\mathbf{v}_1, \ldots, \mathbf{v}_s\}$. Vælg nu vektorer $\mathbf{v}_{s+1}, \ldots, \mathbf{v}_n \in V$, således at $\beta = \mathbf{v}_1, \ldots, \mathbf{v}_s, \mathbf{v}_{s+1}, \ldots, \mathbf{v}_n$ er en ordnet basis for V. Da $L(\mathbf{v}_i) = \lambda \cdot \mathbf{v}_i$ for alle i mellem 1 og s, har vi

$$ {}_{\beta}[L]_{\beta} = \begin{bmatrix} \lambda \cdot \mathbf{I}_s & \mathbf{B} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} $$

for nogle matricer $\mathbf{B} \in \mathbb{F}^{s \times (n-s)}$ og $\mathbf{D} \in \mathbb{F}^{(n-s)\times(n-s)}$, hvor $\mathbf{0}$ betegner en $(n-s) \times s$-nulmatrix. Så har vi

$$ {}_{\beta}[L]_{\beta} - Z \cdot \mathbf{I}_n = \begin{bmatrix} (\lambda - Z) \cdot \mathbf{I}_s & \mathbf{B} \\ \mathbf{0} & \mathbf{D} - Z \cdot \mathbf{I}_{n-s} \end{bmatrix}, $$

og derfor

$$ \begin{aligned} p_L(Z) &= \det({}_{\beta}[L]_{\beta} - Z \cdot \mathbf{I}_n) \\ &= \det\left(\begin{bmatrix} (\lambda - Z) \cdot \mathbf{I}_s & \mathbf{B} \\ \mathbf{0} & \mathbf{D} - Z \cdot \mathbf{I}_{n-s} \end{bmatrix} \right) \\ &= (\lambda - Z)^s \cdot \det(\mathbf{D} - Z \cdot \mathbf{I}_{n-s}). \end{aligned} $$

I den sidste lighed anvendte vi induktion på s og udviklede determinanten fra den første søjle for at bevise induktionens basistrin samt for at udføre induktionstrinnet. Nu er det klart, at multipliciteten af λ i $p_L(Z)$ er mindst s. Med andre ord: $\mathrm{am}(\lambda) \geq s = \mathrm{gm}(\lambda)$, hvilket er præcis, hvad vi ønskede at vise. Den sidste ulighed $\mathrm{am}(\lambda) \leq n$ følger, da $\mathrm{am}(\lambda)$ er multipliciteten af roden λ i polynomiet $p_L(Z)$, og $\deg p_L(Z) = n$. □

11.3 Diagonalisering

I dette afsnit undersøger vi, hvornår en lineær afbildning kan repræsenteres ved en særligt flot matrix: en diagonalmatrix. Vi vil nemlig undersøge, hvornår en lineær afbildning har en diagonalmatrix som afbildningsmatrix. For at opnå dette skal vi kunne vælge en særligt flot ordnet basis. Derfor starter vi med et lemma.

Lemma 11.3.1

Lad $\mathbb{F}$ være et legeme, V et endeligdimensionelt vektorrum over $\mathbb{F}$ og $L : V \to V$ en lineær afbildning. Yderligere antages det, at $\lambda_1, \ldots, \lambda_r \in \mathbb{F}$ er forskellige egenværdier for L, og der benyttes notationen $d_i = \mathrm{gm}(\lambda_i)$ for $i = 1, \ldots, r$. Hvis $(\mathbf{v}_1^{(i)}, \ldots, \mathbf{v}_{d_i}^{(i)})$ for $i = 1, \ldots, r$ er ordnede baser for E_{λ_i}, da er vektorerne

$$\mathbf{v}_1^{(1)}, \ldots, \mathbf{v}_{d_1}^{(1)}, \ldots, \mathbf{v}_1^{(r)}, \ldots, \mathbf{v}_{d_r}^{(r)}$$

lineært uafhængige.

Bevis. Vi vil bevise lemmaet ved hjælp af induktion på r.

Hvis $r = 1$, er der intet at bevise, da vi antager, at $(\mathbf{v}_1^{(1)}, \ldots, \mathbf{v}_{d_1}^{(1)})$ er en ordnet basis for E_{λ_1}. Så er vektorerne $\mathbf{v}_1^{(1)}, \ldots, \mathbf{v}_{d_1}^{(1)}$ selvfølgelig lineært uafhængige. Lad nu $r > 1$, og antag som induktionshypotese, at lemmaet er korrekt, hvis der er $r-1$ forskellige egenværdier. Antag, at

$$\sum_{i=1}^{r} \sum_{j=1}^{d_i} \alpha_{i,j} \cdot \mathbf{v}_j^{(i)} = \mathbf{0} \tag{11.2}$$

for nogle $\alpha_{i,j} \in \mathbb{F}$. Vi skal vise, at $\alpha_{i,j} = 0$ for alle $i = 1, \ldots, r$ og $j = 1, \ldots, d_i$. Ved at anvende den lineære afbildning L på denne ligning og udnytte, at $L(\mathbf{v}_j^{(i)}) = \lambda_i \cdot \mathbf{v}_j^{(i)}$, ser vi, at $\sum_{i=1}^{r} \sum_{j=1}^{d_i} \alpha_{i,j} \cdot \lambda_i \cdot \mathbf{v}_j^{(i)} = \mathbf{0}$, hvilket kan omskrives til

$$\sum_{i=1}^{r} \lambda_i \cdot \sum_{j=1}^{d_i} \alpha_{i,j} \cdot \mathbf{v}_j^{(i)} = \mathbf{0}. \tag{11.3}$$

Ganges Ligning (11.2) med λ_r, og trækkes Ligning (11.3) fra resultatet, udgår leddet svarende til $i = r$, mens højresiden stadig er $\mathbf{0}$. Med andre ord:

$$\lambda_r \cdot \sum_{i=1}^{r-1} \sum_{j=1}^{d_i} \alpha_{i,j} \cdot \mathbf{v}_j^{(i)} - \sum_{i=1}^{r-1} \lambda_i \cdot \sum_{j=1}^{d_i} \alpha_{i,j} \cdot \mathbf{v}_j^{(i)} = \lambda_r \cdot \sum_{i=1}^{r} \sum_{j=1}^{d_i} \alpha_{i,j} \cdot \mathbf{v}_j^{(i)} - \sum_{i=1}^{r} \lambda_i \cdot \sum_{j=1}^{d_i} \alpha_{i,j} \cdot \mathbf{v}_j^{(i)} = \mathbf{0}.$$

Ved at kombinere de to første summer til én, får vi:

$$\sum_{i=1}^{r-1}\sum_{j=1}^{d_i}(\lambda_r-\lambda_i)\cdot\alpha_{i,j}\cdot\mathbf{v}_j^{(i)}=\sum_{i=1}^{r-1}(\lambda_r-\lambda_i)\cdot\sum_{j=1}^{d_i}\alpha_{i,j}\cdot\mathbf{v}_j^{(i)}=\mathbf{0}.$$

Nu kan vi anvende induktionshypotesen og konkludere, at $(\lambda_r-\lambda_i)\cdot\alpha_{i,j}=0$ for $i=1,\ldots,r-1$ og $j=1,\ldots,d_i$. Da vi antog, at alle egenværdier er forskellige, ser vi, at $\lambda_r-\lambda_i\neq 0$ for alle i mellem 1 og $r-1$. Derfor er $\alpha_{i,j}=0$ for $i=1,\ldots,r-1$ og $j=1,\ldots,d_i$. Ved at indsætte dette i Ligning (11.2), får vi, at $\sum_{j=1}^{d_r}\alpha_{r,j}\cdot\mathbf{v}_j^{(r)}=\mathbf{0}$, og så kan vi også konkludere, at $\alpha_{r,j}=0$ for $j=1,\ldots d_r$, da vi antog, at $(\mathbf{v}_1^{(r)},\ldots,\mathbf{v}_{d_r}^{(r)})$ er en ordnet basis for E_{λ_r}. Dette fuldfører induktionstrinnet. Derfor kan vi ved induktionsprincippet konkludere, at lemmaet gælder for alle r. □

Som vi har set tidligere, skal der vælges en ordnet basis β for V, før en lineær afbildning $L:V\to V$ kan repræsenteres ved en matrix ${}_\beta[L]_\beta$. Vektorerne i Lemma 11.3.1 er lineært uafhængige, hvilket er en god start, men de kan ikke nødvendigvis udspænde hele rummet V. Det næste lemma afklarer, hvornår egenvektorerne udspænder V.

Lemma 11.3.2

Lad $\mathbb{F}$ være et legeme, V et vektorrum over $\mathbb{F}$ med dimension n og $L:V\to V$ en lineær afbildning. Da er følgende to udsagn ækvivalente:

(i) Egenvektorerne tilhørende L udspænder V.

(ii) Det karakteristiske polynomium for L er på formen

$$p_L(Z)=(-1)^n\cdot(Z-\lambda_1)^{m_1}\cdots(Z-\lambda_r)^{m_r}$$

for nogle $\lambda_1,\ldots,\lambda_r\in\mathbb{F}$ og positive heltal $m_1,\ldots,m_r$. Yderligere er de algebraiske og geometriske multipliciteter ens for enhver egenværdi λ_i, altså $\mathrm{am}(\lambda_i)=\mathrm{gm}(\lambda_i)$ for $i=1,\ldots,r$.

Bevis. For at vise, at de to punkter er logisk ækvivalente, viser vi først (i) ⇒ (ii) og derefter (ii) ⇒ (i).

(i) ⇒ (ii): antag, at egenvektorerne for L udspænder V. Da kan vi finde en basis S for V, der kun består af egenvektorer. Lad $\lambda_1,\ldots,\lambda_r\in\mathbb{F}$ være egenværdierne for L, og sortér egenvektorerne i S, således at egenvektorerne med egenværdi λ_1 kommer først, derefter dem med egenværdi λ_2 og så videre, afsluttende med egenvektorerne i S med egenværdi λ_r. Vi har så konstrueret en ordnet basis

$$\beta=(\mathbf{v}_1^{(1)},\ldots,\mathbf{v}_{n_1}^{(1)},\ldots,\mathbf{v}_1^{(r)},\ldots,\mathbf{v}_{n_r}^{(r)}),$$

hvor vektorerne $\mathbf{v}_1^{(i)},\ldots,\mathbf{v}_{n_i}^{(i)}$ for $i=1,\ldots,r$ er egenvektorerne i S med egenværdi λ_i.

Nu har vi på den ene side $n_1 + n_2 + \cdots + n_r = n$, da antallet af vektorer i den ordnede basis β er det samme som dimensionen af V. På den anden side har vi for alle i, at $n_i \leq \mathrm{gm}(\lambda_i)$, da $\mathbf{v}_1^{(i)}, \ldots, \mathbf{v}_{n_i}^{(i)}$ er lineært uafhængige vektorer i E_{λ_i}, og $\dim E_{\lambda_i} = \mathrm{gm}(\lambda_i)$. Derfor har vi:

$$\begin{aligned} n &= n_1 + \cdots + n_r \\ &\leq \mathrm{gm}(\lambda_1) + \cdots + \mathrm{gm}(\lambda_r) \\ &\leq \mathrm{am}(\lambda_1) + \cdots + \mathrm{am}(\lambda_r) \\ &\leq \deg p_L(Z) \\ &= n. \end{aligned}$$

Da vi både startede og sluttede med n, skal alle uligheder være ligheder. Dette viser, at $\mathrm{gm}(\lambda_i) = \mathrm{am}(\lambda_i)$ for alle $i = 1, \ldots, r$, og at $p_L(Z)$ er på den form, der er angivet i punkt (ii).

(ii) $\Rightarrow$ (i): antag nu, at $p_L(Z) = (-1)^n \cdot (Z - \lambda_1)^{m_1} \cdots (Z - \lambda_r)^{m_r}$ for nogle forskellige $\lambda_1, \ldots, \lambda_r \in \mathbb{F}$ og for positive heltal $m_1, \ldots, m_r$, hvor $\mathrm{am}(\lambda_i) = \mathrm{gm}(\lambda_i)$ for alle $i = 1, \ldots, r$. Bemærk, at vi per definition har $m_i = \mathrm{am}(\lambda_i)$, hvilket igen medfører, at $m_i = \mathrm{gm}(\lambda_i)$, da vi antager, at $\mathrm{am}(\lambda_i) = \mathrm{gm}(\lambda_i)$ for alle i. Vi konkluderer, at $n = \deg p_L(Z) = \mathrm{gm}(\lambda_1) + \cdots + \mathrm{gm}(\lambda_r)$. På den anden side kan vi ifølge Lemma 11.3.1 finde præcis $\mathrm{gm}(\lambda_1) + \cdots + \mathrm{gm}(\lambda_r)$ lineært uafhængige egenvektorer for L. Ved at kombinere disse udsagn kan vi konkludere, at vi kan finde en ordnet basis for V, der består af egenvektorer. Disse egenvektorerne udspænder V, hvilket er det, vi ønskede at vise. □

Nu er vi klar til at vise hovedresultatet af dette Afsnit.

Definition 11.3.1

Lad en lineær afbildning $L : V \to V$ være givet, hvor V er et endeligdimensionelt vektorrum over et legeme $\mathbb{F}$. Da siger man, at L kan *diagonaliseres*, hvis der findes en ordnet basis β for V, således at den tilsvarende afbildningsmatrix ${}_\beta[L]_\beta$ er en diagonalmatrix. Ligeledes, hvis $\mathbf{A} \in \mathbb{F}^{n \times n}$ er en kvadratisk matrix, så siger man, at $\mathbf{A}$ kan diagonaliseres, hvis $\mathbf{A}$ er similær med en diagonalmatrix.

Sætning 11.3.3

Lad V være et endeligdimensionelt vektorrum over et legeme $\mathbb{F}$. En lineær afbildning $L : V \to V$ kan diagonaliseres, hvis og kun hvis det karakteristiske polynomium for L er på formen $p_L(Z) = (-1)^n \cdot (Z - \lambda_1)^{m_1} \cdots (Z - \lambda_r)^{m_r}$ for nogle $\lambda_1, \ldots, \lambda_r \in \mathbb{F}$, og $\mathrm{am}(\lambda_i) = \mathrm{gm}(\lambda_i)$ for enhver egenværdi λ_i.

Bevis. Ved at bruge Lemma 11.3.2 er det nok at vise, at L kan diagonaliseres, hvis og kun hvis dens egenvektorer udspænder V.

Derfor antager vi først, at L kan diagonaliseres. Da findes der en ordnet basis β for V, således at ${}_\beta[L]_\beta$ er en diagonalmatrix. Men dette medfører, at enhver vektor i β er en egenvektor. Derfor kan V udspændes af egenvektorer.

Antag omvendt, at V kan udspændes af egenvektorer. Da findes der en ordnet basis $\beta = \mathbf{v}_1, \ldots, \mathbf{v}_n$ for V, der kun indeholder egenvektorer. Den tilsvarende matrix ${}_\beta[L]_\beta$ er en diagonalmatrix med egenværdier tilhørende $\mathbf{v}_1, \ldots, \mathbf{v}_n$ i sin diagonal. □

Korollar 11.3.4

En matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$ kan diagonaliseres, hvis og kun hvis det karakteristiske polynomium for $\mathbf{A}$ er på formen $p_\mathbf{A}(Z) = (-1)^n \cdot (Z - \lambda_1)^{m_1} \cdots (Z - \lambda_r)^{m_r}$ for nogle $\lambda_1, \ldots, \lambda_r \in \mathbb{F}$, og $\mathrm{am}(\lambda_i) = \mathrm{gm}(\lambda_i)$ for enhver egenværdi λ_i.

Bevis. Dette følger af Sætning 11.3.3 anvendt på den lineære afbildning $L_\mathbf{A} : \mathbb{F}^n \to \mathbb{F}^n$. □

Korollar 11.3.5

Lad V være et endeligdimensionelt, komplekst vektorrum. En lineær afbildning $L : V \to V$ kan diagonaliseres, hvis og kun hvis $\mathrm{am}(\lambda_i) = \mathrm{gm}(\lambda_i)$ for enhver egenværdi λ_i for L. Ligeledes er en kompleks matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ similær med en diagonalmatrix, hvis og kun hvis $\mathrm{am}(\lambda_i) = \mathrm{gm}(\lambda_i)$ for enhver egenværdi λ_i for $\mathbf{A}$.

Bevis. Hvis legemet, vi arbejder over, er $\mathbb{C}$, følger det af Sætning 4.6.3, at det karakteristiske polynomium $p_L(Z)$ kan skrives som et produkt af dets ledende koefficient samt led på formen $Z - \lambda$. Derfor er denne betingelse i Sætning 11.3.3 altid opfyldt, hvis $\mathbb{F} = \mathbb{C}$, og den kan derfor fjernes. Sætning 11.3.3 medfører derefter det, vi ønsker. Beviset for korollaret i tilfælde af en kompleks matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ udføres på tilsvarende vis. □

Eksempel 11.3.1

Betragt den lineære afbildning $\tilde{D} : V \to V$ introduceret i Eksempel 11.1.6. I det eksempel var V det komplekse vektorrum af polynomier af højst tredje grad, og $\tilde{D} : V \to V$ blev defineret ved $p(Z) \mapsto p(Z)'$.

Spørgsmål: Kan den lineære afbildning $\tilde{D}$ diagonaliseres?

Svar: Fra Eksempel 11.1.6 ved vi, at $p_{\tilde{D}}(Z) = Z^4$. Med notationen fra Sætning 11.3.3 har vi, at $r = 1$, og $\lambda_1 = 0$. Desuden er $\mathrm{am}(0) = 4$, da 0 er en rod med multiplicitet fire i $p_{\tilde{D}}(Z)$. I Eksempel 11.2.2 har vi set, at E_0 er et endimensionelt vektorrum med basis $\{1\}$. Derfor er $\mathrm{gm}(0) = \dim E_0 = 1$. Da $\mathrm{gm}(0) < \mathrm{am}(0)$, medfører Sætning 11.3.3, at den lineære afbildning $\tilde{D}$ ikke kan diagonaliseres.

Eksempel 11.3.2

Betragt matricen

$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

som vi arbejde med i Eksempel 11.2.1, og som også optrådte i Eksempel 11.1.5.

Spørgsmål 1: Kan matricen $\mathbf{A}$ diagonaliseres, når vi arbejder over de reelle tal $\mathbb{R}$? Hvis ja, bestem en matrix $\mathbf{Q} \in \mathbb{R}^{2\times 2}$, således at $\mathbf{Q}^{-1} \cdot \mathbf{A} \cdot \mathbf{Q}$ er en diagonalmatrix.

Spørgsmål 2: Kan matricen $\mathbf{A}$ diagonaliseres, når vi arbejder over de komplekse tal $\mathbb{C}$? Hvis ja, bestem en matrix $\mathbf{Q} \in \mathbb{C}^{2\times 2}$, således at $\mathbf{Q}^{-1} \cdot \mathbf{A} \cdot \mathbf{Q}$ er en diagonalmatrix.

Svar på Spørgsmål 1: Vi fandt i Eksempel 11.1.5 frem til, at $p_{\mathbf{A}}(Z) = Z^2 + 1$. Da $Z^2 + 1$ ikke har nogen reelle rødder, kan det ikke skrives på den form, der kræves i Korollar 11.3.4. Derfor er matricen $\mathbf{A}$ ikke diagonalisérbar over $\mathbb{R}$.

Svar på Spørgsmål 2: Det karakteristiske polynomium $p_{\mathbf{A}}(Z) = Z^2 + 1$ har to komplekse rødder, nemlig i og $-i$. Desuden har vi, at $Z^2 + 1 = (Z - i) \cdot (Z + i)$. Derfor er $\mathrm{am}(i) = 1$ og $\mathrm{am}(-i) = 1$. Da vi ved fra Sætning 11.2.4, at $1 \le \mathrm{gm}(\lambda) \le \mathrm{am}(\lambda)$ for enhver egenværdi λ, konkluderer vi, at $\mathrm{gm}(i) = \mathrm{am}(i) = 1$, og $\mathrm{gm}(-i) = \mathrm{am}(-i) = 1$. Dermed er alle betingelser i Korollar 11.3.4 opfyldt. Vi konkluderer, at den givne matrix $\mathbf{A}$ er diagonalisérbar over de komplekse tal.

Nu bestemmer vi eksplicit en invertibel matrix $\mathbf{Q} \in \mathbb{C}^{2\times 2}$, således at $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ er en diagonalmatrix. Lad os ved ϵ betegne standardbasen for $\mathbb{C}^2$. Da er ${}_{\epsilon}[L_{\mathbf{A}}]_{\epsilon} = \mathbf{A}$. For at diagonalisere $\mathbf{A}$, diagonaliserer vi simpelthen den tilsvarende lineære afbildning $L_{\mathbf{A}}$. For at gøre det, skal vi finde en ordnet basis for $\mathbb{C}^2$ bestående af egenvektorer. I Eksempel 11.2.1 så vi, at:

$$E_i \quad \text{har basis} \quad \left\{ \begin{bmatrix} i \\ 1 \end{bmatrix} \right\}, \qquad \text{og} \qquad E_{-i} \quad \text{har basis} \quad \left\{ \begin{bmatrix} -i \\ 1 \end{bmatrix} \right\}.$$

Derfor er $\beta = \left(\begin{bmatrix} i \\ 1 \end{bmatrix}, \begin{bmatrix} -i \\ 1 \end{bmatrix} \right)$ en ordnet basis for $\mathbb{C}^2$ bestående kun af egenvektorer. Ved anvendelse af denne ordnede basis har vi, at afbildningsmatricen ${}_{\beta}[L_{\mathbf{A}}]_{\beta}$ er en diagonalmatrix med egenværdierne tilhørende vektorerne i den ordnede basis β i sin diagonal. Derfor er

$${}_{\beta}[L_{\mathbf{A}}]_{\beta} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

For at finde matricen $\mathbf{Q}$, således at $\mathbf{Q}^{-1} \cdot \mathbf{A} \cdot \mathbf{Q}$ er en diagonalmatrix, ser vi nu, at

$${}_{\beta}[L_{\mathbf{A}}]_{\beta} = {}_{\beta}[\mathrm{id}_{\mathbb{C}^2}]_{\epsilon} \cdot {}_{\epsilon}[L_{\mathbf{A}}]_{\epsilon} \cdot {}_{\epsilon}[\mathrm{id}_{\mathbb{C}^2}]_{\beta} = {}_{\epsilon}[\mathrm{id}_{\mathbb{C}^2}]_{\beta}^{-1} \cdot \mathbf{A} \cdot {}_{\epsilon}[\mathrm{id}_{\mathbb{C}^2}]_{\beta}.$$

Derfor kan vi simpelthen vælge $\mathbf{Q} = {}_{\epsilon}[\mathrm{id}_{\mathbb{C}^2}]_{\beta}$, koordinatskiftematricen fra β-koordinater til ϵ-koordinater. Denne matrix indeholder egenværdierne i β som søjler. Derfor er

$$\mathbf{Q} = {}_{\epsilon}[\mathrm{id}_{\mathbb{C}^2}]_{\beta} = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$$

den matrix, vi leder efter. Ved en kontrol ser vi ganske rigtigt, at

$$\begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

11.4 Fibonacci-tal genbesøgt

I Eksempel 5.1.2, mere præcist i Ligning (5.3), gav vi et eksempel på en rekursivt defineret sekvens af tal $F_1, F_2, F_3, \ldots$ kaldet Fibonacci-tallene:

$$F_n = \begin{cases} 1 & \text{hvis } n = 1, \\ 1 & \text{hvis } n = 2, \\ F_{n-1} + F_{n-2} & \text{hvis } n \geq 3. \end{cases} \tag{11.4}$$

Denne rekursion kan også udtrykkes ved hjælp af matricer. Faktisk ser man direkte fra Ligning (11.4), at

$$\begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} F_{n-1} \\ F_{n-2} \end{bmatrix} \quad \text{for alle } n \geq 3.$$

Denne matrixform gør det muligt at finde en lukket formel som beskrivelse af Fibonacci-tallene. Først og fremmest har vi følgende:

Lemma 11.4.1

For alle $n \geq 2$ gælder der, at:

$$\begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{n-2} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Bevis. Dette kan vises ved hjælp af induktion på n med basistrin $n = 2$. Detaljerne overlades til læseren. □

For at finde en lukket formel for F_n, er det nok at finde en lukket formel for potenser af matricen

$$\mathbf{P} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Vi vil diagonalisere $\mathbf{P}$ for at gøre dette. Det viser sig nemlig, at hvis en matrix kan diagonaliseres, er det muligt at finde et lukket formeludtryk for dens potenser:

Lemma 11.4.2

Lad $\mathbb{F}$ være et legeme og $\mathbf{A} \in \mathbb{F}^{n \times n}$ en kvadratisk matrix. Lad $\mathbf{Q} \in \mathbb{F}^{n \times n}$ være en invertibel matrix, således at $\mathbf{Q}^{-1} \cdot \mathbf{A} \cdot \mathbf{Q}$ er en diagonalmatrix $\mathbf{D}$ med elementerne $d_1, \ldots, d_n$ i sin diagonal. Da er

$$\mathbf{A}^n = \mathbf{Q} \cdot \mathbf{D}^n \cdot \mathbf{Q}^{-1}.$$

Desuden er $\mathbf{D}^n$ en diagonalmatrix med elementerne $d_1^n, \ldots, d_n^n$ i sin diagonal.

Bevis. Med induktion på n kan man vise, at $(\mathbf{Q}^{-1} \cdot \mathbf{A} \cdot \mathbf{Q})^n = \mathbf{Q}^{-1} \cdot \mathbf{A}^n \cdot \mathbf{Q}$ for alle $n \geq 1$. Da $\mathbf{Q}^{-1} \cdot \mathbf{A} \cdot \mathbf{Q} = \mathbf{D}$, følger resultatet. At vise, at $\mathbf{D}^n$ er en diagonalmatrix med elementerne $d_1^n, \ldots, d_n^n$ i sin diagonal, er let ved induktion på n. □

Pointen med dette lemma er, at det gør beregningen af potenser af en matrix relativt let, hvis matricen er diagonalisérbar. Lad os nu vende tilbage til matricen $\mathbf{P}$. Det karakteristiske polynomium for $\mathbf{P}$ er

$$p_{\mathbf{P}}(Z) = \det\left(\begin{bmatrix} 1-Z & 1 \\ 1 & -Z \end{bmatrix}\right) = Z^2 - Z - 1.$$

Derfor er egenværdierne for $\mathbf{P}$ $\lambda_1 = \frac{1+\sqrt{5}}{2}$ og $\lambda_2 = \frac{1-\sqrt{5}}{2}$. Dette betyder allerede, at matricen $\mathbf{P}$ er diagonalisérbar. For at finde den ønskede koordinatskiftematrix, skal vi finde en basis for egenrummene. For at finde en basis for egenrummet E_{λ_1}, bemærker vi, at

$$\mathbf{P} - \lambda_1 \cdot \mathbf{I}_2 = \begin{bmatrix} \frac{1-\sqrt{5}}{2} & 1 \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix} \underset{R_2 \leftarrow R_2 - \lambda_2 \cdot R_1}{\longrightarrow} \begin{bmatrix} \frac{1-\sqrt{5}}{2} & 1 \\ 0 & 0 \end{bmatrix}.$$

Dermed har vi, at en basis for E_{λ_1} er givet ved

$$\left\{\begin{bmatrix} -1 \\ \frac{1-\sqrt{5}}{2} \end{bmatrix}\right\}.$$

Ligeledes kan man vise, at en basis for E_{λ_2} er givet ved

$$\left\{\begin{bmatrix} -1 \\ \frac{1+\sqrt{5}}{2} \end{bmatrix}\right\}.$$

Derfor er

$$\begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ \frac{1-\sqrt{5}}{2} & \frac{1+\sqrt{5}}{2} \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 & -1 \\ \frac{1-\sqrt{5}}{2} & \frac{1+\sqrt{5}}{2} \end{bmatrix},$$

hvilket medfører, at

$$\mathbf{P} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ \frac{1-\sqrt{5}}{2} & \frac{1+\sqrt{5}}{2} \end{bmatrix} \cdot \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \cdot \begin{bmatrix} -1 & -1 \\ \frac{1-\sqrt{5}}{2} & \frac{1+\sqrt{5}}{2} \end{bmatrix}^{-1}.$$

Ved anvendelse af Lemma 11.4.2 til udregning af potenser af $\mathbf{P}$ samt Lemma 11.4.1 ser vi nu, at

$$\begin{aligned} \begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{n-2} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & -1 \\ \frac{1-\sqrt{5}}{2} & \frac{1+\sqrt{5}}{2} \end{bmatrix} \cdot \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{n-2} & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^{n-2} \end{bmatrix} \cdot \begin{bmatrix} -1 & -1 \\ \frac{1-\sqrt{5}}{2} & \frac{1+\sqrt{5}}{2} \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \end{aligned}$$

Efter at have beregnet alle matrixprodukterne på højre side, opnår man at

$$\begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} \cdot \left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}} \cdot \left(\frac{1-\sqrt{5}}{2}\right)^n \\ \frac{1}{\sqrt{5}} \cdot \left(\frac{1+\sqrt{5}}{2}\right)^{n-1} - \frac{1}{\sqrt{5}} \cdot \left(\frac{1-\sqrt{5}}{2}\right)^{n-1} \end{bmatrix},$$

hvilket forklarer, hvor Ligning (5.4) kom fra.

I dette afsnit fokuserede vi på Fibonacci-tallene, men meget lignende teknikker kan benyttes til at finde lukkede formeludtryk for andre rekursivt definerede sekvenser af tal. Vi vil ikke forfølge dette yderligere her.

11.5 Ekstra: Hvad hvis diagonalisering ikke er mulig?

Dette afsnit er ikke nødvendig læsning og kan springes over. Det er ment som ekstra materiale for en læser, der har tiden og motivationen til det.

Som vi har set i det foregående afsnit, er diagonalisering af en lineær afbildning $L : V \to V$ ikke altid mulig. I dette afsnit vil vi arbejde med den berømte *Jordans normalform*. Nøglen til diagonalisering var at undersøge egenrummet $E_\lambda = \ker(L - \lambda \cdot \mathrm{id}_V)$ for en given egenværdi λ. Vi definerede $\mathrm{gm}(\lambda) = \dim E_\lambda$ og har set, at diagonalisering af en matrix eller lineær afbildning kræver, at betingelsen $\mathrm{gm}(\lambda) = \mathrm{am}(\lambda)$ er opfyldt. Hvis $\mathrm{gm}(\lambda) < \mathrm{am}(\lambda)$, viser det sig, at man kan undersøge kernerne af potenser af den lineære afbildning $L - \lambda \cdot \mathrm{id}_V$. Her skal den i'te potens f^i af en funktion $f : V \to V$ forstås som den i-foldige sammensætning af f med sig selv (altså $f^1 = f$, $f^2 = f \circ f$ og så videre).

Lemma 11.5.1

Lad $\mathbb{F}$ være et legeme, V et n-dimensionelt vektorrum over $\mathbb{F}$ og $L : V \to V$ en lineær afbildning. Antag yderligere, at $\lambda \in \mathbb{F}$ er en egenværdi for L. Da gælder der, at

$$\ker(L - \lambda \cdot \mathrm{id}_V) \subseteq \ker((L - \lambda \cdot \mathrm{id}_V)^2) \subseteq \ker((L - \lambda \cdot \mathrm{id}_V)^3) \subseteq \dots$$

Yderligere,

(i) hvis $\ker((L - \lambda \cdot \mathrm{id}_V)^i) = \ker((L - \lambda \cdot \mathrm{id}_V)^{i+1})$ for et positivt heltal i, så er $\ker((L - \lambda \cdot \mathrm{id}_V)^i) = \ker((L - \lambda \cdot \mathrm{id}_V)^m)$ for alle $m \geq i$, og

(ii) $\ker((L - \lambda \cdot \mathrm{id}_V)^n) = \ker((L - \lambda \cdot \mathrm{id}_V)^{n+1})$.

Bevis. Det er klart, at kernen af $(L - \lambda \cdot \mathrm{id}_V)^{i+1}$ for alle i er et underrum i kernen af $(L - \lambda \cdot \mathrm{id}_V)^i$. Hvis ligheden holder for nogle i, så opnår vi ved dimensionssætningen for lineære afbildninger, se Korollar 10.4.3, at også billederne af de lineære afbildninger

$(L - \lambda \cdot \mathrm{id}_V)^{i+1}$ og $(L - \lambda \cdot \mathrm{id}_V)^i$ er ens. Så har vi, at

$$\begin{aligned}
\mathrm{im}(L - \lambda \cdot \mathrm{id}_V)^{i+2} &= (L - \lambda \cdot \mathrm{id}_V)^{i+2}(V) \\
&= (L - \lambda \cdot \mathrm{id}_V)((L - \lambda \cdot \mathrm{id}_V)^{i+1}(V)) \\
&= (L - \lambda \cdot \mathrm{id}_V)((L - \lambda \cdot \mathrm{id}_V)^{i}(V)) \\
&= (L - \lambda \cdot \mathrm{id}_V)^{i+1}(V) \\
&= \mathrm{im}(L - \lambda \cdot \mathrm{id}_V)^{i+1} \\
&= \mathrm{im}(L - \lambda \cdot \mathrm{id}_V)^{i}.
\end{aligned}$$

Men så ser vi, igen ved at bruge dimensionssætningen for lineære afbildninger, at kernen af $(L - \lambda \cdot \mathrm{id}_V)^{i+2}$ er lig med kernen af $(L - \lambda \cdot \mathrm{id}_V)^i$. Ved induktion på m kan man på samme måde vise, at kernen af $(L - \lambda \cdot \mathrm{id}_V)^m$ for enhver $m \geq i$ er lig med kernen af $(L - \lambda \cdot \mathrm{id}_V)^i$.

Betragt nu sekvensen af underrum:

$$\ker(L - \lambda \cdot \mathrm{id}_V) \subseteq \ker((L - \lambda \cdot \mathrm{id}_V)^2) \subseteq \ker((L - \lambda \cdot \mathrm{id}_V)^3) \subseteq \dots$$

Fra det foregående ved vi, at hvis ligheden holder et eller andet sted i sekvensen, så vil lighederne holde fra da af. Derfor findes der et $e \geq 1$, således at

$$\ker(L - \lambda \cdot \mathrm{id}_V) \subsetneq \cdots \subsetneq \ker((L - \lambda \cdot \mathrm{id}_V)^e) = \ker((L - \lambda \cdot \mathrm{id}_V)^{e+1}) = \dots$$

For hver skarp inklusion øges dimensionen af underrummet med mindst én. Da $\dim V = n$, og $\dim \ker(L - \lambda \cdot \mathrm{id}_V) = \dim E_\lambda \geq 1$, kan dette højst ske n gange. Derfor er $e \leq n$. Særligt er $\ker((L - \lambda \cdot \mathrm{id}_V)^n) = \ker((L - \lambda \cdot \mathrm{id}_V)^{n+1})$. $\square$

Sætning 11.5.2

Lad $\mathbb{F}$ være et legeme, V et n-dimensionelt vektorrum over $\mathbb{F}$ og $L : V \to V$ en lineær afbildning. Antag yderligere, at $\lambda \in \mathbb{F}$ er en egenværdi for L, og skriv $U = \mathrm{im}(L - \lambda \cdot \mathrm{id}_V)^n$ og $W = \ker(L - \lambda \cdot \mathrm{id}_V)^n$. Da gælder der, at

(i) $L(U) \subseteq U$, og $L(W) \subseteq W$.

(ii) $\dim U + \dim W = \dim V$, og $U \cap W = \{\mathbf{0}\}$.

(iii) Enhver vektor i V kan skrives som en sum af en vektor i U og en vektor i W.

Bevis. Vi har set i det tredje punkt i Lemma 11.5.1, at $W = \ker(L - \lambda \cdot \mathrm{id}_V)^n = \ker(L - \lambda \cdot \mathrm{id}_V)^{n+1}$. Vælg nu et $\mathbf{w} \in W$. Da er også $(L - \lambda \cdot \mathrm{id}_V)(\mathbf{w}) \in W$, da $(L - \lambda \cdot \mathrm{id}_V)^n((L - \lambda \cdot \mathrm{id}_V)(\mathbf{w})) = (L - \lambda \cdot \mathrm{id}_V)((L - \lambda \cdot \mathrm{id}_V)^n(\mathbf{w})) = (L - \lambda \cdot \mathrm{id}_V)(\mathbf{0}) = \mathbf{0}$. Derfor er $L(\mathbf{w}) - \lambda \cdot \mathbf{w} \in W$, hvilket medfører, at $L(\mathbf{w}) \in W$. Vi kan konkludere, at $L(W) \subseteq W$. Ligeledes, hvis $\mathbf{u} \in U$, så er $(L - \lambda \cdot \mathrm{id}_V)(\mathbf{u}) \in U$, fordi vi, hvis $\mathbf{u} = (L - \lambda \cdot \mathrm{id}_V)^n(\mathbf{v})$ for nogle $\mathbf{v} \in V$, har, at $(L - \lambda \cdot \mathrm{id}_V)(\mathbf{u}) = (L - \lambda \cdot \mathrm{id}_V)((L - \lambda \cdot \mathrm{id}_V)^n(\mathbf{v})) = (L - \lambda \cdot \mathrm{id}_V)^n((L - \lambda \cdot \mathrm{id}_V)(\mathbf{v})) \in U$. Derfor er $L(\mathbf{u}) - \lambda \cdot \mathbf{u} \in U$, hvilket medfører, at $L(\mathbf{u}) \in U$. Vi kan konkludere, at $L(U) \subseteq U$.

Dimensionssætningen for lineære afbildninger anvendt på den lineære afbildning $(L - \lambda \cdot \mathrm{id}_V)^n : V \to V$ medfører straks, at $\dim U + \dim W = \dim V$. Nu beviser vi, at $U \cap W = \{\mathbf{0}\}$. Lad $\mathbf{u} \in U \cap W$. Vi ønsker at vise, at $\mathbf{u} = \mathbf{0}$. Først og fremmest, da $\mathbf{u} \in U$, findes der $\mathbf{v} \in V$, således at $\mathbf{u} = (L - \lambda \cdot \mathrm{id}_V)^n(\mathbf{v})$. For det andet, da $\mathbf{u} \in W$, har vi $(L - \lambda \cdot \mathrm{id}_V)^n(\mathbf{u}) = \mathbf{0}$. Ved at kombinere disse to, ser vi, at

$$(L - \lambda \cdot \mathrm{id}_V)^{2n}(\mathbf{v}) = (L - \lambda \cdot \mathrm{id}_V)^n(\mathbf{u}) = \mathbf{0}.$$

Med andre ord, $\mathbf{v} \in \ker(L - \lambda \cdot \mathrm{id}_V)^{2n}$. Men Lemma 11.5.1 medfører, at $\ker(L - \lambda \cdot \mathrm{id}_V)^{2n} = \ker(L - \lambda \cdot \mathrm{id}_V)^n$, og derfor $\mathbf{v} \in \ker(L - \lambda \cdot \mathrm{id}_V)^n$. Men så er $\mathbf{u} = (L - \lambda \cdot \mathrm{id}_V)^n(\mathbf{v}) = \mathbf{0}$, hvilket er, hvad vi ønskede at vise.

Givet en ordnet basis $\beta_U = (\mathbf{u}_1, \ldots, \mathbf{u}_r)$ for U og en ordnet basis $\beta_W = (\mathbf{w}_1, \ldots, \mathbf{w}_s)$ for W. Forenes de to, får vi en ordnet basis $\beta = (\beta_U, \beta_W)$ for V. Faktisk kan det faktum, at $U \cap W = \{\mathbf{0}\}$ bruges til at vise, at vektorerne i β er lineært uafhængige, mens identiteten $\dim U + \dim W = \dim V$ medfører, at β indeholder præcis n vektorer. Givet en vilkårligt valgt $\mathbf{v} \in V$ kan vi nu skrive $\mathbf{v}$ på præcis én måde som en linearkombination af $\mathbf{u}_i$ og $\mathbf{w}_j$, nemlig $\mathbf{v} = \sum_i \alpha_i \cdot \mathbf{u}_i + \sum_j \beta_j \cdot \mathbf{w}_j$. Vi ser, at $\sum_i \alpha_i \cdot \mathbf{u}_i \in U$ og $\sum_j \beta_j \cdot \mathbf{w}_j \in W$, hvoraf det sidste punkt i sætningen følger. $\square$

Korollar 11.5.3

Skriv med samme notation som i Sætning 11.5.2 $p_L(Z) = (\lambda - Z)^{\mathrm{am}(\lambda)} \cdot q(Z)$ for en passende valgt $q(Z) \in \mathbb{F}[Z]$. Betegn ved $L|_W : W \to W$, henholdsvis $L|_U : U \to U$, lineære afbildninger opnået ved afgrænsning af definitionsmængden og dispositionsmængden for L til U, henholdsvis W. Da er $p_L(Z) = p_{L|_U}(Z) \cdot p_{L|_W}(Z)$, og λ er ikke en rod i $p_{L|_U}(Z)$.

Bevis. For en givet ordnet basis $\beta_U = (\mathbf{u}_1, \ldots, \mathbf{u}_r)$ for U og en givet ordnet basis $\beta_W = (\mathbf{w}_1, \ldots, \mathbf{w}_s)$ for W har vi allerede set i beviset for Sætning 11.5.2, at $\beta = (\beta_U, \beta_W)$ er en ordnet basis for V. Da vi ved, at $L(U) \subseteq U$, og $L(W) \subseteq W$, vil matricen ${}_\beta[L]_\beta \in \mathbb{F}^{n \times n}$ have formen

$${}_\beta[L]_\beta = \begin{bmatrix} {}_{\beta_U}[L]_{\beta_U} & \mathbf{0} \\ \mathbf{0} & {}_{\beta_W}[L]_{\beta_W} \end{bmatrix}. \tag{11.5}$$

Dette medfører, at $p_L(Z) = p_{L|_U}(Z) \cdot p_{L|_W}(Z)$. Bemærk nu, at λ ikke kan være en rod i $p_{L|_U}(Z)$. Hvis dette var tilfældet, ville der eksistere en $\mathbf{u} \in U$ forskellig fra nul, således at $L|_U(\mathbf{u}) = \lambda \cdot \mathbf{u}$. Da vi per definitionen af den lineære afbildning $L|_U$ har $L|_U(\mathbf{u}) = L(\mathbf{u})$, ville dette medføre, at $(L - \lambda \cdot \mathrm{id}_V)(\mathbf{u}) = \mathbf{0}$. Men så ville vi have $\mathbf{u} \in \ker(L - \lambda \cdot \mathrm{id}_V)$, hvilket ville medføre, at $\mathbf{u} \in \ker(L - \lambda \cdot \mathrm{id}_V)^n = W$. Da vi har set, at $U \cap W = \{\mathbf{0}\}$, ville vi opnå, at $\mathbf{u} = \mathbf{0}$, i modstrid med vores antagelse. $\square$

Lad os nu vende tilbage til det, vi prøver at opnå: at finde en matrix, der repræsenterer L, der er så simpel som mulig. Ligning (11.5) er et vigtigt skridt på vejen. Faktisk har vi reduceret problemstillingen til to enklere opgaver: at finde en simpel matrix, der repræsenterer $L|_U$, og en, der repræsenterer $L|_W$. Desuden er λ ikke en rod i det karakteristiske polynomium for $L|_U$, så for at håndtere egenværdien λ for L behøver vi kun at fortsætte undersøgelsen af den lineære afbildning $L|_W$. Lad os først få en intuitiv idé om, hvad der foregår. Hvis $\mathrm{am}(\lambda) = \mathrm{gm}(\lambda)$, kan vi finde en ordnet basis β_W for W, der består udelukkende af egenvektorer for L, alle med λ som egenværdi. Så er

$$_{\beta_W}[L]_{\beta_W} = [\lambda \cdot \mathbf{I}_s],$$

hvor $s = \dim W$. Hvad vi gjorde i forrige afsnit var essentielt set blot at gentage denne procedure for en anden egenværdi og opdele matricen, der repræsenterer $L|_U$, i mindre blokke. Så længe den algebraiske og geometriske multiplicitet af egenværdierne altid er den samme, vil vi ende med at have diagonaliseret hele matricen.

Men hvad sker der så, hvis $\mathrm{am}(\lambda) > \mathrm{gm}(\lambda)$? Vi kan stadig finde en ordnet basis β_W for W, der består af egenvektorer til egenværdien λ, men vi har også brug for nogle flere vektorer i β_W, der ikke er egenvektorer. Sagt på en anden måde: hvis $\mathrm{am}(\lambda) = \mathrm{gm}(\lambda)$, så er $W = E_\lambda$, således at $\ker(L - \lambda \cdot \mathrm{id}_V) = \ker((L - \lambda \cdot \mathrm{id}_V)^n)$. Men hvis $\mathrm{am}(\lambda) > \mathrm{gm}(\lambda)$, så indeholder W E_λ, men er ikke lig med det. Så tilsyneladende er $\ker(L - \lambda \cdot \mathrm{id}_V) \subsetneq \ker((L - \lambda \cdot \mathrm{id}_V)^n)$. Dette medfører specifikt, at $\ker(L - \lambda \cdot \mathrm{id}_V) \subsetneq \ker((L - \lambda \cdot \mathrm{id}_V)^2)$ ved anvendelse af Lemma 11.5.1. Hvis vi vælger en vektor $\mathbf{w} \in \ker((L - \lambda \cdot \mathrm{id}_V)^2) \setminus \ker(L - \lambda \cdot \mathrm{id}_V)$, har den den pæne egenskab, at $L(\mathbf{w}) - \lambda \cdot \mathbf{w} = (L - \lambda \cdot \mathrm{id}_V)(\mathbf{w}) \in \ker(L - \lambda \cdot \mathrm{id}_V) = E_\lambda$. Lad os definere $\mathbf{v} = L(\mathbf{w}) - \lambda \cdot \mathbf{w}$. Vi ser, at $L(\mathbf{v}) = \lambda \cdot \mathbf{v}$, da $\mathbf{v} \in E_\lambda$, og $L(\mathbf{w}) = \lambda \cdot \mathbf{w} + \mathbf{v}$, ud fra den måde vi definerede $\mathbf{v}$ på. Så hvis vi kun ser på effekten af L på det todimensionelle underrum i V udspændt af $\mathbf{v}$ og $\mathbf{w}$, som har den ordnede basis $(\mathbf{v}, \mathbf{w})$, kan vi repræsentere L ved matricen

$$\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.$$

Dette giver en første idé om, hvad man kan forvente mere generelt om en matrix, der repræsenterer $L|_W$. Specifikt motiverer det til følgende:

Definition 11.5.1

Lad $\mathbb{F}$ være et legeme, og $\lambda \in \mathbb{F}$. En *Jordanblok* af størrelse e er en matrix $\mathbf{J}_e(\lambda) \in \mathbb{F}^{e \times e}$ på formen

$$\mathbf{J}_e(\lambda) = \begin{bmatrix} \lambda & 1 & & \mathbf{0} \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & \lambda \end{bmatrix}.$$

Lemma 11.5.4

Lad $L : V \to V$ være en lineær afbildning, $\dim V = n$ og λ en egenværdi for L. Lad yderligere $W = \ker((L - \lambda \cdot \mathrm{id}_V)^n)$. Da findes der en ordnet basis for W, således at $L|_W : W \to W$, restriktionen af L til W, har en afbildningsmatrix $\mathbf{D}(\lambda)$ på formen

$$\mathbf{D}(\lambda) = \begin{bmatrix} \mathbf{J}_{e_1}(\lambda) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{J}_{e_s}(\lambda) \end{bmatrix}$$

for et positivt heltal s og positive heltal $e_1, \ldots, e_s$.

Bevis. Det vil være praktisk at skrive $W_i = \ker((L - \lambda \cdot \mathrm{id}_V)^i)$ og $r_i = \dim W_i$. Bemærk, at $W_1 = E_\lambda$. Lad nu e være den største eksponent, således at $W_{e-1} \subsetneq W_e$. Så er $W_e = W$ og $r_1 < r_2 < \cdots < r_e = \dim W$. Lad $\beta = (\mathbf{w}_1, \ldots, \mathbf{w}_{r_e})$ være en ordnet basis for W med den yderligere egenskab, at $(\mathbf{w}_1, \ldots, \mathbf{w}_{r_i})$ er en ordnet basis for W_i for alle i.

Vi konstruerer nu gradvist en anden ordnet basis for W, som vi kan kalde γ. Først og fremmest tilføjer vi til γ vektorerne $\mathbf{w}_i$ for ethvert i mellem $r_{e-1} + 1$ og r_e. Grundet denne konstruktion vil vektoren $\mathbf{w}_i$ for ethvert i mellem $r_{e-1} + 1$ og r_e ligge i W_e men ikke i W_{e-1}. Betragt nu vektorerne $\mathbf{w}_{i,j} = (L - \lambda \cdot \mathrm{id}_V)^j(\mathbf{w}_i)$ for $j = 0, \ldots, e-1$. Bemærk, at vektoren $\mathbf{w}_{i,j}$ ligger i W_{e-j} men ikke i W_{e-j-1}. Bemærk også, at $\mathbf{w}_i = \mathbf{w}_{i,0}$ for ethvert i mellem $r_{e-1} + 1$ og r_e.

Først hævder vi, at vektorerne $\mathbf{w}_{i,e-1}$ er lineært uafhængige. Hvis $\sum_{i=r_{e-1}+1}^{r_e} \alpha_i \cdot \mathbf{w}_{i,e-1} = 0$, så er $\sum_{i=r_{e-1}+1}^{r_e} \alpha_i \cdot \mathbf{w}_i \in W_{e-1}$ og derfor i udspændingen af vektorerne $\mathbf{w}_1, \ldots, \mathbf{w}_{r_{e-1}}$. Men da vektorerne $\mathbf{w}_1, \ldots, \mathbf{w}_{r_e}$ er lineært uafhængige, medfører dette, at $\alpha_i = 0$ for alle $i = r_{e-1} + 1, \ldots, r_e$. Dernæst hævder vi, at vektorerne $\mathbf{w}_{i,j}$ med $i = r_{e-1} + 1, \ldots, r_e$ og $j = 0, \ldots, e-1$ er lineært uafhængige. Hvis $\sum_i \sum_{j=0}^{e-1} \alpha_{i,j} \cdot \mathbf{w}_{i,j} = 0$, så anvender vi $(L - \lambda \cdot \mathrm{id}_V)^{e-1}$ og får ligningen $\sum_i \alpha_{i,0} \cdot \mathbf{w}_{i,e-1} = 0$. Derfor er $\alpha_{i,0} = 0$ for alle i. Ved at anvende lavere og lavere potenser af $(L - \lambda \cdot \mathrm{id}_V)$ til ligningen $\sum_i \sum_{j=0}^{e-1} \alpha_{i,j} \cdot \mathbf{w}_{i,j} = 0$ opnår man induktivt, at $\alpha_{i,j} = 0$ for alle i og j.

Derfor giver det mening at inkludere alle vektorerne $\mathbf{w}_{i,j}$ i γ. Mere præcist definerer vi nu $\gamma = (\mathbf{w}_{r_{e-1}+1,e-1}, \ldots, \mathbf{w}_{r_{e-1}+1,0}, \ldots, \mathbf{w}_{r_e,e-1}, \ldots, \mathbf{w}_{r_e,0})$. Vi har $(L - \lambda \cdot \mathrm{id}_V)(\mathbf{w}_{i,j}) = \mathbf{w}_{i,j+1}$ for $j = 0, \ldots, e-2$, og $(L - \lambda \cdot \mathrm{id}_V)(\mathbf{w}_{i,e-1}) = \mathbf{0}$. Dette medfører, at

$$L(\mathbf{w}_{i,j}) = \lambda \cdot \mathbf{w}_{i,j} + \mathbf{w}_{i,j+1} \quad \text{for } j = 0, \ldots, e-2 \text{ og} \quad L(\mathbf{w}_{i,e-1}) = \lambda \cdot \mathbf{w}_{i,e-1}.$$

Vi har nu rent faktisk vist, at restriktionen af L til underrummet udspændt af $\mathbf{w}_{i,j}$ kan repræsenteres ved en blokdiagonalmatrix med $r_e - r_{e-1}$ mange matricer $\mathbf{J}_e(\lambda)$ i sin diagonal.

For ethvert j mellem 0 og $e-1$ udspænder vektorerne $\mathbf{w}_{i,j}$ med i varierende fra $r_{e-1} + 1$ til r_e et underrum i W_{e-j}. Hvis dette underrum er lig med W_{e-j} for alle j, så er γ en ordnet basis for W, der giver anledning til en matrix, der repræsenterer W på Jordans normalform, som vi har set. Hvis ikke, så lad $\tilde{j}$ være den mindste værdi af j, således at dette underrum ikke er hele W_{e-j}, og definér $\tilde{e} = e - \tilde{j}$. Lad så $\tilde{\mathbf{w}}_1, \ldots, \tilde{\mathbf{w}}_d \in W_{\tilde{e}}$ være vektorer, således at

$$(\mathbf{w}_1, \ldots, \mathbf{w}_{r_{\tilde{e}-1}}, \tilde{\mathbf{w}}_1, \ldots, \tilde{\mathbf{w}}_d, \mathbf{w}_{r_{e-1}+1,\tilde{j}}, \ldots, \mathbf{w}_{r_e,\tilde{j}})$$

er en ordnet basis for $W_{\tilde{e}}$. Vi fortsætter på samme måde som før og definerer vektorer $\tilde{\mathbf{w}}_{i,j} = (L - \lambda \cdot \mathrm{id}_V)^j(\tilde{\mathbf{w}}_i)$, som vi tilføjer til γ, og som vil give anledning til Jordanblokke af størrelse $\tilde{e}$ i matricen, der repræsenterer L.

Fortsætter vi på denne måde, vil vi ende med en ordnet basis γ, der giver anledning til en matrix på Jordans normalform, som repræsenterer restriktionen af L til W. □

Sætning 11.5.5

Lad $\mathbb{F}$ være et legeme, V et endeligdimensionelt vektorrum og $L : V \to V$. Antag, at der eksisterer forskellige $\lambda_1, \ldots, \lambda_r \in \mathbb{F}$ samt positive heltal $m_1, \ldots, m_r$, således at $p_L(Z) = (-1)^n \cdot (Z - \lambda_1)^{m_1} \cdots (Z - \lambda_r)^{m_r}$. Da kan den lineære afbildning repræsenteres ved en matrix på formen

$$\begin{bmatrix} \mathbf{D}(\lambda_1) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{D}(\lambda_r) \end{bmatrix},$$

hvor hver matrix $\mathbf{D}(\lambda_i) \in \mathbb{F}^{m_i \times m_i}$ er på formen som i Lemma 11.5.4.

Bevis. Vi beviser sætningen med induktion på antallet af egenværdier. Vi bruger den samme notation for U og W som i Sætning 11.5.2. Hvis $r = 1$, er $W = V$ og derfor $L|_W = L$. Derfor følger resultatet fra Lemma 11.5.4. Lad nu $r > 1$. Lad $\lambda = \lambda_r \in \mathbb{F}$ være en egenværdi for L. Fra Lemma 11.5.4 konkluderer vi, at vi kan vælge en ordnet basis for W, således at $L|_W$ er repræsenteret ved en blokdiagonalmatrix med Jordanblokke $\mathbf{J}_{e_i}(\lambda_r)$ i sin diagonal. Yderligere har vi fra Korollar 11.5.3, at λ_r ikke er en egenværdi for $L|_U$, mens det karakteristiske polynomium for $L|_U$ er en divisor af $P_L(Z)$. Derfor er induktionshypotesen gyldig. □

Matricen givet i Sætning 11.5.5 siges at være *Jordans normalform* af matricen $\mathbf{A}$.

Korollar 11.5.6

Lad $\mathbf{A} \in \mathbb{C}^{n \times n}$ være en kompleks matrix. Da er $\mathbf{A}$ similær med en matrix på Jordans normalform.

Bevis. Hvis vi arbejder over de komplekse tal, følger det fra Sætning 4.6.3, at $p_{\mathbf{A}}(Z)$ kan skrives som et produkt af dets ledende koefficient, som er $(-1)^n$, og led af formen $Z - \lambda$. Derfor gælder Sætning 11.5.5. □

Kapitel 12

Systemer af lineære ordinære differentialligninger af første orden med konstante koefficienter

I dette kapitel vil vi lære nogle familier af differentialligninger at kende. Differentialligninger anvendes ofte til modellering af processer, der forekommer i naturen. De optræder i næsten alle områder af anvendt eksakt videnskab, som (kvante)mekanik, (bio)kemi, dynamik i biologiske systemer, bygningsteknik, studiet af elektriske komponenter og kredsløb og mange flere. Teorien om differentialligninger er omfattende, og vi vil i denne tekst tage et første kig på nogle særlige tilfælde. Inden vi går i gang, vil vi fastsætte et par konventioner og notationsformer, som vi vil benytte i resten af dette kapitel.

Som vi har set i de forgående kapitler, angiver en funktion $f : A \to B$ generelt en afbildning mellem to mængder. I dette kapitel vil vi altid antage, at funktionens definitionsmængde A er mængden af reelle tal $\mathbb{R}$. Hvis dispositionsmængden B er lig med $\mathbb{R}$, vil vi kalde den en reel funktion. Hvis $B = \mathbb{C}$, er der tale om en kompleks funktion med reelt input. Bemærk, at man på engelsk, hvis $B = \mathbb{R}$, kalder funktionen en *real-valued function*, og hvis $B = \mathbb{C}$, kalder den en *complex-valued function*.

Både reelle funktioner og komplekse funktioner med reelt input forekommer mange steder i matematikken, især i analysen. De teknikker og værktøjer fra lineær algebra, som vi har diskuteret hidtil i de foregående kapitler, kan også anvendes i analysen. Vi vil komme til at se, hvordan værktøjer fra lineær algebra kan benyttes til at løse bestemte typer af differentialligninger. I grove træk kan man opfatte en differentialligning som en måde at finde funktioner på, der har yderligere egenskaber, som involverer funktionens afledte. Vi forudsætter, at læseren er bekendt med den afledte af en reel funktion. Når en reel funktion $f : \mathbb{R} \to \mathbb{R}$ er givet, betegner vi ved f' den afledte af f, forudsat at den eksisterer. Funktionen

$f' : \mathbb{R} \to \mathbb{R}$ er igen en reel funktion, og derfor kan man forsøge at bestemme den afledte af f'. Hvis den eksisterer, betegnes den typisk ved f'' eller $f^{(2)}$. På denne måde kan man rekursivt definere for $n \geq 3$, at funktionen $f^{(n)} : \mathbb{R} \to \mathbb{R}$ er den afledte af $f^{(n-1)}$, forudsat at den eksisterer. Vi har set denne notation i Eksempel 9.3.4, hvor vi introducerede det reelle vektorrum $C_\infty(\mathbb{R})$, som består af alle uendeligdifferentiable (reelle) funktioner. Man kan også skrive $f^{(0)} = f$ og $f^{(1)} = f'$. I teorien om reelle funktioner og komplekse funktioner med reelt input er det ganske almindeligt at skrive en funktion som $f(t)$ i stedet for at skrive $f : \mathbb{R} \to \mathbb{R}$ (for reelle funktioner) eller $f : \mathbb{R} \to \mathbb{C}$ (for komplekse funktioner med reelt input). I resten af dette afsnit vil vi ofte benytte denne skrivemåde.

Lad os nu dykke ned i, hvad der menes med en n'te-ordens ordinær differentialligning. Vi vil i resten af denne tekst ofte for nemheds skyld skrive *ODE* i stedet for ordinær differentialligning, hvilket er en forkortelse for det tilsvarende engelske begreb "ordinary differential equation".

Definition 12.0.1

Lad n være et naturligt tal. En n'te-ordens ordinær differentialligning (ODE) er en ligning på formen

$$F(f^{(n)}(t), \ldots, f'(t), f(t), t) = 0,$$

hvor F er en funktion, der tager $n+2$ variable som input.

En løsning til en sådan ODE er en reel funktion $f(t)$, således at

$$F(f^{(n)}(t), \ldots, f'(t), f(t), t) = 0$$

for alle $t \in \mathbb{R}$. Betragt som et første lille eksempel følgende: funktionen $f(t) = e^t$ er en løsning til ODE'en $f'(t) - f(t) = 0$, fordi der gælder, at $(e^t)' = e^t$. Vi kommer til at se mange flere eksempler senere.

Der findes mange variationer og mere raffinerede definitioner. I nogle tilfælde ønsker man for eksempel, at $F(f^{(n)}(t), \ldots, f'(t), f(t), t) = 0$ kun behøver at gælde for alle t i en delmængde af $\mathbb{R}$. Men alt, hvad vi behøver på dette tidspunkt, er en intuitiv forståelse af, hvad en ODE er, og derfor vil vi ikke gå i dybden med sådanne variationer her.

I denne tekst vil vi koncentrere os om en særligt interessant type ODE, hvortil værktøjer fra den lineære algebra kan anvendes, kaldet en lineær ODE. Lad os definerer en sådan i det følgende:

Definition 12.0.2

En *lineær* ODE er en ligning på formen $L(f(t)) = q(t)$, hvor $q(t) \in C_\infty(\mathbb{R})$ er en funktion, og $L : C_\infty(\mathbb{R}) \to C_\infty(\mathbb{R})$ er en lineær afbildning. Hvis $q(t)$ er nulfunktionen, kaldes den lineære ODE en *homogen* lineær ODE, og ellers kaldes den en *inhomogen* lineær ODE.

Også her findes variationer: løsningerne behøver ikke nødvendigvis at være uendeligdifferentiable, ej heller definerede på hele $\mathbb{R}$. Men som nævnt vil vi ikke dykke ned i sådanne detaljer i denne tekst men fremhæver i stedet den lineære algebras rolle.

Tager vi et tilbageblik på den lineære afbildning fra Eksempel 10.2.8, der blev defineret som $L(f(t)) = f(t) - f'(t)$, ser vi nu, at den lineære ODE $L(f(t)) = q(t)$ simpelthen svarer til differentialligningen $f'(t) - f(t) = q(t)$. Pointen med at studere lineære ODE'er i forbindelse med lineær algebra er, at så gælder Sætning 10.4.1, som kan beskrive strukturen af løsningsmængden. Lad os se på nogle flere eksempler.

Eksempel 12.0.1

Er følgende ODE'er lineære eller ej? Afgør for de ODE'er, der er lineære, om de er homogene eller inhomogene.

(a) $f''(t) + 2f'(t) + f(t) = \cos(t)$

(b) $e^t \cdot f'(t) + \cos(t) \cdot f(t) = 0$

(c) $(f'(t))^2 + f(t) = 0$

Svar:

(a) Den givne ODE er lineær med $L(f(t)) = f''(t) + 2f'(t) + f(t)$ og $q(t) = \cos(t)$. Da $q(t)$ ikke er nulfunktionen, er det en inhomogen lineær ODE.

(b) Den givne ODE er lineær med $L(f(t)) = e^t \cdot f'(t) + \cos(t) \cdot f(t)$ og $q(t) = 0$. Da $q(t)$ er nulfunktionen, er det en homogen lineær ODE.

(c) Den givne ODE er ikke lineær på grund af leddet $(f'(t))^2$.

Man er ofte primært interesseret i reelle funktioner som løsninger til en ODE, men nogle gange er det praktisk også at se på løsninger, der er komplekse funktioner med reelt input. For os vil hovedårsagen være, at sådanne funktioner kan benyttes til at finde tilsvarende reelle funktioner som løsninger til en ODE. Lad os derfor forklare, hvordan man bestemmer den afledte af en kompleks funktion med reelt input. For en givet funktion $f : \mathbb{R} \to \mathbb{C}$ kan man for ethvert $t \in \mathbb{R}$ skrive $f(t) = f_1(t) + if_2(t)$, hvor $f_1(t) = \mathrm{Re}(f(t))$ er realdelen af $f(t)$, og $f_2(t) = \mathrm{Im}(f(t))$ er imaginærdelen af $f(t)$. På denne måde giver enhver kompleks funktion af typen $f : \mathbb{R} \to \mathbb{C}$ anledning til de to reelle funktioner $\mathrm{Re}(f) : \mathbb{R} \to \mathbb{R}$ defineret som $t \mapsto \mathrm{Re}(f(t))$ og $\mathrm{Im}(f) : \mathbb{R} \to \mathbb{R}$ defineret som $t \mapsto \mathrm{Im}(f(t))$. Omvendt kan vi, givet to reelle funktioner $f_1 : \mathbb{R} \to \mathbb{R}$ og $f_2 : \mathbb{R} \to \mathbb{R}$, definere en kompleks funktion $f = f_1 + i \cdot f_2$ som $t \mapsto f_1(t) + i \cdot f_2(t)$. Hvis de afledte af f_1 og f_2 eksisterer, kan vi nemlig definere den afledte af f som funktionen $f' = f_1' + i \cdot f_2'$. Tilsvarende kan vi for ethvert ikke-negativt heltal n definere den n'te-afledte som $f^{(n)} = f_1^{(n)} + i \cdot f_2^{(n)}$, forudsat at både $f_1^{(n)}$ og $f_2^{(n)}$ eksisterer. Med disse konventioner på plads kan vi derfor også tale om komplekse funktioner med reelt input som løsninger til en ODE. Vi vil se eksempler på sådanne løsninger senere.

Efter denne korte introduktion til, hvad en lineær ODE er, tager vi et kig på nogle særlige tilfælde og eksempler i de følgende afsnit.

12.1 Lineære førsteordens ODE'er

Som Definition 12.0.1 viser, udgør en førsteordens ODE en relation mellem en funktion $f(t)$ og dens afledte $f'(t)$. For eksempel er $f'(t) = f(t)$ en førsteordens ODE, men også et mere kompliceret udtryk som

$$\sin(f(t)f'(t)) = f'(t)^2 + e^t$$

er en førsteordens ODE. For at bringe disse eksempler på formen som i Definition 12.0.1, omskriver vi blot udtrykkene, så højresiden bliver lig nul. For eksempel kan det førstnævnte udtryk i forrige sætning skrives som $f'(t) - f(t) = 0$, mens det andetnævnte kan skrives som $\sin(f(t)f'(t)) - f'(t)^2 - e^t = 0$. Lad os gennemgå et par eksempler på førsteordens ODE'er.

Eksempel 12.1.1

Undersøg, om funktionen $f(t) = e^{2t}$ er en løsning til en af følgende ODE'er:

(a) $f'(t) - 2f(t) = 0$

(b) $f'(t)^2 - 4f(t) = 0$

(c) $\ln(f'(t)) - \ln(f(t)) = \ln(2)$

Svar:

(a) Ved anvendelse af kædereglen får vi, at $f'(t) = (e^{2t})' = e^{2t}(2t)' = e^{2t}2 = 2e^{2t}$. Derfor gælder der, at

$$f'(t) - 2f(t) = 2e^{2t} - 2e^{2t} = 0.$$

Vi konkluderer, at funktionen $f(t) = e^{2t}$ er en løsning til ODE'en $f'(t) = 2f(t)$.

(b) Vi har set, at $f'(t) = 2e^{2t}$. Derfor gælder der, at

$$f'(t)^2 - 4f(t) = (2e^{2t})^2 - 4e^{2t} = 4(e^{2t})^2 - 4e^{2t} = 4e^{4t} - 4e^{2t} \neq 0.$$

Derfor er funktionen $f(t) = e^{2t}$ ikke en løsning til ODE'en $f'(t)^2 - 4f(t) = 0$.

(c) Hvis $f(t) = e^{2t}$, får vi, at

$$\ln(f'(t)) - \ln(f(t)) = \ln(2e^{2t}) - \ln(e^{2t}) = \ln(2) + \ln(e^{2t}) - \ln(e^{2t}) = \ln(2),$$

så funktionen $f(t) = e^{2t}$ er en løsning til ODE'en $\ln(f'(t)) - \ln(f(t)) = \ln(2)$.

Lad os tage endnu et kig på ODE'en $f'(t) = f(t)$. Vi nævnte tidligere, at funktionen $f(t) = e^t$ er en løsning til denne ODE. Men den er ikke den eneste løsning. For eksempel opfylder begge funktionerne $f(t) = 2e^t$ og $f(t) = -5e^t$ også, at $f'(t) = f(t)$. Faktisk er enhver funktion på formen $f(t) = c \cdot e^t$, hvor $c \in \mathbb{R}$ er en konstant, en løsning til ODE'en $f'(t) = f(t)$.

Man kan vise, at enhver løsning til ODE'en $f'(t) = f(t)$ er på formen $f(t) = c \cdot e^t$. En sådan beskrivelse af alle de mulige løsninger til en ODE kaldes dens *fuldstændige løsning*. Vi benyttede begrebet fuldstændig løsning på en lignende måde, da vi beskrev løsninger til systemer af lineære ligninger i et tidligere kapitel. Ved brug af denne terminologi kan vi sige, at den fuldstændige løsning til ODE'en $f'(t) = f(t)$ er givet ved $f(t) = c \cdot e^t$, hvor $c \in \mathbb{R}$.

Det kan være svært at finde et eksplicit udtryk for den fuldstændige løsning til en ODE. Men for nogle klasser af ODE'er er det muligt. Vi vil nu se på en sådan klasse. En ODE på formen

$$f'(t) = a(t)f(t) + q(t), \tag{12.1}$$

hvor $a(t)$ og $q(t)$ er funktioner i variablen t, kaldes en *lineær førsteordens ODE*. Bemærk, at funktionen $q(t)$ på engelsk kaldes *forcing function*. Da der ikke findes et tilsvarende ofte-benyttet dansk begreb, vil vi i denne tekst benytte det engelske begreb i danske sammenhænge også. En ODE som i Ligning (12.1) er faktisk lineær ifølge Definition 12.0.2: vælg blot en lineær afbildning defineret som $L(f(t)) = f'(t) - a(t)f(t)$ samt en funktion $q(t)$ som givet, og så viser ligningen $L(f(t)) = q(t)$, at $f'(t) - a(t)f(t) = q(t)$, hvilket er ækvivalent med ligningen $f'(t) = a(t)f(t) + q(t)$.

Som et eksempel er ODE'en $f'(t) = f(t)$ en lineær førsteordens ODE: vælg blot $a : \mathbb{R} \to \mathbb{R}$ som funktionen defineret ved $t \mapsto 1$ og $q : \mathbb{R} \to \mathbb{R}$ som funktionen defineret ved $t \mapsto 0$, og Ligning (12.1) er hermed reduceret til ligningen $f'(t) = f(t)$.

Ifølge Definition 12.0.2 kaldes ODE'en fra Ligning (12.1) *homogen*, hvis dens forcing function $q(t)$ er nulfunktionen, og ellers *inhomogen*.

Det viser sig, at man kan opstille en formel for den fuldstændige løsning til en lineær førsteordens ODE. Lad os først få overblik over notation og begreber, som vi vil benytte til denne formel: vi vil ved $P(t)$ betegne en *stamfunktion* for funktionen $a(t)$, hvilket vil sige en funktion, der opfylder $P'(t) = a(t)$. Bemærk, at det tilsvarende engelske begreb er *primitive function*, alternativt *antiderivative*. Vi vil også i resten af dette afsnit antage, at funktionen $a(t)$ har en sådan stamfunktion. Det bliver desuden nødvendigt for os at antage, at funktionen $e^{P(t)}q(t)$ har en stamfunktion. Man kan vise, at disse antagelser er sande, hvis for eksempel både funktionen $a(t)$ og $q(t)$ er differentiable. Er alt dette opfyldt, kan vi opstille følgende resultat:

Sætning 12.1.1

Den fuldstændige løsning til ODE'en $f'(t) = a(t)f(t) + q(t)$ er givet ved

$$f(t) = e^{P(t)} \int e^{-P(t)} q(t) dt.$$

Bevis. Husk, at $P'(t) = a(t)$. Ved først at benytte produktreglen og derefter kædereglen får vi, at

$$\left(e^{-P(t)} f(t)\right)' = \left(e^{-P(t)}\right)' f(t) + e^{-P(t)} f'(t) = -e^{-P(t)} a(t) f(t) + e^{-P(t)} f'(t).$$

Derfor gælder følgende:

$$\begin{aligned} f'(t) = a(t)f(t) + q(t) &\Leftrightarrow e^{-P(t)}f'(t) - e^{-P(t)}a(t)f(t) = e^{-P(t)}q(t) \\ &\Leftrightarrow \left(e^{-P(t)}f(t)\right)' = e^{-P(t)}q(t) \\ &\Leftrightarrow e^{-P(t)}f(t) = \int e^{-P(t)}q(t)dt \\ &\Leftrightarrow f(t) = e^{P(t)}\int e^{-P(t)}q(t)dt. \end{aligned}$$

□

Når man udregner integralet i Sætning 12.1.1, bør man ikke glemme integrationskonstanten, da denne konstant er nødvendig, når man finder den fuldstændige løsning. Lad os se på nogle eksempler.

Eksempel 12.1.2

Bestem den fuldstændige løsning til følgende ODE'er:

(a) $f'(t) = f(t)$

(b) $f'(t) = -\sin(t)f(t) + \sin(t)$

(c) $f'(t) = -t^{-1}f(t) + 1$, med $t > 0$

Svar:

(a) Ved omskrivning af $f'(t) = f(t)$ som $f'(t) - f(t) = 0$ ser vi, at vi kan anvende Sætning 12.1.1 ved at benytte $a(t) = 1$ og $q(t) = 0$. En stamfunktion til $a(t) = 1$ er for eksempel givet ved $P(t) = t$. Da får vi, at den fuldstændige løsning er givet ved

$$f(t) = e^t \int e^{-t} 0\, dt = e^t \int 0\, dt = e^t c = ce^t.$$

Dette stemmer overens med den fuldstændige løsning, vi fandt tidligere for denne ODE.

(b) Vi kan benytte Sætning 12.1.1 med $a(t) = -\sin(t)$ og $q(t) = \sin(t)$. Vi kan vælge $P(t) = \cos(t)$, og vi får derved den ønskede fuldstændige løsning til at være

$$f(t) = e^{\cos(t)} \int e^{-\cos(t)} \sin(t)dt = e^{\cos(t)}\left(e^{-\cos(t)} + c\right) = 1 + ce^{\cos(t)}.$$

(c) Sætning 12.1.1 kan benyttes med $a(t) = -t^{-1} = -1/t$ og $q(t) = 1$. Da $t > 0$, betyder det, at vi kan vælge $P(t) = -\ln(t)$. Den fuldstændige løsning til ODE'en $f'(t) = -t^{-1}f(t) + 1$ bliver så

$$f(t) = e^{-\ln(t)} \int e^{\ln(t)}dt = (1/e^{\ln(t)}) \int t dt = \frac{1}{t}\left(\frac{1}{2}t^2 + c\right) = \frac{t}{2} + \frac{c}{t}.$$

Et vigtigt særtilfælde af Sætning 12.1.1 er, når funktionen a er en konstant funktion, som vi kan betegne $a(t) = a_0$ for alle t. Da bliver Sætning 12.1.1 simplere at arbejde med, hvilket tydeliggøres i følgende korollar.

Korollar 12.1.2

Lad $a_0 \in \mathbb{R}$, og lad $q(t)$ være en differentiabel, reel funktion. Da har ODE'en $f'(t) = a_0 f(t) + q(t)$ den fuldstændige løsning $f(t) = e^{a_0 t} \int e^{-a_0 t} q(t) dt$. Mere konkret, hvis $Q(t)$ er en stamfunktion til $e^{-a_0 t} q(t)$, kan den fuldstændige løsning skrives som $f(t) = c \cdot e^{a_0 t} + e^{a_0 t} Q(t)$, hvor $c \in \mathbb{R}$ er vilkårlig.

Som nævnt tidligere anvendes ODE'er til at modellere processer, der forekommer i naturen. Den fuldstændige løsning til en ODE beskriver alle mulighederne for processens opførsel. For at finde ud af, hvilken af mulighederne der er den rigtige i en given situation, har man brug for yderligere information, som man normalt kan opnå ved at udføre målinger. Én mulighed er at beskrive opførslen af funktionen f for en specifik værdi af variablen t. Man kunne forestille sig, at man måler den præcise tilstand af processen i begyndelsen af et eksperiment. Matematisk set vil vi gøre dette ved at opstille en *begyndelsesværdibetingelse*, det vil sige en betingelse på formen $f(t_0) = y_0$ påført en funktion $f(t)$. Bemærk, at det tilsvarende engelske begreb er *initial-value condition*.

Definition 12.1.1

Givet en reel funktion $f(t)$ og reelle tal t_0 og y_0, således at $f(t_0) = y_0$. Da siges funktionen $f(t)$ at opfylde *begyndelsesværdibetingelsen* $f(t_0) = y_0$.

Det viser sig, at en funktion $f : \mathbb{R} \to \mathbb{R}$ i mange interessante anvendelser er fuldstændig bestemt, hvis den opfylder både en førsteordens ODE og en begyndelsesværdibetingelse. Lad os give en beskrivelse af situationen for generelle ODE'er.

Definition 12.1.2

Lad $f(t)$ være en reel funktion, der opfylder følgende:

(i) $f(t)$ er en løsning til en n'te-ordens ODE $F(f^{(n)}(t), \ldots, f'(t), f(t), t) = 0$.

(ii) $f(t)$ opfylder alle begyndelsesværdibetingelserne $f(t_0) = y_0, f'(t_0) = y_1, \ldots, f^{(n-1)}(t_0) = y_{n-1}$ for givne $t_0 \in \mathbb{R}$ og værdier $y_0, y_1, \ldots, y_{n-1} \in \mathbb{R}$.

Samlet kaldes disse to betingelser et *begyndelsesværdiproblem*. Funktionen $f(t)$ siges at være en løsning på begyndelsesværdiproblemet.

For en førsteordens ODE $F(f'(t), f(t), t) = 0$ svarer dette til at sige, at $f(t)$ er en løsning på begyndelsesværdiproblemet, hvis det opfylder, at både

(i) $F(f'(t), f(t), t) = 0$, og

(ii) $f(t_0) = y_0$ for givne $t_0 \in \mathbb{R}$ og en værdi y_0.

Derfor er Definition 12.1.1 blot et særtilfælde af Definition 12.1.2, nemlig det tilfælde hvor $n = 1$.

Strategien til at løse et begyndelsesværdiproblem følger ofte samme mønster. Først bestemmes den fuldstændige løsning til den givne ODE. Denne fuldstændige løsning vil indeholde nogle parametre. Lad os kalde en sådan parameter c. Derefter benyttes begyndelsesværdibetingelsen til at bestemme c. Den resulterende funktion er den ønskede løsning. Lad os se på to eksempler med førsteordens ODE'er.

Eksempel 12.1.3

Løs følgende begyndelsesværdiproblemer. Det vil sige, bestem i hvert tilfælde den funktion $f(t)$, som opfylder

(a) ODE'en $f'(t) = f(t)$ og begyndelsesværdibetingelsen $f(0) = 7$.

(b) ODE'en $f'(t) + \sin(t)f(t) = \sin(t)$ og begyndelsesværdibetingelsen $f(\pi) = 2$.

Svar:

Bemærk, at vi allerede har bestemt den fuldstændige løsning til de givne to ODE'er i Eksempel 12.1.2. Lad os nu se på hvert begyndelsesværdiproblem separat.

(a) Vi har allerede set, at den fuldstændige løsning til $f'(t) = f(t)$ er givet ved $f(t) = ce^t$. Tricket er nu at evaluere $f(t)$ i 0 og sammenligne resultatet med begyndelsesværdibetingelsen. Vi får, at $f(0) = c$, men ifølge begyndelsesværdibetingelsen skal vi have $f(0) = 7$. Dette betyder, at $c = 7$. Nu hvor vi kender c, har vi, at den ønskede funktion $f : \mathbb{R} \to \mathbb{R}$ er givet ved

$$f(t) = 7e^t.$$

(b) Den fuldstændige løsning er i dette tilfælde givet ved $f(t) = 1 + ce^{\cos(t)}$. Ved at benytte begyndelsesværdibetingelsen får vi, at $2 = f(\pi) = 1 + ce^{\cos(\pi)} = 1 + ce^{-1}$. Dette betyder, at $ce^{-1} = 1$, og derfor $c = e$. Dermed er den ønskede funktion $f : \mathbb{R} \to \mathbb{R}$ givet ved

$$f(t) = 1 + e \cdot e^{\cos(t)} = 1 + e^{1+\cos(t)}.$$

Før vi går i dybden med mere generelle ODE'er, bør vi først etablere en fin egenskab ved den komplekse eksponentialfunktion. Vi ved, at den afledte af den reelle funktion $f(t) = e^{\lambda t}$ blot er $f'(t) = \lambda e^{\lambda t}$ for ethvert $\lambda \in \mathbb{R}$. Det viser sig, at dette også gælder for den komplekse eksponentialfunktion:

Lemma 12.1.3

Lad $\lambda \in \mathbb{C}$, og betragt den komplekse funktion med reelt input $f : \mathbb{R} \to \mathbb{C}$ defineret som $f(t) = e^{\lambda t}$. Da gælder der, at $\mathrm{Re}(f) = e^{\mathrm{Re}(\lambda)t}\cos(\mathrm{Im}(\lambda)t)$, $\mathrm{Im}(f) = e^{\mathrm{Re}(\lambda)t}\sin(\mathrm{Im}(\lambda)t)$, og $f'(t) = \lambda e^{\lambda t}$.

Bevis. Lad os skrive $\lambda = \lambda_1 + i\lambda_2$ på rektangulær form. Da har vi for ethvert $t \in \mathbb{R}$

$$\begin{aligned} e^{\lambda t} &= e^{\lambda_1 t + i \cdot \lambda_2 t} \\ &= e^{\lambda_1 t} \cdot e^{i \cdot \lambda_2 t} \\ &= e^{\lambda_1 t} \cdot (\cos(\lambda_2 t) + i \cdot \sin(\lambda_2 t)) \\ &= e^{\lambda_1 t} \cos(\lambda_2 t) + i \cdot e^{\lambda_1 t} \sin(\lambda_2 t)). \end{aligned}$$

Dette viser, at realdelen af udtrykket $f(t) = e^{\lambda t}$ er givet ved $\mathrm{Re}(f(t)) = e^{\lambda_1 t} \cos(\lambda_2 t)$, mens imaginærdelen er givet ved $\mathrm{Im}(f(t)) = e^{\lambda_1 t} \sin(\lambda_2 t)$. Nu sætter vi $f'(t) = (\mathrm{Re}(f(t)))' + i \cdot (\mathrm{Im}(f(t)))'$. Ved at benytte produkt- og kædereglen til at beregne $\mathrm{Re}(f(t))'$ og $\mathrm{Im}(f(t))'$, får vi

$$\begin{aligned} f'(t) &= \mathrm{Re}(f(t))' + i \cdot \mathrm{Im}(f(t))' \\ &= (e^{\lambda_1 t} \cos(\lambda_2 t))' + i \cdot (e^{\lambda_1 t} \sin(\lambda_2 t))' \\ &= (e^{\lambda_1 t} \lambda_1 \cos(\lambda_2 t) + e^{\lambda_1 t}(-\sin(\lambda_2 t))\lambda_2) + i \cdot (e^{\lambda_1 t} \lambda_1 \sin(\lambda_2 t) + e^{\lambda_1 t} \cos(\lambda_2 t)\lambda_2) \\ &= (\lambda_1 + i\lambda_2) e^{\lambda_1 t} \cos(\lambda_2 t) + (-\lambda_2 + i\lambda_1) e^{\lambda_1 t} \sin(\lambda_2 t) \\ &= (\lambda_1 + i\lambda_2) e^{\lambda_1 t} (\cos(\lambda_2 t) + i \sin(\lambda_2 t)) \\ &= (\lambda_1 + i\lambda_2) e^{\lambda_1 t} e^{i\lambda_2 t} \\ &= \lambda e^{\lambda t}. \end{aligned}$$

□

Denne sætning vil vise sig særdeles nyttig, når vi senere skal finde løsninger til visse typer ODE'er.

12.2 Systemer af lineære førsteordens ODE'er med konstante koefficienter

I det forrige afsnit arbejdede vi med lineære førsteordens ODE'er. Nu tager vi fat på systemer af sådanne ODE'er, men vi vil kun betragte scenarier, hvor alle de funktioner, der optræder som koefficienter, er konstante. Vi vil i næste afsnit vise, at visse højereordens ODE'er kan løses ved hjælp af teorien fra dette afsnit.

Definition 12.2.1

Lad $n > 0$ være et heltal, $q_1(t), \ldots, q_n(t)$ reelle differentiable funktioner, og $\mathbf{A} \in \mathbb{R}^{n \times n}$ en matrix. Da er et system af lineære førsteordens ODE'er en ligning på formen

$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \\ \vdots \\ f_n'(t) \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix} + \begin{bmatrix} q_1(t) \\ q_2(t) \\ \vdots \\ q_n(t) \end{bmatrix} \tag{12.2}$$

Matricen $\mathbf{A}$ kaldes *koefficientmatricen* for systemet, mens funktionerne $q_1(t), \ldots, q_n(t)$ på engelsk kaldes systemets *forcing functions*. Hvis alle disse forcing functions $q_1(t), \ldots, q_n(t)$ er lig med nulfunktionen, kaldes systemet af ODE'er *homogent*, og ellers kaldes det *inhomogent*. En løsning til et inhomogent system af lineære førsteordens ODE'er kaldes en *partikulær løsning*.

Eksempel 12.2.1

Følgende system af lineære førsteordens ODE'er er givet:

$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix} + \begin{bmatrix} e^t \\ 0 \end{bmatrix}. \tag{12.3}$$

(a) Er det givne system af ODE'er i Ligning (12.3) homogent eller inhomogent?

(b) Er $(f_1(t), f_2(t)) = (e^{2t}, 0)$ en løsning til Ligning (12.3)?

(c) Er $(f_1(t), f_2(t)) = (-e^t, 0)$ en løsning til Ligning (12.3)?

Svar:

(a) Systemet af ODE'er i Ligning (12.3) er inhomogent. Dette konkluderer vi hurtigt ved at se på systemets forcing functions $q_2(t)$ og $q_1(t)$. Selvom funktionen $q_2(t)$ er nulfunktionen, er funktionen $q_1(t)$ det ikke. For et homogent system skal alle dets forcing functions være nulfunktionen.

(b) Hvis $(f_1(t), f_2(t)) = (e^{2t}, 0)$, så er

$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \end{bmatrix} = \begin{bmatrix} (e^{2t})' \\ 0 \end{bmatrix} = \begin{bmatrix} 2e^{2t} \\ 0 \end{bmatrix},$$

og

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix} + \begin{bmatrix} e^t \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \cdot e^{2t} + 1 \cdot 0 \\ 0 \cdot e^{2t} + 2 \cdot 0 \end{bmatrix} + \begin{bmatrix} e^t \\ 0 \end{bmatrix} = \begin{bmatrix} 2e^{2t} + e^t \\ 0 \end{bmatrix}.$$

Derfor er $(f_1(t), f_2(t)) = (e^{2t}, 0)$ ikke en løsning til Ligning (12.3).

(c) Hvis $(f_1(t), f_2(t)) = (-e^t, 0)$, da er

$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \end{bmatrix} = \begin{bmatrix} (-e^t)' \\ 0 \end{bmatrix} = \begin{bmatrix} -e^t \\ 0 \end{bmatrix},$$

og

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix} + \begin{bmatrix} e^t \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \cdot (-e^t) + 1 \cdot 0 \\ 0 \cdot (-e^t) + 2 \cdot 0 \end{bmatrix} + \begin{bmatrix} e^t \\ 0 \end{bmatrix} = \begin{bmatrix} -e^t \\ 0 \end{bmatrix}.$$

Derfor er $(f_1(t), f_2(t)) = (-e^t, 0)$ en løsning til Ligning (12.3). Jævnfør definitionen kan vi derfor kalde den en partikulær løsning til differentialligningssystemet.

Vi vil nu beskrive strukturen af løsningerne til systemer af lineære førsteordens ODE'er, hvilket gøres lidt på samme måde, som vi gjorde for systemer af lineære ligninger.

Sætning 12.2.1

Lad et inhomogent system af ODE'er som i Ligning (12.2) være givet, og antag, at $(g_1(t), g_2(t), \ldots, g_n(t))$ er en partikulær løsning til dette system. Da er enhver anden løsning $(\tilde{g}_1(t), \tilde{g}_2(t), \ldots, \tilde{g}_n(t))$ til Ligning (12.2) på formen

$$\begin{bmatrix} \tilde{g}_1(t) \\ \tilde{g}_2(t) \\ \vdots \\ \tilde{g}_n(t) \end{bmatrix} = \begin{bmatrix} g_1(t) \\ g_2(t) \\ \vdots \\ g_n(t) \end{bmatrix} + \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix},$$

hvor $(f_1(t), f_2(t), \ldots, f_n(t))$ er en løsning til det homogene system af ODE'er, der svarer til Ligning (12.2):

$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \\ \vdots \\ f_n'(t) \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix}. \tag{12.4}$$

Bevis. Antag, at $(\tilde{g}_1(t), \tilde{g}_2(t), \ldots, \tilde{g}_n(t))$ er en vilkårlig løsning til Ligning (12.2). Da viser en direkte beregning, at $(\tilde{g}_1(t) - g_1(t), \tilde{g}_2(t) - g_2(t), \ldots, \tilde{g}_n(t) - g_n(t))$ opfylder Ligning (12.4). Hvis vi derefter definerer $f_i(t) = \tilde{g}_i(t) - g_i(t)$ for $i = 1, \ldots, n$, ser vi, at $(\tilde{g}_1(t), \tilde{g}_2(t), \ldots, \tilde{g}_n(t))$ kan skrives som angivet i sætningen.

Omvendt, hvis $(f_1(t), f_2(t), \ldots, f_n(t))$ er en løsning til det homogene system fra Ligning (12.4), viser en direkte beregning, at $(g_1(t) + f_1(t), g_2(t) + f_2(t), \ldots, g_n(t) + f_n(t))$ er en løsning til det inhomogene system fra Ligning (12.2). □

Algoritmisk betyder dette, at vi for at løse et inhomogent system af ODE'er som i Ligning (12.2) skal finde en partikulær løsning dertil og derefter alle løsninger til det tilsvarende homogene system af ODE'er givet i Ligning (12.4). Konceptuelt kan man forstå Sætning 12.2.1 på en anden måde. Lad $C_\infty(\mathbb{R})$ være vektorrummet fra Eksempel 9.3.4. Det består af alle funktioner med definitionsmængde og dispositionsmængde $\mathbb{R}$, der kan differentieres vilkårligt ofte. Betragt for en givet matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ afbildningen $L_\mathbf{A} : C_\infty(\mathbb{R})^n \to C_\infty(\mathbb{R})^n$ defineret ved

$$L_\mathbf{A}\left(\begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix}\right) = \begin{bmatrix} f_1'(t) \\ f_2'(t) \\ \vdots \\ f_n'(t) \end{bmatrix} - \mathbf{A} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix}. \tag{12.5}$$

Man kan vise, at $L_\mathbf{A}$ er en lineær afbildning mellem reelle vektorrum. Kernen af denne afbildning er netop løsningsmængden til det homogene system af ODE'er i Ligning (12.4). Denne observation er en generalisering af det, vi allerede har set i Eksempel 10.2.8. En partikulær løsning er blot en vektor $\mathbf{v}_p = (g_1(t), g_2(t), \ldots, g_n(t)) \in C_\infty(\mathbb{R})^n$, der opfylder, at $L_\mathbf{A}(\mathbf{v}_p) = (q_1(t), \ldots, q_n(t))$. Derfor er Sætning 12.2.1 blot et særtilfælde af andet punkt i Sætning 10.4.1. Bemærk, da kernen af enhver lineær afbildning er et underrum, at løsningsmængden til et homogent system af lineære førsteordens ODE'er (med konstante koefficienter) faktisk er et vektorrum over de reelle tal, da det er kernen af den lineære afbildning $L_\mathbf{A}$. Et meget nyttigt faktum, som vi ikke vil bevise her, er, at dette vektorrum har endelig dimension, som vi kan kalde n. Det er nyttigt at vide, da det betyder, at for at beskrive alle løsninger til systemet i Ligning (12.4) er det nok at finde en basis, altså n lineært uafhængige løsninger. Vi vil benytte dette frit senere. Hvad vi primært vil fokusere på i resten af dette afsnit, er, hvordan man finder en sådan basis. Begrebet fuldstændig løsning, som vi er stødt på igen i Afsnit 12.1 for lineære førsteordens ODE'er, kan nu generaliseres som følger:

Definition 12.2.2

Lad $\mathbf{A} \in \mathbb{R}^{n\times n}$ være givet. Den *fuldstændige løsning* til de homogene ODE'er

$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \\ \vdots \\ f_n'(t) \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix}$$

er et udtryk på formen

$$c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n, \quad c_1, \ldots, c_n \in \mathbb{R},$$

hvor $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ er en ordnet basis for kernen af den lineære afbildning $L_\mathbf{A} : C_\infty(\mathbb{R})^n \to C_\infty(\mathbb{R})^n$ defineret i Ligning (12.5). Hvis $q_1(t), \ldots, q_n(t)$ er forcing functions (der ikke alle er nulfunktionen), og $\mathbf{v}_p = (g_1(t), \ldots, g_n(t)) \in C_\infty(\mathbb{R})^n$ er en partikulær løsning til det inhomogene system af ODE'er

$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \\ \vdots \\ f_n'(t) \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix} + \begin{bmatrix} q_1(t) \\ q_2(t) \\ \vdots \\ q_n(t) \end{bmatrix},$$

så er den fuldstændige løsning til det inhomogene system et udtryk på formen

$$\mathbf{v}_p + c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n\,, \quad c_1, \ldots, c_n \in \mathbb{R}.$$

Et første vigtigt trick herefter er at anvende teorien om egenværdier og egenvektorer for matricen **A**, hvilket vi ser i følgende lemma.

Lemma 12.2.2

Lad $\mathbf{A} \in \mathbb{R}^{n\times n}$ være en matrix, og antag, at $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$ er en egenvektor for **A** med egenværdi $\lambda \in \mathbb{R}$. Da opfylder vektoren af funktioner

$$\begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix} = \begin{bmatrix} v_1 e^{\lambda t} \\ v_2 e^{\lambda t} \\ \vdots \\ v_n e^{\lambda t} \end{bmatrix}$$

det homogene system af ODE'er

$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \\ \vdots \\ f_n'(t) \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix}.$$

Bevis. På den ene side har vi

$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \\ \vdots \\ f_n'(t) \end{bmatrix} = \begin{bmatrix} v_1 (e^{\lambda t})' \\ v_2 (e^{\lambda t})' \\ \vdots \\ v_n (e^{\lambda t})' \end{bmatrix} = \begin{bmatrix} v_1 \lambda e^{\lambda t} \\ v_2 \lambda e^{\lambda t} \\ \vdots \\ v_n \lambda e^{\lambda t} \end{bmatrix} = \lambda \begin{bmatrix} v_1 e^{\lambda t} \\ v_2 e^{\lambda t} \\ \vdots \\ v_n e^{\lambda t} \end{bmatrix} = \lambda \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix}.$$

På den anden side har vi

$$\mathbf{A} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} v_1 e^{\lambda t} \\ v_2 e^{\lambda t} \\ \vdots \\ v_n e^{\lambda t} \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \cdot e^{\lambda t} = \lambda \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \cdot e^{\lambda t} = \lambda \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix}.$$

□

Eksempel 12.2.2
Lad
$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$
Find en løsning til det homogene system af lineære førsteordens ODE'er med koefficientmatrix $\mathbf{A}$.

Svar:
Vi bliver bedt om at finde en løsning til det følgende system af ODE'er:
$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix}. \tag{12.6}$$
Med Lemma 12.2.2 i tankerne lægger vi ud med at finde en egenværdi og tilhørende egenvektor for den givne matrix $\mathbf{A}$. Det karakteristiske polynomium for $\mathbf{A}$ er:
$$p_{\mathbf{A}}(Z) = \det(\mathbf{A} - \lambda \mathbf{I}_2) = \det\left(\begin{bmatrix} 2-\lambda & 1 \\ 0 & 2-\lambda \end{bmatrix}\right) = (2-\lambda)^2 = (\lambda-2)^2.$$
Dermed er 2 den eneste egenværdi, matricen $\mathbf{A}$ har. For at finde en egenvektor for $\mathbf{A}$ tilhørende egenværdien 2 skal vi finde en vektor forskellig fra nulvektoren, der tilhører kernen af matricen $\mathbf{A} - 2\mathbf{I}_2$. Vi bestemmer kernen ved at reducere $\mathbf{A} - 2\mathbf{I}_2$ til sin reducerede trappeform, men det viser sig i dette særtilfælde, at den allerede er på reduceret trappeform:
$$\mathbf{A} - 2\mathbf{I}_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$
Vi konkluderer med det samme, at $\ker(\mathbf{A} - 2\mathbf{I}_2)$ er et endimensionelt vektorrum med en basis givet ved for eksempel
$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}.$$
Nu medfører Lemma 12.2.2, at
$$\begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix} = \begin{bmatrix} 1e^{2t} \\ 0e^{2t} \end{bmatrix} = \begin{bmatrix} e^{2t} \\ 0 \end{bmatrix}$$
er en løsning til Ligning (12.6).

Det er tilstrækkeligt at benytte egenvektorer som i Lemma 12.2.2 til at bestemme den fuldstændige løsning til Ligning (12.4), når matricen $\mathbf{A}$ kan diagonaliseres. Husk på, at dette svarer til at sige, at der eksisterer en invertibel matrix $\mathbf{Q}$, således at $\mathbf{Q}^{-1} \cdot \mathbf{A} \cdot \mathbf{Q}$ er en diagonalmatrix. Mere præcist så vi i forrige kapitel, at søjlerne i matricen $\mathbf{Q}$ er lineært uafhængige egenvektorer for matricen $\mathbf{A}$. Hvis disse søjler tilhører egenværdier $\lambda_1, \ldots, \lambda_n$, så har vi faktisk $\mathbf{Q}^{-1} \cdot \mathbf{A} \cdot \mathbf{Q} = \Lambda$, hvor Λ er diagonalmatricen med egenværdierne $\lambda_1, \ldots, \lambda_n$ i sin diagonal.

Sætning 12.2.3

Antag, at $\mathbf{A} \in \mathbb{R}^{n \times n}$ er en diagonalisérbar matrix. Mere præcist, lad $\mathbf{Q}$ være en invertibel matrix, således at $\mathbf{Q}^{-1} \cdot \mathbf{A} \cdot \mathbf{Q}$ er diagonalmatricen med egenværdierne $\lambda_1, \ldots, \lambda_n$ i sin diagonal. Så har det homogene system af ODE'er i Ligning (12.4) den fuldstændige løsning

$$\begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix} = \mathbf{Q} \cdot \begin{bmatrix} c_1 \cdot e^{\lambda_1 t} \\ c_2 \cdot e^{\lambda_2 t} \\ \vdots \\ c_n \cdot e^{\lambda_n t} \end{bmatrix}, \quad c_1, \ldots, c_n \in \mathbb{R}.$$

Bevis. Bemærk først, at systemet i Ligning (12.4) er ækvivalent med systemet

$$\mathbf{Q}^{-1} \cdot \begin{bmatrix} f_1'(t) \\ f_2'(t) \\ \vdots \\ f_n'(t) \end{bmatrix} = \mathbf{Q}^{-1} \cdot \mathbf{A} \cdot \mathbf{Q} \cdot \mathbf{Q}^{-1} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix}.$$

Definerer vi

$$\begin{bmatrix} \tilde{f}_1(t) \\ \tilde{f}_2(t) \\ \vdots \\ \tilde{f}_n(t) \end{bmatrix} = \mathbf{Q}^{-1} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix}, \tag{12.7}$$

ser vi, at $(f_1(t), \ldots, f_n(t))$ er en løsning til systemet i Ligning (12.4), hvis og kun hvis

$$\begin{bmatrix} \tilde{f}_1'(t) \\ \tilde{f}_2'(t) \\ \vdots \\ \tilde{f}_n'(t) \end{bmatrix} = \mathbf{Q}^{-1} \cdot \mathbf{A} \cdot \mathbf{Q} \cdot \begin{bmatrix} \tilde{f}_1(t) \\ \tilde{f}_2(t) \\ \vdots \\ \tilde{f}_n(t) \end{bmatrix}. \tag{12.8}$$

Men da matricen $\mathbf{Q}^{-1} \cdot \mathbf{A} \cdot \mathbf{Q}$ er en diagonalmatrix med diagonalelementerne $\lambda_1, \ldots, \lambda_n$, reduceres systemet i Ligning (12.8) blot til systemet af differentialligningerne givet ved $\tilde{f}_1'(t) = \lambda_1 \cdot \tilde{f}_1(t), \tilde{f}_2'(t) = \lambda_2 \cdot \tilde{f}_2(t), \ldots, \tilde{f}_n'(t) = \lambda_n \cdot \tilde{f}_n(t)$. Hver af disse differentialligninger er en homogen lineær ODE af første orden og kan derfor løses individuelt ved hjælp af Sætning 12.1.1. Resultatet er, at $\tilde{f}_i(t) = c_i \cdot e^{\lambda_i t}$ for $i = 1, \ldots, n$ og $c_i \in \mathbb{R}$. Med andre ord

$$\begin{bmatrix} \tilde{f}_1(t) \\ \tilde{f}_2(t) \\ \vdots \\ \tilde{f}_n(t) \end{bmatrix} = \begin{bmatrix} c_1 \cdot e^{\lambda_1 t} \\ c_2 \cdot e^{\lambda_2 t} \\ \vdots \\ c_n \cdot e^{\lambda_n t} \end{bmatrix}.$$

Men så følger udsagnet i sætningen ved brug af Ligning (12.7). □

En direkte konsekvens af sætningen er følgende alternative beskrivelse af den fuldstændige løsning.

Korollar 12.2.4

Antag, at $\mathbf{A} \in \mathbb{R}^{n \times n}$ er en diagonalisérbar matrix. Mere præcist, lad $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ være en ordnet basis for $\mathbb{R}^n$ bestående af egenvektorer for $\mathbf{A}$, der svarer til egenværdierne $\lambda_1, \ldots, \lambda_n$. Så har det homogene system i Ligning (12.4) den fuldstændige løsning

$$c_1 \cdot \mathbf{v}_1 e^{\lambda_1 t} + \cdots + c_n \cdot \mathbf{v}_n e^{\lambda_n t}, \quad c_1, \ldots, c_n \in \mathbb{R}.$$

Bevis. Lad $\mathbf{Q}$ være matricen med søjler bestående af egenvektorerne $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Da er $\mathbf{Q}$ en invertibel matrix, der opfylder, at $\mathbf{Q}^{-1} \cdot \mathbf{A} \cdot \mathbf{Q}$ er en diagonalmatrix med egenværdierne $\lambda_1, \ldots, \lambda_n$ i sin diagonal. Ved at anvende Sætning 12.2.3 følger korollaret. □

Bemærk, at Sætning 12.2.3 og Korollar 12.2.4 tillader, at egenværdier kan optræde flere gange. Med andre ord: vi tillader tilfælde, hvor den algebraiske multiplicitet af egenværdierne er større end én. Da vi også antager, at matricen $\mathbf{A}$ er diagonalisérbar, skal den algebraiske og geometriske multiplicitet af en egenværdi dog være ens. Derfor vil sætningen ikke være anvendelig, hvis nogle egenværdier for $\mathbf{A}$ har en mindre geometrisk multiplicitet end algebraisk multiplicitet.

Eksempel 12.2.3

Lad

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Så er $p_{\mathbf{A}}(Z) = (Z-2)^2 \cdot (Z^2 - 1) = (Z-2)^2 \cdot (Z-1) \cdot (Z+1)$. Derfor har $\mathbf{A}$ de tre egenværdier 2, 1 og -1 med algebraiske multipliciteter henholdsvis 2, 1 og 1. Man kan vise, at baser for egenrummene E_2, E_1 og E_{-1} kan gives ved

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}, \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\} \text{ og } \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} \right\}.$$

Bemærk, at den geometriske og algebraiske multiplicitet er den samme for hver egenværdi. Ved at benytte Korollar 12.2.4 ser vi, at den fuldstændige løsning til systemet af lineære førsteordens ODE'er

$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \\ f_3'(t) \\ f_4'(t) \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \\ f_3(t) \\ f_4(t) \end{bmatrix}$$

er

$$\begin{bmatrix} f_1(t) \\ f_2(t) \\ f_3(t) \\ f_4(t) \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} e^{2t} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} e^{2t} + c_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} e^{t} + c_4 \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} e^{-t} = \begin{bmatrix} c_1 e^{2t} \\ c_2 e^{2t} \\ c_3 e^{t} + c_4 e^{-t} \\ c_3 e^{t} - c_4 e^{-t} \end{bmatrix},$$

hvor $c_1, c_2, c_3, c_4 \in \mathbb{R}$.

Før vi går videre til tilfælde, hvor matricen $\mathbf{A}$ ikke kan diagonaliseres, bør vi nævne, at man også for systemer af differentialligninger som dem, vi har undersøgt her, kan opstille begyndelsesværdibetingelser svarende til dem, vi definerede i Definition 12.1.1.

Definition 12.2.3

Givet reelle funktioner $f_1(t), \ldots, f_n(t)$, et reelt tal t_0 og reelle tal $y_1, \ldots, y_n$, således at $f_i(t_0) = y_i$ for $i = 1, \ldots, n$. Da siges funktionerne $f_1(t), \ldots, f_n(t)$ at opfylde *begyndelsesværdibetingelserne* $f_i(t_0) = y_i$ for $i = 1, \ldots, n$.

Det viser sig, at systemet i Ligning (12.4) har netop én løsning, der opfylder begyndelsesværdibetingelserne som angivet i Definition 12.2.3. At bestemme denne løsning kan gøres ved først at bestemme den fuldstændige løsning til systemet i Ligning (12.4) og derefter at bestemme værdierne af konstanterne $c_1, \ldots, c_n$, der forekommer i den fuldstændige løsning, således at begyndelsesværdibetingelserne opfyldes. Vi giver et lille eksempel.

Eksempel 12.2.4

Lad os betragte det samme system af differentialligninger som i Eksempel 12.2.3. Der så vi, at den fuldstændige løsning var givet ved

$$\begin{bmatrix} f_1(t) \\ f_2(t) \\ f_3(t) \\ f_4(t) \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} e^{2t} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} e^{2t} + c_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} e^{t} + c_4 \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} e^{-t} = \begin{bmatrix} c_1 e^{2t} \\ c_2 e^{2t} \\ c_3 e^{t} + c_4 e^{-t} \\ c_3 e^{t} - c_4 e^{-t} \end{bmatrix},$$

hvor $c_1, c_2, c_3, c_4 \in \mathbb{R}$.

Spørgsmål: Bestem den løsning, der opfylder begyndelsesværdibetingelserne $f_1(0) = 1$, $f_2(0) = 2$, $f_3(0) = 3$ og $f_4(0) = 4$.

Svar: Ved at sætte $t = 0$ i den fuldstændige løsning og benytte de givne begyndelsesværdibetingelser, får vi, at konstanterne c_1, c_2, c_3 og c_4 skal opfylde

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 + c_4 \\ c_3 - c_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}.$$

Dette medfører, at

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 + c_4 \\ c_3 - c_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 7/2 \\ -1/2 \end{bmatrix}.$$

Derfor er den løsning, vi leder efter, følgende:

$$\begin{bmatrix} f_1(t) \\ f_2(t) \\ f_3(t) \\ f_4(t) \end{bmatrix} = \begin{bmatrix} e^{2t} \\ 2e^{2t} \\ (7e^t - e^{-t})/2 \\ (7e^t + e^{-t})/2 \end{bmatrix}.$$

Vi vender nu tilbage til studiet af systemet i Ligning (12.4). Kravet i Sætning 12.2.3 om, at der eksisterer en basis af egenvektorer, kan risikere ikke at være opfyldt. Det kan for eksempel ske, hvis det karakteristisk polynomium $p_{\mathbf{A}}(Z)$ ikke kan skrives som et produkt af polynomier af grad én. Med andre ord: $p_{\mathbf{A}}(Z)$ kan have komplekse, ikke-reelle rødder. Den følgende sætning udvider Sætning 12.2.3, så sådanne scenarier er dækket ind.

Sætning 12.2.5

Lad $\mathbf{A} \in \mathbb{C}^{n \times n}$ være en matrix, og lad $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ være en ordnet basis for $\mathbb{C}^n$ bestående af egenvektorer for $\mathbf{A}$, der tilhører (muligvis ikke-reelle) egenværdier $\lambda_1, \ldots, \lambda_n$. Da har det homogene system af ODE'er i Ligning (12.4) over de komplekse tal den fuldstændige løsning

$$c_1 \cdot \mathbf{v}_1 e^{\lambda_1 t} + \cdots + c_n \cdot \mathbf{v}_n e^{\lambda_n t}, \quad c_1, \ldots, c_n \in \mathbb{C}.$$

Bevis. Vi udelader detaljerne i beviset og nævner blot, at beviset praktisk talt er identisk med beviserne for Sætning 12.2.3 og Korollar 12.2.4. Den eneste forskel er, at vi nu arbejder over de komplekse tal. Bemærk, at Lemma 12.1.3 garanterer, at $(e^{\lambda t})' = \lambda e^{\lambda t}$ også for $\lambda \in \mathbb{C}$. $\square$

Antag nu, at $\mathbf{A} \in \mathbb{R}^{n \times n}$, men at dens karakteristiske polynomium $p_{\mathbf{A}}(Z)$ har komplekse rødder. Vi kunne betragte $\mathbf{A}$ som en matrix i $\mathbb{C}^{n \times n}$ og anvende Sætning 12.2.5 til at bestemme den fuldstændige løsning. Problemet med dette vil dog være, at vi så finder den fuldstændige løsning til Ligning 12.4 bestående af komplekse funktioner med reelt input. Man er ofte mere interesseret i en fuldstændig løsning af reelle funktioner i stedet. Heldigvis kan dette opnås med nogle få tricks. Det vigtigste trick er, at da $p_{\mathbf{A}}(Z)$ har koefficienter i $\mathbb{R}$, hvis $\mathbf{A} \in \mathbb{R}^{n \times n}$, så forekommer alle ikke-reelle rødder i par: hvis $\mu \in \mathbb{C} \setminus \mathbb{R}$ er en rod, så er også $\overline{\mu} \in \mathbb{C}$ en rod, hvor $\overline{\mu}$ betegner den komplekst konjugerede af μ (se Lemma 4.3.3). Specifikt kan de reelle rødder i $p_{\mathbf{A}}(Z)$ arrangeres på formen $\lambda_1, \ldots, \lambda_r$ og de komplekse, ikke-reelle rødder på formen $\mu_1, \ldots, \mu_s, \overline{\mu}_1, \ldots, \overline{\mu}_s$. Derfor har vi $n = r + 2s$, hvor vi blot gentager en rod m gange, hvis den forekommer med multiplicitet m. Lad os illustrere dette med et eksempel.

Eksempel 12.2.5

Antag, at $p_{\mathbf{A}}(Z) = (Z-1) \cdot (Z-2)^3 \cdot (Z^2+1)^2$ for en matrix $\mathbf{A} \in \mathbb{R}^{7\times 7}$. Da er rødderne i dette polynomium 1, 2 med multiplicitet 3 og i samt $-i$, begge med multiplicitet 2. Der er to reelle rødder, nemlig 1 og 2, men hvis vi betragter disse rødder med deres multiplicitet, skal vi gentage roden 2 tre gange. Derfor er $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 2$ og $\lambda_4 = 2$. Der er to komplekse, ikke-reelle rødder i og $-i$, som begge skal gentages to gange. Derfor har vi $\mu_1 = i$ og $\mu_2 = i$ samt deraf også $\overline{\mu}_1 = -i$ og $\overline{\mu}_2 = -i$. I dette tilfælde har vi altså $r = 3$ og $s = 2$.

For at kunne beskrive den fuldstændige løsning til Ligning (12.4) i tilfælde af ikke-reelle rødder i $p_{\mathbf{A}}(Z)$, vil det være praktisk at definere den komplekst konjugerede af en vektor $\mathbf{w} \in \mathbb{C}^n$: hvis $\mathbf{w} = (w_1, \ldots, w_n)$, så er $\overline{\mathbf{w}} = (\overline{w}_1, \ldots, \overline{w}_n)$. Pointen her er, at hvis $\mathbf{A} \in \mathbb{R}^{n\times n}$, og $\mathbf{A} \cdot \mathbf{w} = \mu \cdot \mathbf{w}$ for nogle $\mathbf{w} \in \mathbb{C}^n$ og $\mu \in \mathbb{C} \setminus \mathbb{R}$, så ser vi, når vi tager den komplekst konjugerede (og udnytter, at elementerne i $\mathbf{A}$ er reelle tal), at $\mathbf{A} \cdot \overline{\mathbf{w}} = \overline{\mu} \cdot \overline{\mathbf{w}}$. Med dette in mente medfører Sætning 12.2.5 følgende:

Korollar 12.2.6

Antag, at $\mathbf{A} \in \mathbb{R}^{n\times n}$, og at rødderne i dets karakteristiske polynomium $p_{\mathbf{A}}(Z)$ er arrangeret med multiplicitet som $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ og $\mu_1, \ldots, \mu_s, \overline{\mu}_1, \ldots, \overline{\mu}_s$, hvor $\mu_1, \ldots, \mu_s \in \mathbb{C} \setminus \mathbb{R}$. Antag yderligere, at der eksisterer vektorer $\mathbf{v}_i \in \mathbb{R}^n$ for $i = 1, \ldots, s$ og $\mathbf{w}_j \in \mathbb{C}^n$ for $j = 1, \ldots, s$, således at følgende er opfyldt:

(i) $\mathbf{A} \cdot \mathbf{v}_i = \lambda_i \cdot \mathbf{v}_i$ for $i = 1, \ldots, r$,

(ii) $\mathbf{A} \cdot \mathbf{w}_j = \mu_j \cdot \mathbf{w}_j$ for $j = 1, \ldots, s$,

(iii) vektorerne $\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{w}_1, \ldots, \mathbf{w}_s, \overline{\mathbf{w}}_1, \ldots, \overline{\mathbf{w}}_s$ danner en ordnet basis for $\mathbb{C}^n$.

Så har det homogene system af ODE'er i Ligning (12.4) den fuldstændige løsning

$$c_1 \cdot \mathbf{v}_1 e^{\lambda_1 t} + \cdots + c_r \cdot \mathbf{v}_r e^{\lambda_r t} + c_{r+1} \cdot \mathrm{Re}(\mathbf{w}_1 e^{\mu_1 t}) + \cdots + c_{r+s} \cdot \mathrm{Re}(\mathbf{w}_s e^{\mu_s t}) +$$
$$c_{r+s+1} \cdot \mathrm{Im}(\mathbf{w}_1 e^{\mu_1 t}) + \cdots + c_n \cdot \mathrm{Im}(\mathbf{w}_s e^{\mu_s t}), \quad c_1, \ldots, c_n \in \mathbb{R}.$$

Bevis. Når matricen $\mathbf{A}$ opfattes som en matrix over $\mathbb{C}$, er egenværdierne af $\mathbf{A}$ givet ved

$$\lambda_1, \ldots, \lambda_r, \mu_1, \ldots, \mu_s, \overline{\mu}_1, \ldots, \overline{\mu}_s.$$

Derfor medfører Sætning 12.2.5, at

$$\mathbf{v}_1 e^{\lambda_1 t}, \ldots, \mathbf{v}_r e^{\lambda_r t}, \mathbf{w}_1 e^{\mu_1 t}, \ldots, \mathbf{w}_s e^{\mu_s t}, \overline{\mathbf{w}}_1 e^{\overline{\mu}_1 t}, \ldots, \overline{\mathbf{w}}_s e^{\overline{\mu}_s t}$$

danner en basis for løsningerne til Ligning (12.4), når man arbejder over $\mathbb{C}$. For at opnå en basis for denne mængde løsninger, når man arbejder over $\mathbb{R}$, justerer vi basen. Først og fremmest er løsningerne $\mathbf{v}_1e^{\lambda_1 t}, \dots, \mathbf{v}_r e^{\lambda_r t}$ allerede reelle funktioner, så ingen ændring er nødvendig for disse. Givet et par af komplekse funktioner med reelt input som løsninger, $\mathbf{w}_j e^{\mu_j t}$ og $\overline{\mathbf{w}}_j e^{\overline{\mu}_j t}$ for et j, kan vi erstatte dette par med parret

$$\frac{\mathbf{w}_j e^{\mu_j t} + \overline{\mathbf{w}}_j e^{\overline{\mu}_j t}}{2} = \mathrm{Re}(\mathbf{w}_j e^{\mu_j t}) \quad \text{og} \quad \frac{\mathbf{w}_j e^{\mu_j t} - \overline{\mathbf{w}}_j e^{\overline{\mu}_j t}}{2i} = \mathrm{Im}(\mathbf{w}_j e^{\mu_j t}).$$

Da $\mathrm{Re}(\mathbf{w}_j e^{\mu_j t})$ og $\mathrm{Im}(\mathbf{w}_j e^{\mu_j t})$ beskriver reelle funktioner, får vi derfor en basis for alle reelle funktioner, som er løsninger til Ligning (12.4), fra de n løsninger

$$\mathbf{v}_1e^{\lambda_1 t}, \dots, \mathbf{v}_r e^{\lambda_r t}, \mathrm{Re}(\mathbf{w}_1 e^{\mu_1 t}), \dots, \mathrm{Re}(\mathbf{w}_s e^{\mu_s t}), \mathrm{Im}(\mathbf{w}_1 e^{\mu_1 t}), \dots, \mathrm{Im}(\mathbf{w}_s e^{\mu_s t}).$$

□

Det første punkt i Korollar 12.2.6 betyder blot, at vektoren $\mathbf{v}_i$ er en egenvektor for $\mathbf{A}$ med egenværdi λ_i. Det andet punkt betyder, at hvis vi arbejdede over det komplekse tallegeme $\mathbb{C}$ i stedet for $\mathbb{R}$, så ville $\mathbf{w}_j$ være en egenvektor med egenværdi μ_j. I så fald kan det vises, at $\overline{\mathbf{w}}_j$ også er en egenvektor for $\mathbf{A}$ med egenværdi $\overline{\mu}_j$. Endelig betyder det tredje punkt, at der findes en basis for $\mathbb{C}^n$, der består af egenvektorer for $\mathbf{A}$, når man betragter $\mathbf{A}$ som en matrix i $\mathbb{C}^{n\times n}$. Derfor kan de tre punkter samlet omformuleres som: når man betragter $\mathbf{A}$ som en matrix i $\mathbb{C}^{n\times n}$, er matricen $\mathbf{A}$ diagonalisérbar.

Korollar 12.2.6 kan umiddelbart virke kompliceret, men det er praktisk at have ved hånden i konkrete tilfælde. Lad os derfor se på et eksempel.

Eksempel 12.2.6

Lad

$$\mathbf{A} = \begin{bmatrix} 0 & 13 \\ -1 & 4 \end{bmatrix}.$$

Formålet med dette eksempel er at vise, hvordan man bestemmer den fuldstændige løsning til det homogene system af ODE'er

$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \end{bmatrix} = \begin{bmatrix} 0 & 13 \\ -1 & 4 \end{bmatrix} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix}. \tag{12.9}$$

For at være mere præcis ønsker vi at finde den fuldstændige løsning bestående af reelle funktioner.

Først udregner vi, at

$$p_{\mathbf{A}}(Z) = \det(\mathbf{A} - Z\mathbf{I}_2) = \det\left(\begin{bmatrix} -Z & 13 \\ -1 & 4-Z \end{bmatrix}\right) = (-Z)\cdot(4-Z) - 13\cdot(-1) = Z^2 - 4Z + 13.$$

Dette polynomium har rødderne $2+3i$ og $2-3i$ (se Sætning 4.2.1). Da rødderne ikke er reelle, arbejder vi over de komplekse tal for nu. Først finder vi en kompleks egenvektor tilhørende den ikke-reelle rod $2+3i$. Vi gør dette ved at bestemme den reducerede trappeform af matricen $\mathbf{A}-(2+3i)\mathbf{I}_2$:

$$\mathbf{A}-(2+3i)\mathbf{I}_2 = \begin{bmatrix} -2-3i & 13 \\ -1 & 2-3i \end{bmatrix} \underset{R_1 \leftrightarrow R_2}{\longrightarrow} \begin{bmatrix} -1 & 2-3i \\ -2-3i & 13 \end{bmatrix}$$
$$\underset{R_1 \leftarrow -R_1}{\longrightarrow} \begin{bmatrix} 1 & -2+3i \\ -2-3i & 13 \end{bmatrix}$$
$$\underset{R_2 \leftarrow R_2+(2+3i)R_1}{\longrightarrow} \begin{bmatrix} 1 & -2+3i \\ 0 & 0 \end{bmatrix}.$$

Nu ser vi, at E_{2+3i}, altså kernen af $\mathbf{A}-(2+3i)\mathbf{I}_2$ når den betragtes som en matrix i $\mathbb{C}^{2\times 2}$, er lig med $\{(v_1,v_2)\in\mathbb{C}^2 \mid v_1=(2-3i)v_2\}$. Derfor er en basis for E_{2+3i} for eksempel givet ved

$$\left\{\begin{bmatrix} 2-3i \\ 1 \end{bmatrix}\right\}.$$

Tilsvarende viser man, at en mulig basis for E_{2-3i} er

$$\left\{\begin{bmatrix} 2+3i \\ 1 \end{bmatrix}\right\},$$

men vi har faktisk ikke brug for denne anden basis. Nu følger vi opskriften beskrevet i Korollar 12.2.6, og vi udregner først

$$\begin{aligned}
\begin{bmatrix} 2-3i \\ 1 \end{bmatrix} e^{(2+3i)t} &= \begin{bmatrix} 2-3i \\ 1 \end{bmatrix} e^{2t}(\cos(3t)+i\sin(3t)) \\
&= \begin{bmatrix} (2-3i)e^{2t}(\cos(3t)+i\sin(3t)) \\ e^{2t}(\cos(3t)+i\sin(3t)) \end{bmatrix} \\
&= \begin{bmatrix} 2e^{2t}\cos(3t)+3e^{2t}\sin(3t)+i(2e^{2t}\sin(3t)-3e^{2t}\cos(3t)) \\ e^{2t}\cos(3t)+ie^{2t}\sin(3t) \end{bmatrix}
\end{aligned}$$

Derfor er

$$\mathrm{Re}\left(\begin{bmatrix} 2-3i \\ 1 \end{bmatrix} e^{(2+3i)t}\right) = \begin{bmatrix} 2e^{2t}\cos(3t)+3e^{2t}\sin(3t) \\ e^{2t}\cos(3t) \end{bmatrix}$$

og

$$\mathrm{Im}\left(\begin{bmatrix} 2-3i \\ 1 \end{bmatrix} e^{(2+3i)t}\right) = \begin{bmatrix} 2e^{2t}\sin(3t)-3e^{2t}\cos(3t) \\ e^{2t}\sin(3t) \end{bmatrix}.$$

Ved hjælp af Korollar 12.2.6 kan vi konkludere, at den fuldstændige løsning til systemet i Ligning (12.9) er givet ved

$$\begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix} = c_1 \cdot \begin{bmatrix} 2e^{2t}\cos(3t)+3e^{2t}\sin(3t) \\ e^{2t}\cos(3t) \end{bmatrix} + c_2 \cdot \begin{bmatrix} 2e^{2t}\sin(3t)-3e^{2t}\cos(3t) \\ e^{2t}\sin(3t) \end{bmatrix},$$

hvor $c_1, c_2 \in \mathbb{R}$.

Vi har nu beskrevet metoder til at bestemme den fuldstændige løsning, når matricen $\mathbf{A}$ er diagonalisérbar over $\mathbb{R}$ (Sætning 12.2.3), og når den er diagonalisérbar over $\mathbb{C}$ (Korollar 12.2.6). Hvis matricen ikke er diagonalisérbar, ikke engang over $\mathbb{C}$, kan det være noget sværere at løse; der findes en formel for den fuldstændige løsning i et sådant scenarie, men dette ligger udenfor rammerne af denne tekst. Vi vil dog vise et eksempel på et særtilfælde, der berører dette problem.

Eksempel 12.2.7

Lad

$$\mathbf{A} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \text{ med } \lambda \in \mathbb{R}.$$

Denne matrix har λ som egenværdi med algebraisk multiplicitet to og geometrisk multiplicitet ét. Derfor gælder Sætning 12.2.3 ikke, da E_λ er endimensionel med en basis udgjort af for eksempel vektoren $(1, 0)$.

Vi ønsker at bestemme den fuldstændige løsning til systemet af ODE'er

$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \end{bmatrix} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix}. \tag{12.10}$$

Skilles systemet ad, har vi at gøre med de to ODE'er $f_1'(t) = \lambda \cdot f_1(t) + f_2(t)$ og $f_2'(t) = \lambda \cdot f_2(t)$. En løsning findes ved at sætte $f_2(t) = 0$, altså nulfunktionen, og $f_1(t) = e^{\lambda t}$. Med andre ord: vektoren af funktioner $(e^{\lambda t}, 0)$ er en løsning til systemet i Ligning (12.10). En anden løsning kan findes ved at vælge $f_2(t) = e^{\lambda t}$. Da skal $f_1(t)$ opfylde den lineære inhomogene ODE $f_1'(t) = \lambda \cdot f_1(t) + e^{\lambda t}$. Ved at benytte Korollar 12.1.2, har vi, at $f_1(t) = e^{\lambda t} \int e^{-\lambda t} e^{\lambda t} dt = e^{\lambda t} t + c \cdot e^{\lambda t}$, hvor $c \in \mathbb{R}$. Ved at vælge $c = 0$, ser vi, at $(f_1(t), f_2(t)) = (te^{\lambda t}, e^{\lambda t})$ også er en løsning til systemet i Ligning (12.10). Da vi nu har fundet to lineært uafhængige løsninger, kan vi konkludere, at den fuldstændige løsning til systemet i Ligning (12.10) er givet ved

$$\begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix} = c_1 \cdot \begin{bmatrix} e^{\lambda t} \\ 0 \end{bmatrix} + c_2 \cdot \begin{bmatrix} te^{\lambda t} \\ e^{\lambda t} \end{bmatrix}, \quad c_1, c_2 \in \mathbb{R}.$$

12.3 Relationen mellem systemer af lineære førsteordens ODE'er og lineære andenordens ODE'er

Som en anvendelse af teorien i det forrige afsnit vil vi nu kort betragte en helt særlig type af andenordens ODE'er:

Definition 12.3.1

Lad $a_0, a_1 \in \mathbb{R}$ være konstanter og $q : \mathbb{R} \to \mathbb{R}$ en funktion. Da er en lineær andenordens ODE med konstante koefficienter en ODE på formen

$$f''(t) + a_1 \cdot f'(t) + a_0 \cdot f(t) = q(t). \tag{12.11}$$

Funktionen $q(t)$ kaldes på engelsk ODE'ens *forcing function*. Hvis denne forcing function $q(t)$ er nulfunktionen, kaldes ODE'en *homogen*, og ellers kaldes den *inhomogen*.

Som nævnt i Definition 12.1.2 opstiller man ofte begyndelsesværdibetingelser på formen $f(t_0) = y_0, f'(t_0) = y_1$ for en given $t_0 \in \mathbb{R}$ og givne værdier $y_0, y_1 \in \mathbb{R}$. Man kan vise, at hvis $q(t)$ er en differentiabel funktion, så har ODE'en i Ligning (12.11) netop én løsning, der opfylder en given begyndelsesværdibetingelse. For ODE'er som i Ligning (12.11) er en måde at finde denne løsning på først at bestemme dens fuldstændige løsning. Vi vil forklare, hvordan man gør dette, i dette afsnit ved hjælp af teorien fra det forrige afsnit. Følgende sætning indeholder nøgleelementet:

Sætning 12.3.1

Lad en funktion $f : \mathbb{R} \to \mathbb{R}$ være givet. Hvis f er en løsning til ODE'en

$$f''(t) + a_1 \cdot f'(t) + a_0 \cdot f(t) = q(t), \tag{12.12}$$

så er vektoren af funktioner $(f(t), f'(t))$ en løsning til systemet af ODE'er

$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ q(t) \end{bmatrix}. \tag{12.13}$$

Omvendt, hvis $(f_1(t), f_2(t))$ er en løsning til systemet af ODE'er i Ligning (12.13), så er $f_1(t)$ en løsning til ODE'en i Ligning (12.12).

Bevis. Dette overlades til læseren. □

Eksempel 12.3.1

En funktion $f(t)$ er en løsning til den lineære andenordens ODE $f''(t) + 5f'(t) + 6f(t) = 0$, hvis og kun hvis vektoren af funktioner $(f(t), f'(t))$ er en løsning til systemet af ODE'er

$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -6 & -5 \end{bmatrix} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix}.$$

Sætning 12.3.1 har følgende fine konsekvens for strukturen af løsningen til en lineær andenordens ODE.

Korollar 12.3.2

Lad en inhomogen lineær andenordens ODE

$$f''(t) + a_1 \cdot f'(t) + a_0 \cdot f(t) = q(t)$$

være givet, og antag, at $f_p(t)$ er en partikulær løsning til denne differentialligning. Da er enhver anden løsning $f(t)$ på formen $f_p(t) + f_h(t)$, hvor $f_h(t)$ er en løsning til den tilsvarende homogene ODE

$$f''(t) + a_1 \cdot f'(t) + a_0 \cdot f(t) = 0. \tag{12.14}$$

Bevis. Dette følger ved at kombinere Sætningerne 12.2.1 og 12.3.1. □

For at anvende Sætning 12.3.1 til at bestemme løsningen til en lineær andenordens ODE, er det nødvendigt at undersøge egenværdierne og egenvektorerne af den matrix, der optræder i sætningen. Undersøgelse af dens karakteristiske polynomium er derfor et vigtigt første skridt. Vi har

$$\begin{aligned} \det\left(\begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix} - Z\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) &= \det\left(\begin{bmatrix} -Z & 1 \\ -a_0 & -a_1 - Z \end{bmatrix}\right) \\ &= (-Z)\cdot(-a_1 - Z) - 1\cdot(-a_0) = Z^2 + a_1 Z + a_0. \end{aligned}$$

Dette motiverer for følgende definition:

Definition 12.3.2

Polynomiet

$$Z^2 + a_1 Z + a_0$$

kaldes *det karakteristiske polynomium* for ODE'en

$$f''(t) + a_1 \cdot f'(t) + a_0 \cdot f(t) = 0.$$

Med denne definition på plads kan vi helt eksplicit løse en homogen lineær andenordens ODE. Der er tre scenarier at skelne imellem afhængigt af, om polynomiet har to forskellige, reelle rødder, to komplekst konjugerede, ikke-reelle rødder, eller en reel rod med multiplicitet to (se Sætning 4.2.1).

Scenarie 1: Polynomiet $Z^2 + a_1 Z + a_0$ har to forskellige, reelle rødder. Hvis $Z^2 + a_1 Z + a_0$ har to forskellige, reelle rødder, betyder det, at dets diskriminant $D = a_1^2 - 4a_0$ er positiv, og at de reelle rødder er $\lambda_1 = \dfrac{-a_1 + \sqrt{D}}{2}$ og $\lambda_2 = \dfrac{-a_1 - \sqrt{D}}{2}$. Vi kunne nu benytte Sætningerne 12.3.1 og 12.2.3 til at bestemme den fuldstændige løsning til ODE'en i Ligning (12.14), men en direkte tilgang vil være hurtigere. Med vores indsigt i teorien om systemer af ODE'er forventer vi, at den fuldstændige løsning vil involvere funktionerne $e^{\lambda_1 t}$ og $e^{\lambda_2 t}$. Vi hævder derfor, at både $e^{\lambda_1 t}$ og $e^{\lambda_2 t}$ er løsninger til ODE'en i Ligning (12.14). For eksempel ser vi, at

$$\begin{aligned} (e^{\lambda_1 t})'' + a_1(e^{\lambda_1 t})' + a_0 e^{\lambda_1 t} &= \lambda_1^2 e^{\lambda_1 t} + a_1 \lambda_1 e^{\lambda_1 t} + a_0 e^{\lambda_1 t} \\ &= (\lambda_1^2 + a_1 \lambda_1 + a_0) e^{\lambda_1 t} \\ &= 0, \end{aligned}$$

hvor vi i den sidste lighed har udnyttet, at λ_1 er en rod i polynomiet $Z^2 + a_1 Z + a_0$. På meget lignende måde viser man, at funktionen $e^{\lambda_2 t}$ også er en løsning. Hvis $D = a_1^2 - 4a_0 > 0$, vil den fuldstændige løsning til ODE'en i Ligning (12.15) derfor være:

$$c_1 \cdot e^{\lambda_1 t} + c_2 \cdot e^{\lambda_2 t} = c_1 \cdot e^{\left(\frac{-a_1 + \sqrt{D}}{2}\right)t} + c_2 \cdot e^{\left(\frac{-a_1 - \sqrt{D}}{2}\right)t}, \quad c_1, c_2 \in \mathbb{R}. \tag{12.15}$$

Scenarie 2: Polynomiet $Z^2 + a_1 Z + a_0$ har to ikke-reelle rødder. I dette tilfælde er diskriminanten $D = a_1^2 - 4a_0$ negativ, og rødderne i $Z^2 + a_1 Z + a_0$ er $\lambda_1 = \dfrac{-a_1 + i\sqrt{|D|}}{2}$ og $\lambda_2 = \dfrac{-a_1 - i\sqrt{|D|}}{2}$. Meget lignende det forrige scenarie kan man vise, denne gang ved hjælp af Lemma 12.1.3, at både $e^{\lambda_1 t}$ og $e^{\lambda_2 t}$ er komplekse løsninger til ODE'en i Ligning (12.14). For at finde reelle løsninger tager vi blot realdelen og imaginærdelen af en af disse løsninger, inspireret af hvad vi gjorde i Korollar 12.2.6. Vi får

$$\mathrm{Re}(e^{\lambda_1 t}) = \mathrm{Re}(e^{(\frac{-a_1 + i\sqrt{|D|}}{2})t}) = e^{(\frac{-a_1}{2})t}\cos\left(\frac{\sqrt{|D|}}{2}t\right)$$

og på lignende måde

$$\mathrm{Im}(e^{\lambda_1 t}) = \mathrm{Im}(e^{(\frac{-a_1 + i\sqrt{|D|}}{2})t}) = e^{(\frac{-a_1}{2})t}\sin\left(\frac{\sqrt{|D|}}{2}t\right).$$

Hvis $D = a_1^2 - 4a_0 < 0$, vil den fuldstændige løsning til ODE'en i Ligning (12.14) derfor være:

$$c_1 \cdot e^{(\frac{-a_1}{2})t}\cos\left(\frac{\sqrt{|D|}}{2}t\right) + c_2 \cdot e^{(\frac{-a_1}{2})t}\sin\left(\frac{\sqrt{|D|}}{2}t\right), \quad c_1, c_2 \in \mathbb{R}. \tag{12.16}$$

Scenarie 3: Polynomiet $Z^2 + a_1 Z + a_0$ har en reel rod med multiplicitet to. I dette tilfælde er diskriminanten $D = a_1^2 - 4a_0$ lig nul, og dobbeltroden er givet ved $\lambda = -a_1/2$. Som i de forrige scenarier kan man direkte vise, at $e^{\lambda t}$ er en løsning til ODE'en i Ligning (12.14), men i dette scenarie mangler der endnu en løsning. Heldigvis kan vi igen få inspiration fra, hvad vi har været igennem med systemer af lineære ODE'er. I Eksempel 12.2.7 var vi i den situation, at den algebraiske multiplicitet af en egenværdi var to, men dens geometriske multiplicitet kun var én. Vi er i en lignende situation her. Hvis $D = 0$, så har matricen $\begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix}$, der forekommer i Sætning 12.3.1, egenværdien λ med algebraisk multiplicitet to, men man kan vise, at dens geometriske multiplicitet kun er én. Da funktionen $te^{\lambda t}$ dukkede op i Eksempel 12.2.7, vil det være naturligt at prøve at se, om denne funktion er en løsning til ODE'en i Ligning (12.14). Og det er den faktisk:

$$\begin{aligned}
(te^{\lambda t})'' + a_1(te^{\lambda t})' + a_0 te^{\lambda t} &= (e^{\lambda t} + t\lambda e^{\lambda t})' + a_1(e^{\lambda t} + t\lambda e^{\lambda t}) + a_0 te^{\lambda t} \\
&= (\lambda e^{\lambda t} + \lambda e^{\lambda t} + t\lambda^2 e^{\lambda t}) + a_1(e^{\lambda t} + t\lambda e^{\lambda t}) + a_0 te^{\lambda t} \\
&= (\lambda^2 + a_1\lambda + a_0)te^{\lambda t} + (2\lambda + a_1)e^{\lambda t} \\
&= (2\lambda + a_1)e^{\lambda t} \\
&= 0,
\end{aligned}$$

hvor vi i de sidste to ligheder har benyttet, at $\lambda^2 + a_1\lambda + a_0 = 0$, og $\lambda = -a_1/2$. Vi konkluderer, at hvis $D = a_1^2 - 4a_0 = 0$, så er den fuldstændige løsning til ODE'en i Ligning (12.14):

$$c_1 \cdot e^{\lambda t} + c_2 \cdot te^{\lambda t} = c_1 \cdot e^{\left(\frac{-a_1}{2}\right)t} + c_2 \cdot t \cdot e^{\left(\frac{-a_1}{2}\right)t}, \quad c_1, c_2 \in \mathbb{R}. \tag{12.17}$$

Vi afslutter afsnittet med en gennemgang af nogle flere eksempler.

Eksempel 12.3.2

Bestem den fuldstændige løsning til differentialligningen $f''(t) - 5f'(t) + 6f(t) = 0$.

Svar: Det karakteristiske polynomium for differentialligningen er $Z^2 - 5Z + 6$. Dette polynomium har en diskriminant lig 1 og derfor to forskellige reelle rødder. Udregnes disse rødder på den sædvanlige måde får man, at de er 2 og 3.

Ved at benytte Ligning (12.15) har vi derefter følgende fuldstændige løsning:

$$f(t) = c_1 e^{2t} + c_2 e^{3t}, \quad c_1, c_2 \in \mathbb{R}.$$

Eksempel 12.3.3

Bestem den fuldstændige løsning til differentialligningen $f''(t) - 4f'(t) + 4f(t) = 0$.

Svar: Det karakteristiske polynomium for differentialligningen er $Z^2 - 4Z + 4$, som har en diskriminant på nul. Mere præcist har det roden 2 med multiplicitet to. Ligning (12.17) fortæller nu, at den fuldstændige løsning, vi søger, er givet ved:

$$f(t) = c_1 e^{2t} + c_2 t e^{2t}, \quad c_1, c_2 \in \mathbb{R}.$$

Eksempel 12.3.4

Bestem den fuldstændige løsning til differentialligningen $f''(t) - 4f'(t) + 13f(t) = 0$.

Svar: I dette tilfælde er det karakteristiske polynomium for differentialligningen $Z^2 - 4Z + 13$, som har en negativ diskriminant, nemlig $D = (-4)^2 - 4 \cdot 13 = -36$. Derfor har det karakteristiske polynomium to ikke-reelle rødder, som viser sig at være $2 + 3i$ og $2 - 3i$. Ifølge Ligning (12.16) er den ønskede fuldstændige løsning:

$$f(t) = c_1 e^{2t} \cos(3t) + c_2 e^{2t} \sin(3t), \quad c_1, c_2 \in \mathbb{R}.$$

Afslutningsvis giver vi nogle eksempler på inhomogene lineære andenordens differentialligninger.

Eksempel 12.3.5

Bestem den fuldstændige løsning til følgende differentialligninger:

1. $f''(t) - 5f'(t) + 6f(t) = t$. Det er givet, at der findes en partikulær løsning på formen $f(t) = at + b$ med $a, b \in \mathbb{R}$.
2. $f''(t) - 4f'(t) + 4f(t) = e^t$. Det er givet, at $f(t) = e^t$ er en løsning.
3. $f''(t) - 4f'(t) + 13f(t) = 1$. Det er givet, at der findes en løsning på formen $f(t) = a$ med $a \in \mathbb{R}$.

Svar:

Jævnfør Korollar 12.4.2 og de tidligere eksempler er det nok at finde en partikulær løsning til hver af differentialligningerne.

1. Lad os forsøge at finde en partikulær løsning på formen $f(t) = at + b$ med $a, b \in \mathbb{R}$. Ved at indsætte dette i differentialligningen ser vi, at $0 - 4a + 4(at + b) = t$. Derfor får vi $4a = 1$ og $-4a + 4b = 0$. Vi ser, at $f(t) = t/4 + 1/4$ er en partikulær løsning. Ved at benytte Eksempel 12.3.2 og Korollar 12.4.2 konkluderer vi, at den fuldstændige løsning er givet ved:

$$f(t) = \frac{t}{4} + \frac{1}{4} + c_1 e^{2t} + c_2 e^{3t}, \quad c_1, c_2 \in \mathbb{R}.$$

2. Da vi er givet en partikulær løsning, kan vi finde den fuldstændige løsning direkte fra Eksempel 12.3.3 ved at benytte Korollar 12.4.2. Resultatet er:

$$f(t) = e^t + c_1 e^{2t} + c_2 t e^{2t}, \quad c_1, c_2 \in \mathbb{R}.$$

3. Først finder vi en partikulær løsning på formen $f(t) = a$. Ved at indsætte dette i differentialligningen ser vi, at $0 - 4 \cdot 0 + 13a = 1$, og derfor er $f(t) = 1/13$ en partikulær løsning. Ligesom tidligere kombinerer vi nu denne partikulære løsning med den fuldstændige løsning til den tilsvarende homogene ODE givet i Eksempel 12.3.4, og vi får følgende som den ønskede fuldstændige løsning til den givne inhomogene ligning:

$$f(t) = \frac{1}{13} + c_1 e^{2t} \cos(3t) + c_2 e^{2t} \sin(3t), \quad c_1, c_2 \in \mathbb{R}.$$

12.4 Relationen mellem systemer af lineære førsteordens ODE'er og lineære differentialligninger af højere orden

De samme ideer, som blev anvendt i det foregående afsnit, kan også anvendes på visse n'te-ordens differentialligninger. For fuldstændighedens skyld viser vi i dette afsnit, hvordan det kan foregå. Dette afsnit er ikke obligatorisk læsning og er kun ment som ekstra materiale til læsere, der ønsker at vide lidt mere. Vi starter på samme måde som for lineære andenordens ODE'er.

Definition 12.4.1

Lad n være et naturligt tal, $a_0, \ldots, a_{n-1} \in \mathbb{R}$ konstanter og $q : \mathbb{R} \to \mathbb{R}$ en funktion. En lineær n'te-ordens ODE med konstante koefficienter er en ODE på formen

$$f^{(n)}(t) + a_{n-1} \cdot f^{(n-1)}(t) + \cdots + a_1 \cdot f'(t) + a_0 \cdot f(t) = q(t). \tag{12.18}$$

Funktionen $q(t)$ kaldes på engelsk ODE'ens *forcing function*. Hvis denne forcing function $q(t)$ er nulfunktionen, kaldes ODE'en *homogen*, og ellers kaldes den *inhomogen*.

Som nævnt i Definition 12.1.2 arbejdes der ofte med begyndelsesværdibetingelser på formen $f(t_0) = y_0, f'(t_0) = y_1, \ldots, f^{(n-1)}(t_0) = y_{n-1}$ for et givet $t_0 \in \mathbb{R}$ og givne værdier $y_0, y_1, \ldots, y_{n-1} \in \mathbb{R}$. Man kan vise, at hvis $q(t)$ er en differentiabel funktion, så har ODE'en i Ligning (12.18) netop én løsning, der opfylder en given begyndelsesværdibetingelse. For

ODE'er som i Ligning (12.18) er en måde at finde denne løsning på at bestemme dens fuldstændige løsning først. Vi vil i dette afsnit beskrive, hvordan dette kan gøres.

Den primære ide er at relatere en løsning til en lineær n'te-ordens ODE med konstante koefficienter til en løsning til et passende valgt system af lineære førsteordens ODE'er.

Sætning 12.4.1

Lad en reel funktion $f : \mathbb{R} \to \mathbb{R}$ være givet. Hvis f er en løsning til ODE'en

$$f^{(n)}(t) + a_{n-1} \cdot f^{(n-1)}(t) + \cdots + a_1 \cdot f'(t) + a_0 \cdot f(t) = q(t), \tag{12.19}$$

så er vektoren af funktioner $(f(t), f'(t), \ldots, f^{(n-1)}(t))$ en løsning til følgende system af ODE'er:

$$\begin{bmatrix} f_1'(t) \\ f_2'(t) \\ \vdots \\ f_n'(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ -a_0 & \cdots & \cdots & -a_{n-2} & -a_{n-1} \end{bmatrix} \cdot \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ q(t) \end{bmatrix}. \tag{12.20}$$

Omvendt, hvis $(f_1(t), \ldots, f_n(t))$ er en løsning til systemet af ODE'er i Ligning (12.20), så er $f_1(t)$ en løsning til ODE'en i Ligning (12.19).

Bevis. Dette overlades til læseren. □

Sætning 12.4.1 medfører, at man kan anvende al den teori, vi har udviklet i det forrige afsnit, når man undersøger en ODE som Ligning (12.18). For eksempel kan vi konkludere følgende.

Korollar 12.4.2

Lad en inhomogen lineær n'te-ordens ODE

$$f^{(n)}(t) + a_{n-1} \cdot f^{(n-1)}(t) + \cdots + a_1 \cdot f'(t) + a_0 \cdot f(t) = q(t)$$

være givet, og antag, at $f_p(t)$ er en partikulær løsning til denne differentialligning. Da er enhver anden løsning $f(t)$ på formen $f_p(t) + f_h(t)$, hvor $f_h(t)$ er en løsning til den tilsvarende homogene ODE

$$f^{(n)}(t) + a_{n-1} \cdot f^{(n-1)}(t) + \cdots + a_1 \cdot f'(t) + a_0 \cdot f(t) = 0. \tag{12.21}$$

Bevis. Dette følger ved at kombinere Sætningerne 12.2.1 og 12.4.1. □

Som for systemer af n lineære førsteordens ODE'er kan man også her vise, at løsningsmængden til en homogen lineær n'te-ordens ODE udgør et vektorrum af dimension n. For at beskrive den fuldstændige løsning skal man derfor finde n lineært uafhængige løsninger. Som med systemer af lineære førsteordens ODE'er er første skridt mod bestemmelse af den fuldstændige løsning til en lineær n'te-ordens ODE at finde den fuldstændige løsning til den tilsvarende homogene ODE. Hvis vi vil benytte Sætning 12.4.1, vil første skridt være at bestemme det karakteristiske polynomium for matricer på den form, der optræder i Sætning 12.4.1. Heldigvis findes der en praktisk formel for de karakteristiske polynomier af sådanne matricer. Den virker endda over ethvert legeme $\mathbb{F}$.

Lemma 12.4.3

Lad $\mathbb{F}$ være et legeme, $n \geq 2$ et heltal og $a_0, \ldots, a_{n-1} \in \mathbb{F}$. Da er det karakteristiske polynomium for matricen

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ -a_0 & \cdots & \cdots & -a_{n-2} & -a_{n-1} \end{bmatrix}$$

lig med

$$p_{\mathbf{A}}(Z) = (-1)^n \cdot (Z^n + a_{n-1}Z^{n-1} + \cdots + a_1 Z + a_0).$$

Bevis. Vi beviser dette ved induktion på n for $n \neq 2$. Hvis $n = 2$, kan vi direkte se, at

$$p_{\mathbf{A}}(Z) = \det\left(\begin{bmatrix} -Z & 1 \\ -a_0 & -a_1 - Z \end{bmatrix}\right) = (-Z)\cdot(-a_1 - Z) - 1\cdot(-a_0) = Z^2 + a_1 Z + a_0.$$

Antag nu, at $n > 2$, og at resultatet er sandt for $n-1$. Udvikles determinanten af $\mathbf{A} - Z\mathbf{I}_n$ langs den første søjle får vi, at:

$$\det(\mathbf{A} - Z\mathbf{I}_n) = -Z \cdot \det\left(\begin{bmatrix} -Z & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & -Z & 1 & 0 \\ 0 & \cdots & 0 & -Z & 1 \\ -a_1 & \cdots & \cdots & -a_{n-2} & -a_{n-1} - Z \end{bmatrix}\right)$$
$$+ (-1)^n \cdot (-a_0) \cdot \det\left(\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -Z & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & -Z & 1 & 0 \\ 0 & \cdots & 0 & -Z & 1 \end{bmatrix}\right).$$

Benyttes induktionshypotesen på den første determinant efter ligheden og Sætning 8.1.4 på den anden determinant får vi, at

$$\begin{aligned} \det(\mathbf{A} - Z\mathbf{I}_n) &= (-Z) \cdot (-1)^{n-1} \cdot (Z^{n-1+}a_{n-1}Z^{n-2} + \cdots + a_1) + (-1)^{n-1} \cdot (-a_0) \cdot 1 \\ &= (-1)^n \cdot (Z^n + a_{n-1}Z^{n-1} + \ldots f a_1 Z + a_0). \end{aligned}$$

Dette fuldender induktionstrinnet. Dermed er lemmaet sandt for ethvert heltal $n \geq 2$. □

Bemærk, at matricen i Lemma 12.4.3 på engelsk kaldes en *companion matrix* for polynomiet $Z^n + a_{n-1}Z^{n-1} + \cdots + a_0$. Lemma 12.4.3 medfører, at når den lineære n'te-ordens ODE i Ligning (12.18) skal løses, så er det første, man bør gøre, at finde rødderne i polynomiet $Z^n + a_{n-1}Z^{n-1} + \ldots f a_1 Z + a_0$. Polynomiet

$$Z^n + a_{n-1}Z^{n-1} + \cdots + a_1 Z + a_0$$

kaldes ofte *det karakteristiske polynomium* for ODE'en

$$f^{(n)}(t) + a_{n-1} \cdot f^{(n-1)}(t) + \cdots + a_1 \cdot f'(t) + a_0 \cdot f(t) = 0.$$

Vi kunne fortsætte med at udvikle teorien om lineære n'te-ordens ODE'er yderligere, men i denne tekst vil vi afslutte vores studier af differentialligninger her.