Matematisk Metode

Jesper Lützen og Ian Kiming

17. oktober 2008

Contents

Introduktion. Den aksiomatisk-deduktive metode **ix**

1 Logik **1**
1.1 Udsagn og prædikater 1
1.2 Sammensatte udsagn 2
1.3 Sammensætning af prædikater 5
1.4 Kvantorer 6
1.5 Flere kvantorer 7
1.6 Bemærkning om brug af udsagn og prædikater 9
1.7 Definitioner 10
1.8 Opgaver 11

2 Beviser **13**
2.1 Gyldige slutninger (deduktioner). 13
2.2 Beviser 13
2.3 Direkte beviser 14
2.4 Modeksempler. 15
2.5 Formodninger og deres behandling. 15
2.6 Bevis ved kontraposition. 16
2.6.1 Eksempler på bevis ved kontraposition. 16
2.7 Bevis ved modstrid. 17
2.7.1 Eksempler på beviser ved modstrid: 18
2.8 Beviser delt op i tilfælde. 20
2.9 Eksistenssætninger 21
2.9.1 Eksempler på eksistenssætninger 22
2.9.2 Ikke-konstruktive beviser 22
2.10 Entydighedssætninger 23
2.10.1 Eksempel på entydighedssætning 23
2.11 Eksistens og entydighed 24
2.12 Opgaver 24

3 Euklids aksiomatisering af geometrien **25**
3.1 Bemærkninger til definitionerne. 25
3.2 Bemærkninger til postulaterne og de almindelige begreber. 26

3.3 Bemærkninger til sætningerne 27
3.4 Resten af Euklids Elementer 30

4 Analyse og Syntese. 31
4.1 Matematisk kreativitet 31
4.2 Eksistensproblemer 32
4.3 Analyse og syntese 34
4.4 Ligningsløsning. Et eksempel på analyse - syntese 34
4.5 Ikke stringent analyse 36
4.6 Geometriske eksempler 37
4.7 Advarsler 39
4.8 Terminologi 41
4.9 Problem- og sætningsanalyse 41

5 Induktionsbeviser 43
5.1 Simpel induktion 43
5.2 Fuldstændig induktion 46
5.3 Induktionsaksiomet 48
5.4 Rekursion 50
5.5 Opgaver 52

6 Mængdelære 55
6.1 Hvad er en mængde? 55
6.2 Delmængder 57
6.3 Fællesmængde og foreningsmængde 60
6.3.1 Familier af mængder. 62
6.4 Mængdedifferens og komplementærmængde 65
6.5 Mængdealgebra 66
6.6 Produktmængde 69
6.7 Potensmængden 71
6.8 Russels paradoks 71
6.9 Opgaver 73

7 Relationer 75
7.1 Ækvivalensrelationer 77
7.2 Ordningsrelationer 83
7.3 Opgaver 92

8 Afbildninger, funktioner 97
8.1 Injektivitet og surjektivitet 100
8.2 Billeder og urbilleder 102
8.3 Sammensætning af afbildninger, invers afbildning 105
8.4 Kardinalitet 110
8.5 Opgaver 110

9 Kompositionsregler, grupper og isomorfier 113
9.1 Kompositionsregler . . . 113
9.2 Grupper . . . 116
9.3 Gruppeisomorfier . . . 120
9.4 Ordningsisomorfier . . . 123

10 Aksiomatisk beskrivelse af de reelle tal 127
10.1 Legemer . . . 127
10.2 Ordnede legemer . . . 133
10.3 Fuldstændigt ordnede legemer. De reelle tal . . . 143
10.4 Opgaver . . . 145

11 Beviser i reel analyse, i tilknytning til Lindstrøms "Kalkulus" 147
11.1 Reelle talfølger . . . 148
11.2 Kontinuerte funktioner . . . 148
11.2.1 Kontinuitet af et produkt. Et bevis, og en bevisanalyse. . . 148
11.3 Hvordan man skal læse matematik . . . 151
11.4 Mellemværdisætningen . . . 153
11.5 Ekstremalværdisætningen . . . 156
11.6 Opgaver . . . 158

12 Construction of the real numbers. 159
12.1 Motivation. . . . 159
12.2 Definition of the real numbers. . . . 161
12.2.1 Fundamental definitions and basic properties. . . . 161
12.2.2 Definition of the set of real numbers. . . . 165
12.3 The order in $\mathbb{R}$ and further properties. . . . 170
12.3.1 Order and absolute value. . . . 170
12.3.2 Sequences in $\mathbb{R}$ and completeness. . . . 174
12.3.3 The least upper bound property. . . . 177
12.4 Final remarks. . . . 181
12.5 Exercises . . . 182

A Talsystemets opbygning 185

Forord

Disse noter er skrevet til undervisningen i kurset Matematisk Metode ved Institut for Matematiske Fag, Københavns Universitet. De er udarbejdet af Jesper Lützen og inkluderer omredigerede versioner af tidligere noter skrevet til kurset af Ian Kiming. Kapitlet om konstruktion af de reelle tal er skrevet af Ian Kiming

Introduktion. Den aksiomatisk-deduktive metode

Mange videnskaber er karakteriseret ved de objekter de omhandler: Botanik handler om planter og astronomi om himmelleger. Matematik er derimod karakteriseret ved sin metode. Det karakteristiske ved matematikken er at dens resultater kræver beviser. Naturligvis argumenterer man også i andre videnskaber, men i matematik er argumenterne eller beviserne mere stringente (dvs. strenge, præcise) end i andre videnskaber. Det er nok en væsentlig grund til at matematiske resultater har vist sig mere langtidsholdbare end resultaterne i andre videnskaber.

En matematisk sætning regnes for sand hvis den kan bevises ved logisk gyldige argumenter ud fra andre sætninger, som man ved er sande. Men disse andre sætninger bør jo så også bevises ud fra endnu andre o.s.v. For at bevisprocessen kan komme i gang, er det derfor klart at man må begynde med sætninger, som man ikke kræver beviser for, men som man postulerer. Det er de såkaldte aksiomer. Den her beskrevne metode kaldes den aksiomatisk-deduktive metode, fordi den fra aksiomer successivt deducerer (udleder, beviser) nye resultater.

Denne beskrivelse af den matematiske metode rejser to væsentlige problemer: Hvilke aksiomer skal man lægge til grund for matematikken og hvilke argumenter (slutningsregler) skal man acceptere som gyldige? Disse spørgsmål om matematikkens grundlag er emnet for kurset i matematisk logik og mængdelære. I Mat M vil vi nok opstille regler for korrekte slutninger (argumenter) og vi vil opstille aksiomer for visse dele af matematikken (geometrien og de reelle tal). Men vi vil ikke gå til bunds med grundlagsproblemerne. For eksempel vil vi ikke præcisere slutningsreglerne i mindste detalje, og vi vil ikke opstille aksiomerne for mængdelæren, selv om disse aksiomer ligger til grund for den videre aksiomatisering af for eksempel de reelle tal.

Kurset er et håndværkskursus. Vi vil introducere logikken som et værktøj, der bruges til daglig af arbejdende matematikere. Ligeså vil vi introducere mængder som samlinger af ting uden at kære os om de subtile problemer, der måtte være resultatet af en sådan "naiv" tilgang til mængdelæren. Det er

formålet med bogen at læseren skal kunne lære, hvordan han/hun skal læse og selv fremstille beviser for sætninger på et niveau af stringens, som er sædvanligt i matematisk forsknings- og lærebogslitteratur. Til det brug diskuterer vi særligt udbredte bevistyper og -strategier.

Et andet formål er at introducere og præcicere nogle matematiske begreber, som indgår i mange matematiske sammenhænge: mængder, relationer, ækvivalensrelationer, ordningsrelationer, funktioner, grupper, legemer, isomorfier, de reelle tal, konvergens og kontinuitet. En del af disse begreber (specielt de reelle tal og funktionsbegrebet) er i en vis forstand velkendte fra skolen, og spiller en dobbelt rolle i denne bog. I begyndelsen af bogen optræder de som eksempelmateriale. Når vi skal illustrere forskellige bevisstrategier, skal vi have et emne at bevise sætninger om. I denne del af bogen anser vi altså de reelle (og de hele og rationale) tal for velkendte, og benytter deres sædvanlige egenskaber. Senere i bogen vil vi præcicere disse begreber, og i de kapitler vil vi starte helt fra bunden, og systematisk opbygge de reelle tal, uden at gøre brug af tidligere tillært viden.

Chapter 1

Logik

Dette kapitel omhandler matematiske udsagn og prædikater. I et formelt kursus om logik opstiller man helt præcise regler for hvilke tegnstrenge, der kan tillades i opbygningen af udsagn og prædikater. I disse noter vil vi blot præcisere dagligdags logik.

1.1 Udsagn og prædikater

Definition 1 *Et (matematisk)* ***udsagn*** *er en udtalelse som er enten sand eller falsk.*

Eksempel 2

$$1 < 2 \quad og \quad 1 > 2 \tag{1.1}$$

er begge udsagn. Det første er sandt det andet er falsk. Derimod er

$$1 \int 2 \quad og \quad "matematik\ er\ smukt" \tag{1.2}$$

ikke udsagn.

Notation 3 *Vi betegner normalt udsagn med små bogstaver: $p, q, ...$*

Eksempel 4 *Betragt udtalelsen: $x < 2$. Det er ikke et udsagn, for når vi ikke har fastlagt værdien af x, er det hverken sandt eller falsk. Derimod bliver det et udsagn, når vi tillægger x en bestemt reel værdi. Vi siger da at x er en* ***fri variabel*** *og kalder $x < 2$ for et prædikat i denne variabel.*

Definition 5 *En udtalelse, der indeholder en fri variabel, kaldes et* ***prædikat*** *(eller et åbent udsagn) om elementerne i en given mængde. Det bliver et udsagn, når den frie variabel erstattes med et bestemt element i den givne mængde.*

Bemærkning 6 *Man kan på helt analog måde definere prædikater med flere frie variable. For eksempel er $x^2 > y$ et prædikat i to reelle variable.*

Notation 7 *Vi betegner normalt et prædikat i den variable x med $p(x)$, $q(x)$,.... Prædikater i to variable x, og y betegnes med $p(x, y)$, $q(x, y)$, og så videre.*

1.2 Sammensatte udsagn

Man kan lave sammensatte udsagn ud fra simple udsagn ved at bruge de logiske **konnektiver** $\wedge, \vee, \neg, \Rightarrow, \Leftrightarrow$. Sandheden af de sammensatte udsagn afhænger alene af sandheden af de indgående simple udsagn.

Definition 8 *Lad p og q være udsagn. Da defineres følgende sammensatte udsagn:*

- Konjunktion: $p \wedge q$ (læses "p og q") er sandt når både p og q er sande og ellers falsk.
- Disjunktion: $p \vee q$ (læses "p eller q") er sandt når enten p eller q eller de begge er sande og falsk når både p og q er falske.
- Negation: $\neg p$ (læses "non p") er sand når p er falsk og falsk når p er sand. I nogle bøger skrives $\sim p$ i stedet for $\neg p$.
- Implikation: $p \Rightarrow q$ (læses "p medfører q" eller "hvis p så q" eller "p kun hvis q") er falsk når p er sand og q er falsk; ellers er det sandt. Med andre ord siger implikationen $p \Rightarrow q$ at når p er sand så er q også sand, og den siger ikke mere end det. Man kan også skrive $q \Leftarrow p$ i stedet for $p \Rightarrow q$. Man kalder p for hypotesen og q for konklusionen.
- Biimplikation (ækvivalens): $p \Leftrightarrow q$ (læses "p er ensbetydende med q" eller "p hvis og kun hvis q") er sand hvis p og q har samme sandhedsværdi.

Sandhedstabeller. Man kan illustrere og præcisere disse definitioner i en tabel, hvori man anfører alle kombinationer af p's og q's sandhedsværdi (s for sand og f for falsk) og de tilhørende sandhedsværdier af de sammensatte udsagn:

p	q	$p \wedge q$	$p \vee q$	$\neg p$	$p \Rightarrow q$	$p \Leftrightarrow q$
s	s	s	s	f	s	s
s	f	f	s	f	f	f
f	s	f	s	s	s	f
f	f	f	f	s	s	s

(1.3)

Mere komplekse sammensatte udsagn. Man kan danne mere komplekse sammensatte udsagn ved at bruge tegnene $\wedge, \vee, \neg, \Rightarrow$ og $\Leftrightarrow$ efter hinanden. For eksempel kan man fra de tre simple udsagn p, q og r danne udsagnet: $(\neg(p \vee q)) \Rightarrow r$. En sandhedstabel for dette udsagn kan fås ved at kombinere informationerne i den ovenstående sandhedstabel:

p	q	r	$p \vee q$	$\neg(p \vee q)$	$(\neg(p \vee q)) \Rightarrow r$
s	s	s	s	f	s
s	s	f	s	f	s
s	f	s	s	f	s
s	f	f	s	f	s
f	s	s	s	f	s
f	s	f	s	f	s
f	f	s	f	s	s
f	f	f	f	s	f

(1.4)

Definition 9 *Et sammensat udsagn, som er falsk for alle sandhedsværdier af de indgående simple udsagn kaldes en* ***modstrid****.*

Eksempel 10 *Det sammensatte udsagn $p \wedge (\neg p)$ er en modstrid. Det ses let af sandhedstabellen:*

p	$\neg p$	$p \wedge (\neg p)$
s	f	f
f	s	f

(1.5)

Definition 11 *Et sammensat udsagn, som er sandt for alle sandhedsværdier af de indgående simple udsagn kaldes en* ***tautologi****.*

Eksempel 12 *Det sammensatte udsagn $((p \Rightarrow q) \wedge (q \Rightarrow p)) \Leftrightarrow (p \Leftrightarrow q)$ er en tautologi. Det ses af nedenstående sandhedstabel:*

p	q	$p \Rightarrow q$	$q \Rightarrow p$	$(p \Rightarrow q)$ $\wedge(q \Rightarrow p)$	$(p \Leftrightarrow q)$	$((p \Rightarrow q) \wedge (q \Rightarrow p))$ $\Leftrightarrow (p \Leftrightarrow q)$
s	s	s	s	s	s	s
s	f	f	s	f	f	s
f	s	s	f	f	f	s
f	f	s	s	s	s	s

(1.6)

Det her betragtede udsagn $((p \Rightarrow q) \wedge (q \Rightarrow p)) \Leftrightarrow (p \Leftrightarrow q)$ er en tautologi, fordi $((p \Rightarrow q) \wedge (q \Rightarrow p))$ og $(p \Leftrightarrow q)$ er sande for de samme kombinationer af sandhedsværdier af p og q. Man kan da opfatte dem som det samme udsagn, og vi siger at de er logisk ækvivalente. Man kan m.a.o. opfatte $(p \Leftrightarrow q)$ som en forkortelse af $(p \Rightarrow q) \wedge (q \Rightarrow p)$ eller kort skrevet: $\Leftrightarrow$ er en forkortelse for $\Rightarrow \wedge \Leftarrow$.

Notation 13 *Ligesom i almindelig algebra er der konventioner for i hvilken rækkefølge, man skal læse de logiske konnektiver. I et udtryk af formen $xy + z$ skal man foretage multiplikationen før additionen. På samme måde skal $\neg p \wedge q$*

læses som $(\neg p) \wedge q$ *og ikke som* $\neg (p \wedge q)$*. Vi siger at* $\neg$ *er mindst dominerende (skal bruges først) og* $\wedge$ *er mere dominerende (skal bruges bagefter). I følgende tabel er angivet hvilke af konnektiverne, der dominerer over hvilke:*

mindst dominerende	$\neg$*, negation*	*brug først*
	$\wedge$*, konjunktion;* $\vee$*, disjunktion*	
	$\Rightarrow$*, implikation*	
mest dominerende	$\Leftrightarrow$*, biimplikation*	*brug sidst*

(1.7)

Bemærk, at der ikke er nogen vedtagen rækkefølge for $\wedge$ *og* $\vee$*. Et udtryk af formen* $p \wedge q \vee r$ *har altså ingen mening, før man sætter parenteser: enten* $(p \wedge q) \vee r$ *eller* $p \wedge (q \vee r)$*. Selv i tilfælde, hvor der er en konvention om rækkefølgen af konnektiverne, kan det lette læsningen at sætte parenteser i udtryk for sammensatte udsagn.*

Øvelse 14 *Sæt parenteser i følgende sammensatte udsagn:*

$$p \vee \neg q \tag{1.8}$$

$$p \Rightarrow q \vee r \tag{1.9}$$

$$p \Rightarrow q \Leftrightarrow r \tag{1.10}$$

$$p \wedge q \Leftrightarrow r \Rightarrow \neg q \tag{1.11}$$

Definition 15 *To sammensatte udsagn* p *og* q *kaldes* ***logisk ækvivalente****, og vi skriver* $p \equiv q$*, hvis* $p \Leftrightarrow q$ *er en tautologi.*

Sætning 16 ***Logisk huskeseddel.*** *Det er praktisk at huske følgende logiske ækvivalenser udenad*

$$\neg\neg p \equiv p \ , \tag{1.12}$$

$$p \ \wedge \ q \equiv q \ \wedge \ p \ , \qquad p \ \vee \ q \equiv q \ \vee \ p \ , \tag{1.13}$$

$$(p \ \wedge \ q) \wedge \ r \equiv p \ \wedge (q \ \wedge \ r) \ , \qquad (p \ \vee \ q) \vee \ r \equiv p \ \vee (q \ \vee \ r) \ , \tag{1.14}$$

$$(p \ \wedge \ q) \vee \ r \equiv (p \ \vee \ r) \wedge (q \ \vee \ r) \ , \qquad (p \ \vee \ q) \wedge \ r \equiv (p \ \wedge \ r) \vee (q \ \wedge \ r) \ , \tag{1.15}$$

$$\neg(p \ \wedge \ q) \equiv \ \neg p \ \vee \neg q \ , \qquad \neg(p \ \vee \ q) \equiv \ \neg p \ \wedge \neg q \ , \tag{1.16}$$

$$p \Rightarrow q \ \equiv \ \neg p \ \vee \ q \ , \tag{1.17}$$

$$p \Rightarrow q \ \equiv \ \neg q \Rightarrow \neg p \tag{1.18}$$

Bevis. Disse logiske ækvivalenser kan eftervises ved at betragte sandhedstabellerne for udsagnene. ■

Definition 17 *Når* $p \Rightarrow q$ *er en implikation kaldes implikationen* $\neg q \Rightarrow \neg p$ *den* ***kontraponerede*** *implikation.*

Ifølge (1.18) er en implikation og dens kontraponerede logisk ækvivalente.

Definition 18 *Når $p \Rightarrow q$ er en implikation kaldes $q \Rightarrow p$ den omvendte implikation.*

Øvelse 19 *Vis ved et eksempel, at en implikation og dens omvendte ikke er logisk ækvivalente.*

Bemærkning 20 *For at fremme forståelsen omformer man helst sammensatte udsagn og prædikater så de* ***så vidt muligt ikke indeholder negationer****. For eksempel vil man om et naturligt tal hellere sige at "x er et lige primtal" end at sige: "x er hverken sammensat eller ulige".*

Øvelse 21 *Vis ved hjælp af reglerne på huskesedlen, at de to prædikater i bemærkning (20) ovenfor er logisk ækvivalente.*

Øvelse 22 *Vis, at $((p \Rightarrow q) \wedge (q \Rightarrow r)) \Rightarrow (p \Rightarrow r)$ er en tautologi.*

Notation 23 *På grund af ovenstående resultat kan man tillade sig at forkorte udsagnet $(p \Rightarrow q) \wedge (q \Rightarrow r)$ til $p \Rightarrow q \Rightarrow r$.*

1.3 Sammensætning af prædikater

Man kan naturligvis sammensætte prædikater på samme måde som udsagn. Hvis $p(x)$ og $q(x)$ er prædikater om elementerne i mængden M, kan man for eksempel danne prædikatet $p(x) \wedge q(x)$. Når x erstattes af et bestemt element i M, bliver $p(x)$ og $q(x)$ til udsagn, og dermed bliver også $p(x) \wedge q(x)$ til et udsagn.

Konvention. På tilsvarende vis kan man ud fra prædikaterne $p(x)$ og $q(x)$ danne prædikaterne $p(x) \Rightarrow q(x)$ og $p(x) \Leftrightarrow q(x)$. Men her er der en konvention om at læse $p(x) \Rightarrow q(x)$ og $p(x) \Leftrightarrow q(x)$ ikke som prædikater i den frie variabel x, men som udsagnene: "$p(x) \Rightarrow q(x)$ for alle x i M" og "$p(x) \Leftrightarrow q(x)$ for alle x i M". Med denne gængse fortolkning betyder $p(x) \Rightarrow q(x)$ altså: "$q(x)$ er sand for alle de værdier af x, som gør $p(x)$ sand" eller "Hvis $p(x)$ er sand for en værdi af x, så er $q(x)$ sand for den samme værdi af x". På samme vis læses $p(x) \Leftrightarrow q(x)$ altså som udsagnet: "$p(x)$ og $q(x)$ er sande for de samme værdier af x".

Bemærkning 24 *Det kan virke underligt, at man definerer at $p \Rightarrow q$ er sand når både p og q er falske.*

Betragt for eksempel følgende implikationer:

$$2 < 1 \Rightarrow \sin 2 = 10^5. \tag{1.19}$$

"Hvis Anders Fogh Rasmussen er præsident i USA, (1.20)

så er månen lavet af grøn ost" (1.21)

I vores almindelige omgang med sproget vil vi nok karakterisere disse udsagn som noget vrøvl, dels fordi hypotesen ikke har noget at gøre med konklusionen, og dels fordi de fire elementære udsagn, implikationerne er sammensat af, alle er falske. Men i matematisk logik er udsagnene sande.

For at forstå hvorfor man har valgt at noget falsk kan medføre både noget falsk og noget sandt, kan vi se på implikationer $p(x) \Rightarrow q(x)$ mellem prædikater, for her svarer konventionen meget bedre til vores umiddelbare dagligdags omgang med sproget. Betragt for eksempel følgende implikation om reelle tal:

$$x > 2 \Rightarrow x^2 > 4 \tag{1.22}$$

Hvis vi skal vise at denne implikation er sand er vores umiddelbare strategi at undersøge x'er som er større end 2 og vise at de har $x^2 > 4$. Da dette er korrekt, slutter vi at implikationen er sand. Vi undersøger slet ikke hvad der sker når x er mindre end eller lig med 2, fordi vi opfatter disse værdier af x som irrelevante for implikationens sandhedsværdi. Denne strategi er korrekt, netop fordi vi har defineret, at $p \Rightarrow q$ er sand, hvis p er falsk (uanset sandhedsværdien af q). Konventionen ovenfor betyder jo, at implikationen $x > 2 \Rightarrow x^2 > 4$ skal læses "$x > 2 \Rightarrow x^2 > 4$ for alle $x \in \mathbb{R}$". Vi burde altså egentligt undersøge alle reelle x, men netop fordi vi har defineret at $x > 2 \Rightarrow x^2 > 4$ er sand for de x, der ikke opfylder $x > 2$, er der ingen grund til at undersøge sådanne x'er. Altså er det nok at undersøge de værdier af x, som gør hypotesen sand.

1.4 Kvantorer

Definition 25 *Ud fra et prædikat $p(x)$ om elementerne i en mængde M kan vi danne to udsagn:*

Udsagnet

$$\forall x \in M : p(x) \tag{1.23}$$

er sandt, netop når $p(x)$ er sand for alle elementerne x i M. Man siger (og skriver): "for alle (eller for ethvert, eller for et vilkårligt) x i M (gælder) $p(x)$".

Udsagnet

$$\exists x \in M : p(x) \tag{1.24}$$

er sandt, netop når der eksisterer (mindst) et element x i M, som gør $p(x)$ sand. Man siger (og skriver): "Der eksisterer et x i M, så $p(x)$".

*Tegnene $\forall$ og $\exists$ kaldes **kvantorer**. $\forall$ kaldes **al-kvantoren** og $\exists$ kaldes **eksistens-kvantoren**.*

Eksempel 26 *Betragt prædikatet $p(x)$: $x^2 \geq 0$. Dette prædikat bliver et sandt udsagn, uanset hvilket reelt tal vi sætter ind på x's plads. Altså er udsagnet $\forall x \in \mathbb{R} : x^2 \geq 0$ sandt. Vi siger at "for alle reelle x gælder $x^2 \geq 0$".*

Eksempel 27 *Udsagnet $\exists x \in \mathbb{R} : x^2 < 1$ er sandt, da der findes et x (for eksempel $x = 0$), som gør prædikatet $x^2 < 1$ sandt. Derimod er $\forall x \in \mathbb{R} : x^2 < 1$ et falsk udsagn, da prædikatet $x^2 < 1$ ikke er sandt for alle $x \in \mathbb{R}$. Det er jo for eksempel falsk for $x = 5$.*

Bemærkning 28 *Når man sætter en kvantor foran den frie variabel i et prædikat i én variabel, får man et udsagn og ikke et prædikat. Variablen er ikke længere fri og kan ikke tillægges forskellige værdier. Man siger da at den variable er bunden.*

Eksempel 29 *Betragt tegnstrengen:* $\forall x \in \mathbb{R} : x^2 \geq 0$. *Det er ikke et prædikat men et udsagn (som er sandt), og da der står en alkvantor foran* x*, er dette en bunden variabel.*

Konvention. Den ovennævnte konvention kan nu formuleres som følger: $p(x) \Rightarrow q(x)$ skal normalt fortolkes som udsagnet: $\forall x : p(x) \Rightarrow q(x)$; og $p(x) \Leftrightarrow q(x)$ skal normalt fortolkes som $\forall x : p(x) \Leftrightarrow q(x)$.

Bemærkning 30 *Hvis det er helt klart hvilken mængde variablen* x *varierer over, kan man udelade at angive denne når man kvantificerer over* x*. Hvis det for eksempel er klart at vi arbejder med de reelle tal, kan man skrive* $\forall x : x^2 \geq 0$

1.5 Flere kvantorer

Hvis $p(x, y)$ er et prædikat i de to variable x og y, og vi kvantificerer over den ene variabel (f.eks x), så vil resultatet være et prædikat i den anden variabel.

Eksempel 31 *Prædikatet* $\forall x \in \mathbb{R} : x^2 > y$ *er et prædikat i den fri reelle variabel* y*. Derimod er* x *en bunden variabel. Prædikatet er sandt for negative værdier af* y *og falsk for positive værdier af* y*.*

Bemærkning om flere kvantorer. Da $\forall x \in \mathbb{R} : x^2 > y$ er et prædikat i den reelle variabel y, kan vi kvantificere over y. Dermed opnås et udsagn, hvori begge de to variable er bundne. For eksempel kan vi danne udsagnet: $\exists y \in \mathbb{R}(\forall x \in \mathbb{R} : x^2 > y)$. Dette er et sandt udsagn, thi der eksisterer jo et y (for eksempel $y = -1$), som gør prædikatet $\forall x \in \mathbb{R} : x^2 > y$ sandt. Derimod er udsagnet $\forall y \in \mathbb{R}(\forall x \in \mathbb{R} : x^2 > y)$ falsk. Man udelader normalt parentesen og skriver for eksempel $\exists y \in \mathbb{R} \forall x \in \mathbb{R} : x^2 > y$.

Bemærkning om kvantorernes rækkefølge. Det er vigtigt at bemærke at kvantorernes rækkefølge ikke er ligegyldig. For eksempel er udsagnet

$$\forall y \in \mathbb{R} \exists x \in \mathbb{R} : x^2 > y \tag{1.25}$$

sandt, da man altid kan vælge x så stor, at x^2 bliver større end et givet y, uanset hvilken værdi y tillægges. Derimod er udsagnet

$$\exists x \in \mathbb{R} \forall y \in \mathbb{R} : x^2 > y \tag{1.26}$$

falsk. Det er jo ikke muligt at finde et x, så x^2 bliver større end alle reelle tal y.

Der er dog nogle tilfælde, hvor man gerne må bytte om på kvantorerne:

Sætning 32 *Lad $p(x,y)$ være et prædikat i de frie variabel x og y. Der gælder følgende implikationer:*

$$(\exists x\ \exists y:\ p(x,y)) \quad \Leftrightarrow \quad (\exists y\ \exists x:\ p(x,y))\ , \tag{1.27}$$

$$(\forall x\ \forall y:\ p(x,y)) \quad \Leftrightarrow \quad (\forall y\ \forall x:\ p(x,y))\ , \tag{1.28}$$

samt

$$(\exists x\ \forall y:\ p(x,y)) \quad \Rightarrow \quad (\forall y\ \exists x:\ p(x,y))\ . \tag{1.29}$$

Formuleres de to første i ord, er de intuitivt klare. Det gælder også den tredie regel: Antag, at der findes x (kald et sådant x_0), således at der for alle y gælder $p(x_0,y)$. Da vil der for ethvert y findes x, (for eksempel det førnævnte x_0), så $p(x_0,y)$.

Bemærkning 33 *Udsagnet $(\forall y\ \exists x:\ p(x,y)) \Rightarrow (\exists x\ \forall y:\ p(x,y))$ er derimod ikke i almindelighed sandt:*

Eksemplet i ovenstående bemærkning giver et modeksempel. Som et andet simpelt modeksempel kunne man eksempelvis lade $p(x,y)$ være prædikatet $x^2 = y$ hvor de frie variable x og y tillades at løbe over de positive reelle tal $\mathbb{R}_+$. I så fald er udsagnet $\forall y\ \exists x : x^2 = y$ sandt: Det siger blot, at ethvert positivt reelt tal har en positiv kvadratrod. Men udsagnet $\exists x\ \forall y : x^2 = y$ er jo klart falsk: Der findes naturligvis intet positivt reelt tal x, hvis kvadrat er lig ethvert positivt reelt tal y.

Eksempel 34 *Punktvis og uniform kontinuitet:*

De ovennævnte eksempler på at kvantorer ikke kan ombyttes er legetøjseksempler. Her skal nævnes et tilfælde, hvor problemet bliver akut i en vigtig matematisk sammenhæng. Lad den reelle funktion f være defineret på en delmængde M af $\mathbb{R}$. Som bekendt siges f da at være kontinuert i punktet $x_0 \in M$, hvis

$$\forall \epsilon \in \mathbb{R}_+ \exists \delta \in \mathbb{R}_+ \forall x \in M : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon \tag{1.30}$$

Endvidere siges f at være **(punktvis) kontinuert** i M, hvis den er kontinuert i alle punkter $x_0 \in M$, dvs. hvis

$$\forall x_0 \in M \forall \epsilon \in \mathbb{R}_+ \exists \delta \in \mathbb{R}_+ \forall x \in M : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon \tag{1.31}$$

I følge ovenstående regler har vi lov til at bytte om på de to første alkvantorer i denne definition. Men hvad hvis vi flytter $\forall x_0 \in M$ helt hen efter $\exists \delta \in \mathbb{R}_+$? Så fås følgende betingelse på f:

$$\forall \epsilon \in \mathbb{R}_+ \exists \delta \in \mathbb{R}_+ \forall x_0 \in M \forall x \in M : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon \tag{1.32}$$

Hvis f opfylder dette kaldes den pr. definition **uniformt kontinuert** på mængden M. Det følger af (1.29) at hvis en funktion er uniform kontinuert på en mængde M, så er den også punktvis kontinuert. Derimod er $f(x) = 1/x$ defineret på intervallet $]0,1]$ et eksempel på en funktion, som er punktvis kontinuert men ikke uniform kontinuert. En hovedsætning siger dog, at på en lukket og begrænset mængde er uniform kontinuitet det samme som punktvis kontinuitet.

Øvelse 35 *Argumenter for at $f(x) = 1/x$ defineret på intervallet $]0,1]$ ikke er uniformt kontinuert.*

Sætning 36 ***Logisk huskeseddel fortsat.*** *Man har følgende regler:*

$$\neg(\forall x \ : \ p(x)) \ \equiv \ \exists x \ : \ \neg p(x) \ , \qquad \neg(\exists x \ : \ p(x)) \ \equiv \ \forall x \ : \ \neg p(x) \ . \quad (1.33)$$

Disse tillader negation af længere strenge med al– og/eller eksistenskvantorer. Eksempelvis:

$$\neg(\forall x \ \exists y \ \forall z \ : \ p(x,y,z)) \ \equiv \ \exists x \ \forall y \ \exists z \ : \ \neg p(x,y,z) \ . \quad (1.34)$$

Endelig gælder følgende regel, hvis $p(x)$ er et prædikat og q et udsagn:

$$(\forall x \ : \ p(x)) \wedge q \ \equiv \ \forall x \ : \ p(x) \wedge q \ , \quad (1.35)$$

samt de tilsvarende, der fås hvis $\wedge$ og/eller $\forall$ udskiftes med $\vee$ hhv. $\exists$.

1.6 Bemærkning om brug af udsagn og prædikater

Når der i en matematisk tekst står et udsagn, betyder det da "dette udsagn er sandt" eller betyder det "her står et udsagn, som kan være sandt eller falsk"? Det korte svar på spørgsmålet er, at det afhænger af sammenhængen. I en tekst om logik vil der ofte stå udsagn, som kan være sande eller falske. For eksempel har vi ovenfor skrevet forskellige udsagn uden at implicere at de var sande. Når man derimod formulerer en sætning i matematik, er meningen at udsagnet i sætningen er sand. Nedenfor formuleres for eksempel De Morgan's love og meningen er naturligvis, at disse er sande.

Man kan sige noget lignende om prædikater. Når man skriver $x > 2$, kan man mene: Her står et prædikat i den reelle variable x. Men oftere mener man "x er et reelt tal, som gør prædikatet $x > 2$ sandt".

Denne usystematiske brug af udsagn i matematik giver sjældent anledning til forvirring. Men jeg vil dog fremhæve en argumentationsform, hvor der ofte opstår forvirring: Hvis der midt i et matematisk bevis står en implikation f.eks. $p \Rightarrow q$, så betyder det at følgende udsagn er sandt: "Hvis p er sand da er q sand". Det betyder *ikke* "Da p er sand er også q sand". Hvis man vil sige at p er sand, må man skrive det eksplicit. Lad os se på et eksempel:

Sætning: Hvis et kvadrat har en side, der er større end 2 så er dets areal større end 4.

En udbredt fejlagtig præsentation af beviset går på følgende vis: Hvis siden i kvadratet kaldes x, er $x > 2$. Da $x > 2$ gælder at

$$x > 2 \Rightarrow x^2 > 4, \quad (1.36)$$

hvorfor arealet $x^2 > 4$.

Forfatteren af argumentet mente nok følgende slutning (det, der står i parenteserne behøver man ikke skrive. Det kan underforstås): "Da $x > 2$ (er sand)

er $x^2 > 4$ (sand)". Men når forfatteren skriver at $x > 2 \Rightarrow x^2 > 4$ "gælder" så betyder det bare, at *hvis* $x > 2$ (er sand) *så* er $x^2 > 4$ (sand). Men dette udsagn er sandt uanset værdien af $x \in \mathbb{R}$. Der er derfor ikke nogen grund til at skrive, at *da* $x > 2$ *så* gælder $x > 2 \Rightarrow x^2 > 4$. Og argumentet er slet ikke færdigt når vi har udledt at $x > 2 \Rightarrow x^2 > 4$ for det betyder jo ikke at $x^2 > 4$ er sandt, sådan som vi skulle konkludere i beviset.

Vi kan dog fra $x > 2$ og $x > 2 \Rightarrow x^2 > 4$ slutte at $x^2 > 4$ (modus ponens, se senere). Men det bliver meget tungt at bruge pile i sådanne tilfælde. Det er derfor ofte sikrest og mest elegant at undgå brug af pile.

Hovedreglen synes at være at når der står et længere sammensat udsagn i en matematisk tekst, så mener vi, at dette udsagn er sandt, men vi mener *ikke* at de simplere udsagn, som indgår i det sammensatte udsagn er sande. Når vi skriver $p \Rightarrow q$ så mener vi (ofte) at det er sandt, at hvis p er sand, så er q også sand. Men det betyder ikke at vi mener at p er sand (og derfor også q er sand).

Tilsvarende, når vi om en funktion der er kontinuert i a skriver at

$$\forall \epsilon > 0 \exists \delta > 0 : |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon \tag{1.37}$$

så mener vi naturligvis at hele dette sammensatte udsagn er sandt, men vi mener *ikke* at udsagnet $|x - a| < \delta$ er sandt. Hvis vi derfor vil antage dette i et bevis, må vi eksplicit sige det: "Lad $|x - a| < \delta$".

1.7 Definitioner

Definitioner er sætninger som fortæller, hvad ord eller tegn betyder. I modsætning til aksiomerne indeholder definitionerne ikke egentlig ny information om den matematiske teori, og i modsætning til sætningerne kræver de ikke noget bevis. Lad os se på nogle eksempler på definitioner, som vi skal bruge i det følgende. Det drejer sig om definitioner inden for teorien for de naturlige tals aritmetik. Vi skal ikke nævne aksiomerne for denne teori eksplicit. For vores formål her er det nok at vide at de blot specificerer de regler, vi er vokset op med fra barns ben.

Definition 37 *Et helt tal x kaldes **lige**, hvis der findes et helt tal n så $x = 2n$.*

Definition 38 *Et helt tal kaldes **ulige**, hvis det ikke er lige.*

Definition 39 *Et naturligt tal x kaldes et **sammensat** tal, hvis det har en faktorisering $x = nm$, hvor n og m er naturlige tal med $1 < n, m < x$.*

Definition 40 *Et naturligt tal forskelligt fra 1 kaldes et **primtal**, hvis det ikke er et sammensat tal.*

Bemærkninger om definitioner. I formuleringen af definitionerne indgår der ordet "hvis". Der burde egentlig stå "hvis og kun hvis". Den første definition skal for eksempel ikke blot betyde at tal af formen $2n$ er lige, men også at tal,

der ikke er af denne form, ikke er lige. Der er dog tradition for kun at skrive "hvis" i definitioner.

Selv om man i matematikken ofte benytter ord, som også har en dagligdags betydning betyder de i matematik kun det, som er specificeret i definitionen. For eksempel har ordene "grænse" og "kontinuitet" i matematikken en betydning, som kun til en vis grad stemmer overens med den dagligdags betydning af ordene. Det er naturligvis vigtigt at danne sig intuitive billeder af hvad de matematiske ord og begreber dækker, men i sidste ende er det kun definitionerne, der bestemmer betydningen af begreberne og ordene. Man kan for eksempel have en intuitiv fornemmelse af, at kontinuitet af en funktion betyder, at dens graf hænger sammen, og det er også en nyttig intuition. Men man kan blive snydt af den. For eksempel kan den vildlede en til at tro, at funktionen f defineret ved

$$f(x) = \left\{ \begin{array}{c} \sin \frac{1}{x} \text{ for } x \neq 0 \\ 0 \text{ for } x = 0 \end{array} \right\} \tag{1.38}$$

er kontinuert i 0. Men definitionen af kontinuitet afgør, at funktionen er diskontinuert i 0.

Bemærk også at definition af et matematisk begreb ikke sikrer, at der i teorien eksisterer et objekt af den slags som defineres. Eksistens må etableres ud fra aksiomerne i teorien. For eksempel kan man godt definere, at et lige primtal større end 2 kaldes et stor-lige primtal. Der er dog ingen naturlige tal som er stor-lige primtal. Ligeså kan man godt i euklidisk geometri definere en retvinklet femkant som en femkant med fem rette vinkler. Det viser sig bare at sådanne femkanter ikke findes. Omvendt, når man har defineret et kvadrat som en firkant med fire rette vinkler og fire lige store sider, så må man først bruge sådanne firkanter i sin teori, når man har vist deres eksistens, uanset hvor intuitivt det kan synes, at der findes kvadrater.

Når et lighedstegn bruges til at definere et matematisk objekt skrives ofte $:=$. For eksempel kan man skrive at intervallet $[a, \infty[$ defineres som

$$[a, \infty[:= \{x \in \mathbb{R} \mid a \leq x\}\,. \tag{1.39}$$

Når en biimplikationspil bruges til at definere et udsagn eller et prædikat, skriver man ofte $\overset{def}{\Leftrightarrow}$. For eksempel vil vi senere definere relationen $<$ som følger:

$$x < y \overset{def}{\Leftrightarrow} (x \leq y) \wedge x \neq y \tag{1.40}$$

1.8 Opgaver

1. Opskriv sandhedstabeller for følgende sammensatte udsagn:

$$(p \vee \neg q) \wedge (\neg p \vee q) \quad \text{og} \quad (p \vee q) \wedge (\neg p \vee \neg q) \tag{1.41}$$

2. Hvilke af følgende sammensatte udsagn er logisk ækvivalente?

$$p \quad \Rightarrow \quad q, \;\; q \Rightarrow p, \;\; \neg\,(p \Rightarrow q)\,, \tag{1.42}$$

$$p \quad \Rightarrow \quad \neg q, \;\; \neg p \Rightarrow q, \;\; \neg p \Rightarrow \neg q. \tag{1.43}$$

3. Vis at følgende sammensatte udsagn er en tautologi:

$$(p \Rightarrow q) \vee (\neg p \Rightarrow q) \tag{1.44}$$

4. Skriv følgende udsagn ved brug af kvantorer:
a. Ligningen $x^3 = 7$ har mindst en rod.
b. Ligningen $x^2 - 2x - 5 = 0$ har ingen rational rod.
c. Enhver ligning af formen $x^3 = a$ har en rod.
d. Der findes ingen ligninger af formen $x^n = a$, der ikke har rødder.
e. Der findes intet helt tal, som er større end alle andre hele tal.
5. Lad $p(x)$ og $q(x)$ være prædikater om elementerne i en mængde M.
Er følgende udsagn logisk ækvivalente?
a. $\forall x \in M : p(x) \wedge q(x)$ og $(\forall x \in M : p(x)) \wedge (\forall x \in M : q(x))$
b. $\forall x \in M : p(x) \vee q(x)$ og $(\forall x \in M : p(x)) \vee (\forall x \in M : q(x))$
c. $\exists x \in M : p(x) \vee q(x)$ og $(\exists x \in M : p(x)) \vee (\exists x \in M : q(x))$
d. $\exists x \in M : p(x) \wedge q(x)$ og $(\exists x \in M : p(x)) \wedge (\exists x \in M : q(x))$
6. Lad det være givet, at Kurt kun spiser is, når solen skinner. Lad $p(t)$ og $q(t)$ være følgende prædikater:
$p(t)$: Solen skinner til tidspunktet t.
$q(t)$: Kurt spiser en is til tidspunktet t.
Hvilken implikation gælder mellem $p(t)$ og $q(t)$.
7. Bestem de kontraponerede til følgende udsagn. Brug positiv udtryksmåde hvis det er muligt.
a. Hvis $x < 0$, så er $x^2 > 0$.
b. Hvis $x \neq 0$, så eksisterer der et y så $xy = 1$.
c. Hvis x er et lige helt tal, så er x^2 et lige helt tal.
d. Hvis $x + y$ er ulige og $y + z$ er ulige, så er $x + z$ lige
e. Hvis f er et polynomium af ulige grad, så har f mindst én reel rod.
8. Bevis nogle af de logiske ækvivalenser i Sætning 16, ved at opskrive sandhedstabeller.
9. Giv et eksempel på en sand implikation, hvis omvendte implikation er sand, og én hvor den omvendte implikation er falsk.
10. Neger følgende udsagn:
a. $x > 0$ og x er rational.
b. l er enten parallel med m, eller også er l lig med m. (l, m er linjer)
11. Opskriv med kvantorer hvad det betyder at en reel funktion ikke er kontinuert i punktet a. (neger (1.37))

Chapter 2

Beviser

En matematisk teori består af en række udsagn (spilleregler) kaldet *aksiomerne,* som man regner for sande i teorien, og hvorfra man beviser "sætninger", det vil sige andre udsagn, som er sande i teorien. Vi skal i dette afsnit se på, hvordan man beviser sætninger. Lad os først formalisere bevisbegrebet lidt mere:

2.1 Gyldige slutninger (deduktioner).

Hvis $p_1, p_2, \ldots, p_n$ og q er udsagn, siger vi, at q *kan sluttes af* $p_1, p_2, \ldots, p_n$, eller at vi har en *gyldig slutning (en deduktion) af* q *fra udsagnene* $p_1, p_2, \ldots, p_n$, såfremt:

$$(p_1 \wedge \ldots \wedge p_n) \Rightarrow q \text{ er en tautologi} \tag{2.1}$$

Her er nogle vigtige eksempler på gyldige slutninger:

$$\text{Af } (p \Rightarrow q \text{ og } p) \text{ kan } q \text{ sluttes} \quad \text{(Modus ponens).} \tag{2.2}$$

$$\text{Af } (p \Rightarrow q \text{ og } \neg q) \text{ kan } \neg p \text{ sluttes} \quad \text{(Modus tollens).} \tag{2.3}$$

$$\text{Af } (p \vee q \text{ og } \neg p) \text{ kan } q \text{ sluttes} \quad \text{(Disjunktiv syllogisme).} \tag{2.4}$$

$$\text{Af } (p \Rightarrow q \text{ og } q \Rightarrow r) \text{ kan } p \Rightarrow r \text{ sluttes} \quad \text{(Hypotetisk syllogisme).} \tag{2.5}$$

$$\text{Af } (p \vee q \text{ og } p \Rightarrow r \text{ og } q \Rightarrow r) \text{ kan } r \text{ sluttes} \quad \text{(Dilemma).} \tag{2.6}$$

Øvelse: Vis ved brug af sandhedstabeller at disse slutninger er gyldige.

2.2 Beviser

Et **bevis** for et udsagn q består af en kæde af udsagn $p_1, p_2, \ldots, p_n$, således at:

- $p_n = q$

- For hvert $i = 1, \ldots, n$ er udsagnet p_i enten et aksiom i vores teori, eller et tidligere bevist udsagn, eller fremgår ved en *gyldig slutning* af udsagnene $p_1, \ldots, p_{i-1}$.

Når et udsagn er blevet bevist, er det blevet en såkaldt "sætning" (teorem, proposition) i teorien.

2.3 Direkte beviser

De fleste matematiske sætninger er af formen: "Hvis ... så ...", altså af formen $p \Rightarrow q$ eller $p(x) \Rightarrow q(x)$. Som understreget ovenfor er der i sådan en sætning en gemt alkvantor, så sætningen er af formen $\forall x \in M : p(x) \Rightarrow q(x)$. Som vi også gjorde opmærksom på i forrige kapitel bevises en sådan sætning ved at vise, at $q(x)$ er sand for alle de x som gør $p(x)$ sand. Vi behøver altså ikke kære os om de x som gør $p(x)$ falsk, men selv da kan der jo være uendeligt mange x'er at checke. Lad os for eksempel se på sætningen:

Sætning 41 *Kvadratet på et lige tal er lige.*

Bemærkning. For at se at denne sætning er af formen $p(x) \Rightarrow q(x)$, kan vi skrive den på den mindre elegante form: "Hvis x er et lige naturligt tal, så er x^2 et lige tal". For at bevise sætningen skal vi altså checke, alle lige naturlige tal og kontrollere, at deres kvadrat er lige. Vi kunne så begynde forfra: $2^2 = 4$ er lige, $4^2 = 16$ er lige,..., men vi ville aldrig blive færdige. I stedet bruger vi fleksibiliteten i bogstavregningen, som tillader os at behandle alle lige tal på én gang. Vi antager blot, at x er et lige naturligt tal, og viser ud fra definitionen af lige tal og de kendte regneregler for naturlige tal, at x^2 er lige. Vi opererer altså med x, som om det var et bestemt tal , men er omhyggelige med kun at bruge de egenskaber, som alle lige naturlige tal har. Beviset kan forløbe således:

Bevis for Sætning (41) Lad x være et lige naturligt tal. Bestem et naturligt tal n så $x = 2n$. Dette er muligt ifølge definitionen af et lige tal. Ifølge regnereglerne for naturlige tal gælder da, at $x^2 = (2n)^2 = 2^2n^2 = 2(2n^2)$. Da $2n^2$ er et naturligt tal, er x^2 altså af formen $2m$ for et naturligt tal m og er derfor lige. QED.

Bemærkninger. Bemærk at beviset begynder med ordene "Lad x være et lige naturligt tal". Sådan begynder et typisk direkte bevis med at "lade" hypotesen være sand. Det er bedre end at starte med ordene "for alle", fordi vi efter ordet "lad" kan operere med x, som om det er et bestemt naturligt tal. Ligeså er det også bedre at skrive "bestem et naturligt tal n så $x = 2n$" end at skrive "da eksisterer et naturligt tal n så $x = 2n$", fordi vi efter at have "bestemt" n, kan operere med det som med en kendt størrelse.

Beviset slutter med bogstaverne QED. Det er en forkortelse for det latinske "quod erat demonstrandum", som betyder: hvad der skulle bevises. Det er dog også blevet almindeligt at slutte beviser med en firkant.

Bemærk også at beviset ikke bruger pile. Man kunne måske fristes til at skrive beviset på følgende vis: Da x er et lige tal, kan vi bestemme et naturligt

tal n så $x = 2n$. Derfor gælder at

$$\begin{aligned} x^2 &= (2n)^2 && (2.7)\\ &\Rightarrow x^2 = 2^2n^2 && (2.8)\\ &\Rightarrow x^2 = 2(2n^2). && (2.9) \end{aligned}$$

Men som bemærket ovenfor ville det være forkert. Implikationerne gælder jo altid og ikke *fordi* $x = 2n$. Og når man som her kun regner på den ene side af lighedstegnet, er det meget mere elegant at skrive udregningen med en række lighedstegn, som vi gjorde i beviset: $x^2 = (2n)^2 = 2^2n^2 = 2(2n^2)$.

Bemærkning: Ifølge den ovenstående forklaring af hvad et bevis er, skulle ovenstående bevis bestå af gyldige slutninger ud fra aksiomerne og de allerede beviste sætninger i teorien. Faktisk består beviset i gyldige slutninger ud fra 1. definitionen af et lige naturligt tal og 2. aksiomer og sætninger i teorien for aritmetikken for de naturlige tal. Vi har dog her kun eksplicit formuleret definitionen af et lige tal, hvorimod vi ikke har præsenteret aksiomerne og de simple sætninger for de naturlige tals aritmetik. Denne unøjagtighed vil vi også tillade os i det følgende. Det er i øvrigt karakteristisk for megen matematik, at den tager sit udgangspunkt i en ikke helt formaliseret teori.

2.4 Modeksempler.

Vi har ovenfor set hvordan man kan vise at et "hvis...så..." udsagn er sandt, altså er en sætning i en bestemt matematisk teori. Hvordan kan man da indse at et sådant udsagn er falsk? Jo, et udsagn af formen $(\forall x :)p(x) \Rightarrow q(x)$ er jo kun sandt, hvis alle x der gør hypotesen $p(x)$ sand også gør konklusionen $q(x)$ sand. Vi viser altså at udsagnet er falsk, hvis vi bare finder et eneste x, som gør $p(x)$ sand, men som gør $q(x)$ falsk. Et sådant x kaldes et **modeksempel**.

Eksempel 42 *For at modbevise udsagnet "For alle naturlige tal n er $n^2 > n$" er det nok at bemærke at* 1 *er et naturligt tal, medens udsagnet "*$1^2 > 1$*" er falsk. Tallet* 1 *er altså et modeksempel.*

2.5 Formodninger og deres behandling.

I lærebøger som denne er opgaverne ofte formuleret som: "Bevis ..." eller "Find et modeksempel mod ..." For den kreative matematiske forsker er situationen mere kompliceret. Han eller hun vil ofte have en intuitiv formodning om, at et bestemt udsagn $p(x) \Rightarrow q(x)$ er en sætning i den teori han eller hun arbejder med. For at afgøre sagen vil matematikeren først prøve at finde et bevis for udsagnets sandhed. Hvis det mislykkes, vil bevisforsøgene måske have afdækket nogle mulige modeksempler. Disse vil så blive prøvet af. Hvis de viser sig at være modeksempler, er sagen klar: udsagnet er ikke en sand sætning. Hvis det derimod viser sig, at eksemplerne alligevel ikke er modeksempler, så vil

matematikeren nok endnu en gang prøve at finde et bevis osv. Hvis denne dialektiske proces ender med et bevis, er udsagnet blevet en sætning i teorien. Hvis processen ender med et modeksempel er udsagnet falsk. Hvis det er falsk, kan matematikeren vælge at vende sig mod andre ting eller prøve at modificere udsagnet, så eksemplet ikke længere er et modeksempel. Hvis det sidste lykkes fortsætter den dialektiske afprøvningsproces med det nye udsagn. Hvis afprøvningsprocessen ender uden at der er fundet modeksempler eller beviser, må udsagnet forblive en formodning. Hvis udsagnet er særligt interessant og genstridigt kan det blive en berømt formodning som for eksempel

Goldbachs formodning: Ethvert lige tal større end to er sum af to primtal.

2.6 Bevis ved kontraposition.

Hvis man skal vise $p \Rightarrow q$, er det nogle gange simplere at vise det kontraponerede udsagn $\neg q \Rightarrow \neg p$. I så fald er man færdig, for af $\neg q \Rightarrow \neg p$ kan man slutte $p \Rightarrow q$, da disse udsagn er logisk ækvivalente (se den logiske huskeseddel).

2.6.1 Eksempler på bevis ved kontraposition.

Sætning 43 *Hvis x^2 er ulige, så er x ulige.*

Bevis. Beviset føres ved kontraposition: Lad x være lige. Vi skal da vise at x^2 er lige. Det følger af sætning (41) ■

Sætning 44 *Kvadratet på et ulige tal er ulige. Eller sagt anderledes: Hvis x er ulige, så er x^2 ulige.*

Bevisskitse. Da et ulige tal er defineret som et tal, der ikke er lige, er det nærliggende at forsøge at lave et bevis ved kontraposition, altså at bevise det kontraponerede udsagn:.

Sætning 45 *Hvis x^2 er lige, er x lige.*

I dette bevis kan man bruge følgende sætning som vi ikke vil bevise:

Sætning 46 *Hvis et primtal går op i et produkt af to naturlige tal, da går det op i et af tallene.*

Øvelse 47 *Gennemfør beviset for sætning (44) ved kontraposition.*

Alternativ bevisskitse. Man kunne fristes til at bevise sætning (44) direkte ved at bruge, at ulige tal er tal af formen $2n + 1$ for et naturligt tal n. Men da vi har defineret et ulige tal til at være et tal der ikke er lige, kan vi ikke bruge denne anden karakterisering af ulige tal, før vi har bevist, at den er ækvivalent med definitionen. Det vil vi gøre nedenfor (sætning 55)

Øvelse 48 *Gennemfør beviset for sætning (44) på grundlag af den alternative karakterisering af et ulige tal.*

Vi vil nu give et mere interessant eksempel på et bevis ved kontraposition:

Sætning 49 *For ethvert $n \in \mathbb{N}$ gælder: Hvis $2^n - 1$ er et primtal, da er n et primtal.*

Bevis. Antag, at $n \in \mathbb{N}$ ikke er et primtal. Vi vil vise, at $2^n - 1$ da heller ikke er et primtal. Såfremt $n = 1$, er dette klart, idet i så fald $2^n - 1 = 1$, som ikke er et primtal.

Vi kan altså gerne antage $n > 1$. Da n ikke er et primtal, findes en ikke-triviel faktorisering:

$$n = a \cdot b \tag{2.10}$$

hvor $a, b \in \mathbb{N}$ med $1 < a, b < n$.

Men nu verificerer man let, at vi i så fald har:

$$2^n - 1 = 2^{ab} - 1 = (2^a - 1) \cdot (2^{a(b-1)} + \ldots + 2^a + 1) \; . \tag{2.11}$$

Idet $a > 1$, er $2^a - 1 > 1$. Idet $b - 1 \geq 1$ og $a > 1$, er $2^{a(b-1)} + \ldots + 2^a + 1 > 1$. Altså har vi i (2.11) en ikke-triviel faktorisering af $2^n - 1$. Følgelig er $2^n - 1$ ikke et primtal. ■

Bemærkning. Sætningen siger altså, at for, at $2^n - 1$ er et primtal, er det en nødvendig betingelse, at n er et primtal. Med andre ord: Ønsker vi at finde primtal af form $2^n - 1$, behøver vi kun at se på primtalseksponenter n.

Der findes faktisk primtal af form $2^p - 1$ (hvor altså p så selv må være et primtal): $p = 2$, $2^p - 1 = 3$ er et eksempel. Dog er ikke alle tal af form $2^p - 1$, hvor p er et primtal, selv et primtal: $p = 11$ giver et modeksempel, idet $2^{11} - 1 = 2047 = 23 \cdot 89$.

Primtal af form $2^p - 1$ kaldes *Mersenne-primtal* efter den franske matematiker Marin Mersenne (1588–1648).

Det i skrivende stund (29/6. 2008) største, kendte primtal er et Mersenne-primtal, nemlig $2^{32,582,657} - 1$ et tal på 9,808,358 cifre. For mere information angående Mersenne-primtal, se:
http://www.mersenne.org/

2.7 Bevis ved modstrid.

Hvis man vil vise at et udsagn q er sandt, kan man gøre det ved at antage, at udsagnet er falsk (altså at $\neg q$ sand) og så vise at det fører til modstrid. Intuitivt er det jo klart, at hvis vi opnår en modstrid må vi have antaget noget falsk. $\neg q$ er altså falsk, hvorfor q er sand. Mere formelt kan bevismetoden beskrives således:

Man ønsker at bevise et udsagn q. Kan man bevise:

$$\neg q \Rightarrow (p \; \wedge \neg p) \; , \tag{2.12}$$

for et eller andet udsagn p, er man færdig: For det checkes let, at:

$$q \equiv (\neg q \Rightarrow (p \wedge \neg p)) , \tag{2.13}$$

hvorfor q kan sluttes af $\neg q \Rightarrow (p \wedge \sim p)$.

2.7.1 Eksempler på beviser ved modstrid:

Sætning 50 $\sqrt{2}$ *er irrational.*

Her skal vi bruge følgende definitioner og en sætning.

Definition 51 *Et reelt tal x kaldes **rationalt**[1], hvis der eksisterer hele tal m, n så $x = \frac{m}{n}$.*

Definition 52 *Et reelt tal kaldes **irrationalt**, hvis det ikke er rationalt.*

Definition 53 *To hele tal kaldes **indbyrdes primiske**, hvis der ikke findes et primtal som går op i dem begge.*

Sætning 54 *Hvis m, n er hele tal, så findes der to indbyrdes primiske hele tal m_1 og n_1 så $\frac{m}{n} = \frac{m_1}{n_1}$. Hvis endvidere $\frac{m}{n} > 0$ kan m_1 og n_1 vælges så de er naturlige tal. Man siger at $\frac{m_1}{n_1}$ er **uforkortelig**.*

"Bevis". Forkort $\frac{m}{n}$ så meget som muligt.
Bevis. for sætning 50. Vi viser sætningen ved et modstridsargument. Antag altså at $\sqrt{2}$ ikke er irrational, dvs. at $\sqrt{2}$ er rational. Bestem da hele tal m, n så $\sqrt{2} = \frac{m}{n}$. Ifølge Sætning 54 kan vi antage at m og n er positive og indbyrdes primiske, Det følger da af definitionen af $\sqrt{2}$ at

$$2 = (\sqrt{2})^2 = (\frac{m}{n})^2 = \frac{m^2}{n^2} \tag{2.14}$$

hvorfra vi slutter at

$$m^2 = 2n^2 \tag{2.15}$$

I følge definitionen på lige tal betyder dette, at m^2 er lige, hvorfor m ifølge Sætning 45 selv er lige. Altså findes et naturligt tal p så $m = 2p$. Men så får vi fra (2.15), at

$$2^2 p^2 = (2p)^2 = m^2 = 2n^2, \tag{2.16}$$

hvoraf

$$2p^2 = n^2. \tag{2.17}$$

Heraf ses, at n^2 er lige og derfor iflg. Sætning 45, at n er lige og altså af formen $2q$ for et naturligt tal q. Men det betyder at 2 går op i både m og n, i modstrid med at vi havde antaget, at de var indbyrdes primiske. Altså har antagelsen om at $\sqrt{2}$ var rational ført til en modstrid, og vi slutter at $\sqrt{2}$ er irrational. ■

[1] Bemærk at det *ikke* hedder rationelt!

Sætning 55 *Et helt tal x er ulige, hvis og kun hvis der findes et helt tal n så $x = 2n + 1$*

Bevis. Vi skal vise

$$x \text{ er ulige} \Leftrightarrow \exists n \in \mathbb{Z} : x = 2n + 1 \tag{2.18}$$

Vi viser først $\Rightarrow$: Dette gør vi ved et direkte bevis. Antag altså at x er ulige, dvs at det ikke er lige. Vi skal vise at x kan skrives på formen $2n + 1$ for et n i $\mathbb{Z}$. Bestem det største hele tal m så $2m < x$.[2] Da er

$$2m < x \leq 2(m + 1). \tag{2.19}$$

Men da x er ulige, er x pr. definition ikke lige, og altså ikke lig med $2(m + 1)$. Altså er

$$2m < x < 2(m + 1) = 2m + 2, \tag{2.20}$$

og da $2m + 1$ er det eneste hele tal mellem $2m$ og $2m + 2$, har vi at $x = 2m + 1$.

Dernæst viser vi $\Leftarrow$: Igen begynder vi beviset som et direkte bevis: Vi antager altså at $x = 2n + 1$ og $n \in \mathbb{Z}$ og skal vise at x er ulige. Vi skal altså vise at x ikke er lige. Det gør vi ved modstrid. Antag altså at x er lige, dvs. kan skrives $x = 2m$ for $m \in \mathbb{Z}$. Men så får vi

$$2n + 1 = x = 2m, \tag{2.21}$$

hvoraf

$$1 = 2(m - n). \tag{2.22}$$

Heraf ses at 2 går op i 1; men vi ved at 2 ikke går op i 1. Altså har vi sluttet os til en modstrid og kan derfor konkludere, at x ikke er lige, altså at x er ulige. ■

Sætning 56 *Der findes uendeligt mange primtal.*

I beviset for sætningen vil vi tillade os at bruge nogle talteoretiske sætninger og begreber, som vi ikke kommer udførligt ind på i dette kursus:

Man kan vise følgende: Hvis $n > 1$ er et naturligt tal, da findes der et primtal p, der går op i n, dvs., således, at n kan skrives $n = p \cdot m$ for et $m \in \mathbb{N}$. Videre viser man, at der *ikke* findes noget primtal, der går op i det naturlige tal 1.

Desuden skal vi bruge begrebet '(ikke tom) endelig mængde': Vi vil forstå dette begreb intuitivt som betydende, at mængdens elementer kan skrives op i en liste $a_1, \ldots, a_k$ nummereret ved de første k naturlige tal (for et eller andet $k \in \mathbb{N}$).

Bevis. Beviset føres som et modstridsbevis, så vi starter med at antage, at konklusionen er falsk, dvs. vi antager at mængden af primtal er endelig.

Mængden af primtal er dog ikke tom: eksempelvis er 2 klart et primtal. På grund af vores antagelse kan vi nu stille samtlige primtal op i en endelig liste

[2] På dette sted vil vi uden bevis bruge at et sådant største tal findes.

$p_1, p_2, ..., p_k$. Med andre ord har vi nu - på grund af vores antagelse - følgende implikation:

$$p \text{ primtal} \Rightarrow p \in \{p_1, p_2, ..., p_k\} \tag{2.23}$$

Nu, givet de endeligt mange tal $p_1, p_2, ..., p_k$ kan vi betragte deres produkt $p_1 \cdot p_2 \cdot ... \cdot p_k$, som er et naturligt tal. Vi har dermed også følgende naturlige tal:

$$N = p_1 \cdot p_2 \cdot ... \cdot p_k + 1 \tag{2.24}$$

Da klart $N > 1$, ved vi (fra talteorien), at der findes et primtal p, som går op i N. Dvs. vi kan skrive:

$$N = p \cdot m \tag{2.25}$$

med et naturligt tal m.

På den anden side: Idet p er et primtal, følger af (2.23), at p er et af tallene $p_1, p_2, ..., p_k$; dvs. $p = p_i$ for et $i \in \{1, 2, ..., k\}$. Men i så fald har vi:

$$p_1 \cdot p_2 \cdot ... \cdot p_k = p \cdot n \tag{2.26}$$

for et vist $n \in \mathbb{N}$, nemlig produktet af alle tallene $p_1 \cdot p_2 \cdot ... \cdot p_k$ på nær p_i.

Men sammenligner vi nu (2.24), (2.25) og (2.26), finder vi:

$$p \cdot n + 1 = p_1 \cdot p_2 \cdot ... \cdot p_k + 1 = N = p \cdot m \tag{2.27}$$

hvorfor $1 = p(m - n)$. Således går p op i 1, og p er derfor ikke et primtal. På grundlag af vores antagelse om at sætningen er falsk, har vi nu sluttet os til eksistensen af et naturligt tal p, således at

$$(p \text{ er primtal}) \wedge (p \text{ er ikke primtal}). \tag{2.28}$$

Da dette er en modstrid, er sætningen dermed bevist. ■

Notation 57 *Beviser ved kontraposition og ved modstrid kaldes ofte* ***indirekte beviser****.*

2.8 Beviser delt op i tilfælde.

I nogle beviser kan det være nødvendigt eller bekvemt at dele op i forskellige tilfælde.

Sætning 58 *Udsagnet* $(p \vee q) \Rightarrow r$ *er ækvivalent med udsagnet* $(p \Rightarrow r) \wedge (q \Rightarrow r)$.

Bevis. Lav en sandhedstabel. ■

Det betyder at man kan bevise $p \vee q \Rightarrow r$ ved at bevise $p \Rightarrow r$ og $q \Rightarrow r$.

Eksempel 59 *Hvis* n *er et naturligt tal, er* $n^2 + n$ *et lige tal.*

Bevis. Ifølge definition (38) er ethvert naturligt tal enten lige eller ulige. Derfor kan sætningen omformes til:

$$((n \text{ lige}) \vee (n \text{ ulige})) \Rightarrow (n^2 + n) \text{ lige}. \tag{2.29}$$

Det kan vi altså vise ved at bevise

$$((n \text{ lige}) \Rightarrow (n^2 + n) \text{ lige}) \wedge \ ((n \text{ ulige}) \Rightarrow (n^2 + n) \text{ lige}). \tag{2.30}$$

Først beviser vi at $(n \text{ lige}) \Rightarrow (n^2 + n)$ lige. Antag altså at n er lige. Ifølge Definition 37 betyder det at vi kan finde et $m \in \mathbb{N}$ så $n = 2m$ (overvej). Men så er

$$n^2 + n = (2m)^2 + 2m = 2(2m^2 + m), \tag{2.31}$$

og da $2m^2 + m \in \mathbb{N}$, er $n^2 + n$ altså et lige tal i dette tilfælde.

Dernæst beviser vi at $(n \text{ ulige}) \Rightarrow (n^2 + n)$ lige. Antag altså at n er ulige. Så findes der ifølge Sætning (55) et $m \in \mathbb{N} \cup \{0\}$ så $n = 2m + 1$. Men så er

$$\begin{aligned} n^2 + n &= (2m + 1)^2 + (2m + 1) = 2(2m^2) + 2(2m) + 1 + 2m + 1 && (2.32) \\ &= 2(2m^2 + 3m + 1), && (2.33) \end{aligned}$$

og da $2m^2 + 3m + 1 \in \mathbb{N}$ er $n^2 + n$ altså også lige i dette tilfælde.

Vi har altså bevist (2.30) og dermed sætningen. ■

Øvelse 60 *Brug den samme teknik til at bevise følgende sætning:*

Sætning 61 *Hvis n er et naturligt tal, så går 4 op i enten n^2 eller $n^2 - 1$.*

2.9 Eksistenssætninger

En særlig slags sætninger er de der udsiger eksistensen af et objekt med bestemte egenskaber, altså sætninger af formen:

$$\exists x \in M : p(x) \tag{2.34}$$

Sådanne sætninger kaldes eksistenssætninger.

Man skulle tro at eksistenssætninger var lettere at vise end sætninger af formen $p(x) \Rightarrow q(x)$, hvor man jo skal undersøge om $q(x)$ er sand ikke bare for ét x, men for alle x som opfylder $p(x)$. Almindeligvis er eksistenssætninger dog sværere at vise end universelle udsagn i den forstand, at de kræver mere kreativitet. Et eksistensbevis falder nemlig oftest i to dele. Først finder man en kandidat x_0, og dernæst beviser man at $p(x_0)$ er sand. Den sidste del af beviset går ofte ret let. Her følges logikkens sædvanlige regler. Derimod er der ingen faste regler for, hvordan man finder en kandidat. Det er aldeles ligegyldigt, hvordan det sker. Det kan ske ved et inspireret gæt eller ved en mere systematisk undersøgelse. Og man behøver i beviset ikke fortælle, hvordan kandidaten er fundet. Så længe man efterfølgende kan vise at kandidaten x_0 opfylder det ønskede ($p(x_0)$), er beviset i hus.

2.9.1 Eksempler på eksistenssætninger

Sætning 62 *Der findes et naturligt tal x så $x = x^2$.*

Bevis. Betragt tallet 1. Da 1 er et naturligt tal og $1 = 1^2$, er sætningen vist. ■

Sætning 63 *Der findes et lige primtal.*

Bevis. Betragt tallet 2. Det er et primtal, og da $2 = 2 \cdot 1$ er det også lige. ■

Sætning 64 *Der eksisterer en rational rod i polynomiet $P(x) = x^4 - 2x^3 - 3x^2 + 5x + 2$.*

Bevis. Da 2 er et rationalt tal, og $P(2) = 2^4 - 2 \cdot 2^3 - 3 \cdot 2^2 + 5 \cdot 2 + 2 = 0$ er 2 en rational rod i P. ■

Bemærkning. Hvor universelle udsagn modbevises ved at finde et modeksempel (altså ved at løse et eksistensproblem), vil eksistensudsagn modbevises ved at bevise et universelt udsagn. Negationen af udsagnet $\exists x \in M : p(x)$ er jo udsagnet $\forall x \in M : \neg p(x)$.

Bemærkning: En eksistenssætning $\exists x \in M : p(x)$ udtaler sig ikke om, hvor mange elementer i M, der opfylder $p(x)$. Den siger blot at der mindst er et. Hvis man vil vise, at der højst er et, skal man vise en entydighedssætning.

Bemærkning. De beviser vi har omtalt ovenfor er *konstruktive* i den forstand, at de ikke bare fortæller at der findes et objekt med de ønskede egenskaber, men også angiver et sådant objekt. Eksistenssætninger kan dog også bevises ikke-konstruktivt.

2.9.2 Ikke-konstruktive beviser

Hvis et eksistensbevis godtgør at der eksisterer et objekt med givne egenskaber uden at fortælle, hvilket objekt der er tale om, så siges beviset at være ikke-konstruktivt.

Eksempel 65 *Ligningen*

$$x^3 + 2x^2 + x + 7 = 0 \tag{2.35}$$

har en reel løsning.

Bevis. Polynomiet $P(x) = x^3+2x^2+x+7$ er en kontinuert funktion af x på hele den reelle akse. Da $P(1) > 0$ og $P(-10) < 0$ ved vi fra mellemværdisætningen (Skjæringssætningen 5.2.1 i Lindstrøm), at der findes et reelt tal $x_0 \in]-10, 1[$, så $P(x_0) = 0$. Dette x_0 er altså en rod i ligningen. ■

Eksempel 66 *Der findes et naturligt tal n som er større end e^{100}.*

Bevis. Vi vil føre beviset ved modstrid. Antag derfor, at alle naturlige tal er mindre eller lig med e^{100}. Vi vil senere vise, at det fører til modstrid, så vi slutter, at der er et naturligt tal større end e^{100}. ■

Beviserne i de to sidste eksempler er ikke-konstruktive, idet de ikke fortæller hvilket tal der opfylder det ønskede.

Man kan også lave ikke-konstruktive eksistensbeviser ved at argumentere indirekte: Betragt eksistenssætningen $\exists x \in M : p(x)$ (hvor $p(x)$ er et prædikat i variablen $x \in M$). Vi kan da indirekte bevise, at sætningen er sand, idet vi beviser, at dens negation $\forall x \in M : \neg p(x)$ fører til modstrid. I dette tilfælde giver beviset altså et eksistensbevis uden at give en ide til, hvordan vi finder et objekt x, som opfylder det ønskede $p(x)$. Normalt foretrækker man konstruktive eksistensbeviser netop fordi man så ved, hvordan man finder det eksisterende objekt.

2.10 Entydighedssætninger

Entydighedssætninger er sætninger der udsiger, at der højst er et objekt med bestemte egenskaber, altså sætninger af formen:

$$\text{Der findes højst ét } x \in M \text{ så } p(x) \tag{2.36}$$

En sådan entydighedssætning kan vises ved at udlede nogle konsekvenser fra $p(x)$ som fastlægger x entydigt.

Bemærkning 67 *En anden meget brugt metode til at bevise at der højst findes ét $x \in M$ så $p(x)$, er at antage, at x og y begge har egenskaben p og derfra udlede at $x = y$. Udtrykt mere formelt bevises entydigheden ved at vise at for $x, y \in M$ gælder:*

$$(p(x) \wedge p(y)) \Rightarrow (x = y). \tag{2.37}$$

2.10.1 Eksempel på entydighedssætning

Sætning 68 *Hvis q er et positivt rationalt tal da findes højst ét positivt rationalt tal x så*

$$x^2 + x = q. \tag{2.38}$$

Bevis. Antag at x og y er to positive rationale tal, som opfylder at

$$x^2 + x = y^2 + y = q. \tag{2.39}$$

Det følger af aksiomerne for regning med de rationale tal, at funktionen x^2+x er voksende på de positive rationale tal (dette vil vi ikke bevise). Fra $x^2+x = y^2+y$ kan vi derfor slutte at $x = y$. ■

Bemærk at sætningen i lighed med alle andre entydighedssætninger ikke udtaler sig om, hvorvidt der *eksisterer* positive rationale løsninger til ligningen (2.38). Entydighedssætningen siger blot, at *hvis* der findes en positiv rational løsning, er der kun én. Hvis $q = 2$ er det let at se at $x = 1$ er en positiv rational løsning, og sætningen siger derfor at der ikke er andre. Hvis $q = 1$, findes der derimod ingen rationale løsninger (prøv bare at løse ligningen $x^2 + x = 1$).

2.11 Eksistens og entydighed

Mange sætninger udsiger både eksistens og entydighed. De begynder ofte med ordene "der findes én og kun én..." eller "der findes netop én..."

Sætning 69 *Der findes netop en positiv rod i ligningen* $x^2 + x = 2$.

Bevis. Da $1^2 + 1 = 2$ er 1 en positiv rod i ligningen, og da $x^2 + x$ er voksende for $x > 0$ er 1 den eneste positive rod. ■

2.12 Opgaver

1.Afgør om følgende udsagn er sande eller falske (giv bevis eller modeksempel):

1. $a^2 + b^2 = (a + b)^2$ for $a, b \in \mathbb{R}$
2. Der eksisterer naturlige tal x og y, så at $5x + 2y = 27$

2. Lad x og y være positive reelle tal. Vis at

$$\frac{x}{y} + \frac{y}{x} \leq 2 \Rightarrow (x - y)^2 \leq 0 \tag{2.40}$$

Brug dette til at give et modstridsbevis for følgende sætning:
Hvis x og y er forskellige positive reelle tal, da er

$$\frac{x}{y} + \frac{y}{x} > 2 \tag{2.41}$$

3. Lad x og y være hele tal. Bevis følgende sætning:

$$x \cdot y \text{ er ulige, hvis og kun hvis } x \text{ er ulige og } y \text{ er ulige} \tag{2.42}$$

I beviset må du gerne bruge Sætning 55.
4. a. Bevis at $\sqrt{3}$ er irrational
b. Bevis at $\sqrt{6}$ er irrational
c. Gælder der følgende sætning?: Summen af to irrationale reelle tal er irrational.
d. Afgør om $\sqrt{2} + \sqrt{3}$ er irrational.
I beviserne må du gerne bruge Sætning 46
5. Bevis at for et vilkårligt helt tal n er $n^2 - 5n + 7$ ulige.

Chapter 3

Euklids aksiomatisering af geometrien

Historiens første overleverede aksiomatiske system er Euklid's Elementer, skrevet omkring 300 f.Kr. Indtil ca. 1850 blev dette værk betragtet som mønstret på eksakt matematisk fremstilling. Vi skal gennemgå de indledende definitioner og aksiomer fra den første af Elementernes 13 bøger samt de første 6 sætninger. Formålet er at opøve geometrisk bevisteknik, og dermed træne matematisk eksakthed. Det er ikke formålet at undersøge de historiske spørgsmål, som værket rejser. De diskuteres i kurset i matematikkens historie.

Vi skal fremhæve Euklids meget omhyggelige beviser, men også påpege en del mangler og problemer i fremstillingen. Disse mangler blev opdaget i slutningen af 1800-tallet. De blev udbedret af David Hilbert i *Grundlagen der Geometrie* (første udgave 1899). Vi har her foretrukket Euklid frem for Hilbert, dels af kulturhistoriske grunde, dels for at opøve læserens kritiske sans, og dels fordi Hilberts fuldstændig stringente fremstilling er mere omstændelig end de fleste har lyst til at være.

Bog 1 af Euklids Elementer til og med sætning I.6 er gengivet på kursushjemmesiden i Taisbak et al's nye danske oversættelse.

3.1 Bemærkninger til definitionerne.

Det ser ud som om Euklid vil definere alt. Men det er et håbløst forehavende. Se f.eks. på definitionen af en ret linje. Det er en linje, som ligger lige mellem sine punkter. Men så kan man med rette spørge, hvad det betyder at "ligge lige". Det kan jo ikke så godt betyde, at de ligger på en ret linje, for så er definitionen cirkulær. Det er klart at for at definere hvad et ord betyder skal man bruge andre ord, hvis betydning man så atter skal definere ved endnu andre ord. Hilbert valgte at kortslutte denne uendelige regres ved at operere med udefinerede ord, som for eksempel punkt, linje, plan, ligge på, kongruent med osv. hvorfra definitionsprocessen kan tage sit udgangspunkt. Disse ord er

hos Hilbert kun fastlagt ved, at de skal opfylde aksiomerne. Vi må derfor i dag se kritisk på Euklids indledende definitioner.

Efter Euklids første kritisable definitioner kommer dog en række præcise definitioner af begreber i termer af de tidligere begreber. Bemærk for eksempel definition 10 af en ret vinkel, definition 15 af en cirkel, definition 20 af en ligesidet trekant og definition 23 af parallelle linjer.

3.2 Bemærkninger til postulaterne og de almindelige begreber.

Hvor Euklid tilsyneladende prøver at definere alt, har han indset at man ikke kan bevise alt. Han starter bevisprocessen med en række ubeviste sætninger, de sætninger vi vil kalde **aksiomer**. Euklid har delt aksiomerne op i to grupper: Først postulaterne, som er de geometriske aksiomer, og dernæst de almindelige begreber, som for størstedelens vedkommende er mere almindelige slutningsregler. Vi skal ikke hænge os i denne opdeling, og skal opfatte både postulater og almindelige begreber som aksiomer, altså som sætninger, som ligger til grund for teorien, men som ikke bevises.

De første tre postulater kaldes ofte konstruktionspostulaterne. De postulerer, at man i bestemte situationer kan konstruere bestemte linjer og cirkler. Man siger ofte at Euklid konstruerer med passer og lineal. De konstruktioner, som beskrives i de tre første postulater, vil man jo lave med disse instrumenter. Men Euklid nævner ikke instrumenterne med ét ord. Det eneste vigtige for ham er at de betragtede linjer og cirkler kan konstrueres. Man kan opfatte dette som eksistenspostulater. For selv om Euklid i definitionerne definerede rette linjer, cirkler, rette vinkler osv. så ligger der intet eksistenspostulat gemt i definitionerne. Postulaterne siger netop, at visse linjer og cirkler eksisterer. Eksistensen af andre figurer som for eksempel en ret vinkel, en ligesidet trekant, et kvadrat osv. må så vises ved at figurerne konstrueres ud fra postulaterne for linje og cirkel.

Det femte postulat er det berømte parallelpostulat.

Det første almindelige begreb siger, med en sprogbrug vi vil indføre senere, at lighedsrelationen er transitiv, altså at $(a = b) \wedge (b = c) \Rightarrow a = c$. Vi bemærker her at $=$ ikke betyder, at de to størrelser er ens. Det betyder bare at de er lige store.

Det andet og tredje almindelige begreb kan sammenfattes til: $(a = b) \wedge (c = d) \Rightarrow a \pm c = b \pm d$.

I de fleste udgaver af Elementerne er de almindelige begreber, som her har numrene $5, 6, 8, 9$, normalt udeladt.

3.3 Bemærkninger til sætningerne

Euklids Elementer indeholder to slags sætninger: Konstruktionsproblemerne og de almindelige sætninger. Af de her gengivne sætninger[1] er I.1, I.2, og I.3 konstruktionsproblemer, medens I.4, I.5, I.6 er almindelige sætninger. Sætningers beviser afsluttes med ordene "Hvad der skulle bevises" (QED, quod erat demonstrandum), medens konstruktionernes beviser afsluttes med ordene "Hvad der skulle gøres" (QEF, quod erat faciendum).

Euklids sætninger præsenteres efter en fast skabelon, som i den ny danske oversættelse er fremhævet i forbindelse med den første sætning:

[*i*]: den generelle formulering af sætningen eller opgaven.

[*ii*]: der sættes bogstaver på sætningens eller opgavens givne.

[*iii*]: der sættes bogstaver på det der skal vises eller konstrueres-

[*iv*]: konstruktion; dette afsnit findes kun i konstruktionsproblemerne.

[*v*]: bevis.

[*vi*]: sammenfatning eller konklusion.

I den gengivne oversættelse er der indsat firkantede parenteser med henvisninger til de tidligere definitioner, aksiomer og sætninger, som bruges i konstruktionerne og beviserne. Henvisningerne stod ikke eksplicit i Euklids originale tekst, men da det netop er disse henvisninger, som viser den deduktive struktur i Elementerne, bør du under læsningen være meget opmærksom på dem. Jeg vil kun undtagelsesvist kommentere henvisningerne i det følgende.

Sætning I.1. I definition 20 defineres, hvad en ligesidet trekant er. Her konstrueres den, og dermed vises dens eksistens. Bemærk, hvordan Euklid først bruger postulat 3 til at tegne de to cirkler, og dernæst bruger definition 15 til at slutte at radierne i hver cirkel er lige store. Han har ganske vist ikke fået defineret "radius", men det er klart, at det skal defineres som de rette linjer, som nævnes i definition 15. Vi bemærker også hans omhyggelige (nogen vil måske sige pedantiske) brug af lighedsrelationens refleksivitet (A.1) til at vise, at $AC = CB$.

Der er dog et væsentligt hul i konstruktionen. Det drejer sig om konstruktionen af punktet C som skæringspunktet mellem de to cirkler. Et trivielt problem er, at Euklid ikke har defineret et "skæringspunkt", men det skal klart betyde et punkt, der ligger på begge cirkler. Hovedproblemet er, om de to cirkler overhovedet har et skæringspunkt. Det kan man ikke definere sig fra. Her kræves et postulat (aksiom), men et sådant postulat finder man ikke hos Euklid. Det eneste skæringspostulat i Elementerne er det femte, som postulerer at to linjer skærer hinanden under visse forudsætninger, men der er ingen skæringspostulater for cirkler. Skæringspostulatet kunne lyde: "Når linjestykket mellem to cirklers centre er mindre end de to radier tilsammen og større end forskellen mellem de to radier, så skærer cirklerne hinanden". Alternativt kunne man formulere en kontinuitetssætning, som siger noget i stil med, at hvis en kurve (linje i Euklids sprogbrug) går fra det indre til det ydre af en cirkel, må den skære

[1] Når man henviser til sætningerne i Euklids elementer, sker det normalt ved angivelse af bogens nummer med romertal, efterfulgt af sætningens nummer skrevet med arabertal.

cirklen (her kræves så en definition af indre og ydre). Dette kontinuitetsaksiom ville udelukke at cirklen (og kurven) har "huller", sådan som de rationale tal har huller, når de anbringes på den reelle tallinje. (herom senere).

Sætning I.2 og I.3. Disse konstruktioner viser, hvordan man kan trække et linjestykke fra et større. Men hvorfor er det gjort så besværligt? Det ville være meget simplere (med betegnelserne i sætning I.3) at bruge postulat 3 til at tegne en cirkel med centrum i A og radius lig C. Dette har Euklid naturligvis vidst, så der må være en grund til at han foretrækker den mere besværlige konstruktion. Grunden ligger i fortolkningen af postulat 3. Vi bemærker nemlig, at Euklid i konstruktionerne i de tre første sætninger kun bruger postulatet til at konstruere cirkler med radius afsat ud fra centrum. Det ser altså ud til, at man skal læse postulatet på følgende snævre måde: "man kan tegne en cirkel med et hvilket som helst centrum og en hvilket som helst radius afsat ud fra centrum". Hvis vi skal læse postulatet sådan, duer den simplere konstruktion ikke. Det kræves at radien C transporteres hen til A, og det er jo netop det som sker i sætning I.2.

Man siger somme tider, at Euklid konstruerer med sammenklappende passer, altså en passer som klapper sammen, når den løftes fra papiret. Det betyder jo netop, at man kun kan tegne cirkler med radius afsat ud fra centrum. Men Euklids snævre læsning af postulat 3 er naturligvis ikke et resultat af, at antikke græske passere var underlige, men et resultat af en bevidst stræben efter at gøre aksiomerne så få og så svage som muligt. Vi bemærker i øvrigt at Euklid i sætning I.3 netop får bevist, at man selv med den svage læsning af postulat 3 kan tegne en cirkel med et givet centrum og en radius liggende et andet sted. Han viser altså at den svage læsning medfører den stærke.

Sætning I.4. Dette er den første kongruenssætning. Man siger at to trekanter er kongruente, hvis de har alle sider og vinkler parvist lige store (og Euklid tilføjer også at trekanterne selv har samme størrelse). Der er i alt fire kongruenssætninger i Elementerne. De siger alle, at når tre sider eller vinkler i en trekant er parvist lige så store som de tre tilsvarende sider eller vinkler i den anden trekant, så er de to trekanter kongruente. Sætning I.4 er den såkaldte SVS (side-vinkel-side) kongruenssætning. Den siger at to trekanter er kongruente, hvis de har en vinkel og de to hosliggende sider parvist lige store.

Euklids bevis kan synes meget overbevisende, men er dog basalt råddent. Euklid "anbringer" den ene trekant oven på den anden. Men hvis trekanterne ikke i forvejen ligger oven i hinanden, kræver det en flytning fra et sted i planen til et andet sted. Sådanne flytninger er slet ikke omtalt i aksiomerne og hører derfor ikke hjemme i Euklids opbygning af geometrien. Man kan godt bygge geometrien på flytningsbegrebet, men så skal der nogle flytnings-aksiomer til. Uden ekstra aksiomer kunne det ske, at når man forsøger at flytte en trekant fra et sted til et andet, så ville den være tvunget til at skifte form. Betragt for eksempel en flade som er plan et sted, medens den har en kugleformet bakke et andet sted. Hvis man da ville prøve at flytte en trekant fra den kugleformede bakke hen på den plane del af fladen, ville det ikke kunne lade sig gøre uden at ændre på nogle af vinklerne eller siderne. Hvis man for eksempel som Euklid ville prøve at bevare vinkel A og siderne AB og AC, så ville siden BC nødvendigvis

skulle gøres større, for at få fladet trekanten ud.

For at undgå sådanne problemer skal man til Euklids aksiomer tilføje et aksiom, der sikrer, at planen er "lige flad" over det hele. Den simpleste måde at opnå det på er simpelthen at gøre sætning I.4 (eller en del af den) til et aksiom. Hilbert har således medtaget følgende aksiom:

Hvis to trekanter har en vinkel og dens to hosliggende sider parvis lige store, så er de to andre vinkler også parvis lige store.

Derfra kan han så *bevise* at også det sidste par sider er parvist lige store.

Sætning I.5. Figuren i denne sætning har fået sit eget navn: æselbroen.

Man kan vise sætningen på en simplere måde: Betragt den oprindelige trekant som to trekanter BAC og CAB. Disse har $\angle BAC = \angle CAB$, $BA = CA$, og $CA = BA$. Ifølge kongruenssætningen I.4 er da også $\angle ABC = \angle ACB$.

Ud over den pædagogiske vanskelighed, der ligger i at opfatte den samme trekant som to trekanter, er der dog et problem ved dette bevis. Vi får jo ikke umiddelbart resultatet om at de udvendige vinkler ved grundlinjen er lige store. Man kunne tro at beviset for dette resultat var en simpel konsekvens af, at vinklerne ved grundlinjen er lige store. Man kunne vel bare bemærke at

$$\angle DBC + \angle ABC = 2R = \angle ECB + \angle ACB \tag{3.1}$$

hvor R står for en ret vinkel. Ved fra denne lighed at fratrække de to lige store vinkler $\angle ABC$ og $\angle ACB$, får man jo fra A.3 at resterne $\angle DBC$ og $\angle ECB$ er lige store.

Men dette bevis er ikke korrekt. For hvordan ved vi at $\angle DBC + \angle ABC = 2R$? Med andre ord, hvordan ved vi at når en ret linje DA går gennem et punkt B, så vil vinklen ved punktet mellem de to modsatrettede halvlinjer BA og BD være to rette? Man kunne tro at det var en konsekvens af definitionen af en ret vinkel. Problemet er dog at for at vi kan vide at der ved punktet B ligger to rette vinkler, skal der gennem B være oprejst en linje på DA, så de to vinkler, der ligger ved siden af hinanden er lige store. Det kræver definitionen jo. Men denne rette linjes eksistens, kan vi ikke tage for givet. Vi skal først konstruere den. Konstruktionen af den vinkelrette på en ret linje gennem et punkt på linjen (oprejsning af den vinkelrette) kommer dog først i sætning I.11 i Elementerne, og beviset for konstruktionens rigtighed bruger netop sætning I.4. Derfor ville det være en cirkelslutning, hvis vi uden videre brugte, at der ved et punkt på en ret linje er to rette vinkler mellem de to modsatrettede halvlinjer.

Euklids bevis undgår på elegant vis denne cirkelslutning.

Sætning I.6. Denne sætning er den omvendte af Sætning I.5:

$$\begin{array}{ll} \text{Sætning I.5} & (AB = AC) \Rightarrow (\angle ABC = \angle ACB) \\ \text{Sætning I.6} & (\angle ABC = \angle ACB) \Rightarrow (AB = AC) \end{array} \tag{3.2}$$

Sætning I.6 vises indirekte ved et modstridsbevis. Modstriden opnås ved at konstruere en trekant, som på den ene side er en del af den oprindelige trekant og derfor iflg aksiom 8 er mindre end den, men som kan bevises at være kongruent med, og dermed lige så stor som den oprindelige. Her mangler Euklid dog et

argument for, at den konstruerede trekant er en del af den oprindelige. Det ser klart nok ud på tegningen, men der er ikke aksiomer nok til at sikre det.

Et mere tilfredsstillende bevis kunne have ført ved kontraposition. Prøv at gennemfør dette.

3.4 Resten af Euklids Elementer

Euklid fortsætter med at deducere mere og mere komplekse sætninger. Bog 1 slutter med Pythagoras' læresætning og dens omvendte sætning:

Sætning I.47. I en retvinklet trekant er kvadratet på den side, der ligger over for den rette vinkel, lig summen af kvadraterne på de sider, der indeslutter den rette vinkel.

Sætning I.48. Hvis kvadratet på en af siderne i en trekant er lig summen af kvadraterne på trekantens to andre sider, er den vinkel, der indesluttes af trekantens to andre sider, ret.

Elementerne består af 13 bøger. I den sidste konstrueres de 5 regulære polyedere.

Chapter 4

Analyse og Syntese.

4.1 Matematisk kreativitet

Videnskabsteoretikere skelner mellem en "context of justification" og en "context of discovery". Den første handler om, hvordan man argumenterer for allerede indhøstet viden, den anden om, hvordan man opdager eller skaber ny viden. I matematik argumenterer man for sin viden ved hjælp af de meget formaliserede og stringente beviser, som MatM især handler om. Men ny viden indhøstes sjældent på denne måde. I princippet kunne man fodre en computer med alle en teoris aksiomer og alle de tilladte slutningsregler, og så sætte den til at bevise nye sætninger. Det vil den også kunne gøre, men langt størstedelen af de sætninger maskinen vil udlede, ville vi mennesker opfatte som aldeles uinteressante. Hvis man for eksempel bad en computer (eller en fantasiløs person) om at deducere sætninger om de naturlige tal, kunne den måske gå i gang med at bevise at $1 + (1 + 1) = (1 + 1) + 1$, $1 + (1 + (1 + 1)) = ((1 + 1) + 1) + 1$, og så videre. Computeren ville dermed have udstukket et "forskningsprojekt", som til dommedag ville blive ved med at spytte sande sætninger ud. Men ikke én af disse sætninger ville forekomme os at være interessante.

Dette eksempel viser, hvad der sker, hvis man på må og få deducerer sætninger fra aksiomerne. Problemet er at vi ikke kan formalisere, hvilke sætninger der er interessante og hvilke der er uinteressante. Det er her matematikere af kød og blod kommer ind. Matematikeren bruger sin intuition, sin fantasi, analogier mm. til at udtænke sætninger, som er interessante. Om de så er sande, må man checke ved formelle beviser. Men også udtænkningen af disse beviser kræver fantasi og intuition.

Mange matematikere, psykologer og filosoffer har diskuteret, hvordan matematikere får ideer til nye resultater og deres beviser. I princippet er alt tilladt i denne kreative fase af matematikken. Så længe den kreative proces ender med et stringent bevis for sætningen, er vi tilfredse. Matematisk kreativitet er lige så uforklarlig som kreativitet i andre områder af menneskelivet og er derfor svær at undervise i. (Se dog for eksempel Polya's bog "How to solve it"). Der er

dog visse heuristiske metoder, som man kan lære sig. Dette kapitel handler om en af dem, nemlig den matematiske analyse. Den kommer i to udgaver: en til brug ved problemløsning, og en anden, som man bruger, når man skal finde på beviser for sætninger. *Begge er kendetegnet ved at de tager udgangspunkt i det søgte, ubekendte eller ubeviste og arbejder sig tilbage til det kendte eller beviste.*

4.2 Eksistensproblemer

Problem 70 *Undersøg om polynomiet* $P(x) = x^4 - 2x^3 - 3x^2 + 5x + 2$ *har rationale rødder, og i så fald hvilke.*

Her er altså tale om et eksistensproblem, efterfulgt af et problem om at bestemme samtlige objekter, som opfylder det ønskede. I forrige afsnit beviste vi at 2 var en rational rod i polynomiet. Men vi afslørede ikke, hvordan tallet 2 var fremkommet. Vi trak det bare op af hatten. Her skal vi se, hvordan man kan komme frem til denne kandidat.

Man kunne måske tro, at problemet mest effektivt løses ved at bruge en formel for løsningen af ligningen. Det er imidlertid ikke tilfældet. Ganske vidst findes der en formel for løsningen af en fjerdegradsligning. Men den er meget kompliceret, og det vil være meget svært ad den vej at bestemme om rødderne er rationale tal. Mere metodologisk kan man også sige at det vil være at skyde gråspurve med kanoner hvis man brugte en formel som frembringer samtlige fire komplekse rødder, når nu spørgsmålet kun handler om rationale løsninger.

Lad os i stedet begynde med eksistensproblemet: Findes der et rationalt tal, som er rod i $P(x)$? Hvis vi kan finde et rationalt tal p/q så $P(p/q) = 0$, så har vi besvaret spørgsmålet med ja. Beviset går da bare ud på at indsætte p/q i $P(x)$ og vise, at resultatet er nul. Men hvordan skal vi finde en kandidat p/q?

Analyse.

Det gøres mest effektivt ved at lave en *analyse* af problemet. En analyse går ud på, at vi *antager,* vi har bestemt et objekt, der har den ønskede egenskab (en løsning til problemet), og så ser vi, hvad vi kan udlede om dette objekt. I det forelagte problem antager vi, at p/q er en rod i polynomiet og undersøger, om vi deraf kan slutte os til noget om p og q.

Vi vil først gå lidt mere generelt til værks, idet vi antager at p/q er en rod i polynomiet $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ med heltallige koefficienter. Vi antager altså at

$$a_n \frac{p^n}{q^n} + a_{n-1}\frac{p^{n-1}}{q^{n-1}} + \cdots + a_1 \frac{p}{q} + a_0 = 0. \tag{4.1}$$

Hvis vi ganger igennem med q^n, ser vi at

$$a_n p^n + a_{n-1}p^{n-1}q^1 + \cdots + a_1 p q^{n-1} + a_0 q^n = 0 \tag{4.2}$$

Ifølge sætning 54 kan vi endvidere antage, at p/q er forkortet mest muligt så de to hele tal p og q er indbyrdes primiske (dvs. at de ikke har nogen fælles hele

divisorer ud over 1 og -1). Hvis vi nu isolerer første led i ligningen fås:

$$a_n p^n = -a_{n-1}p^{n-1}q^1 - \cdots - a_1 pq^{n-1} - a_0 q^n. \tag{4.3}$$

Da q går op i alle led på højresiden, og derfor i hele højresiden, går det også op i venstresiden $a_n p^n$. Men da q er indbyrdes primisk med p og derfor med p^n, må q gå op i a_n. Her har vi brugt, at hvis et helt tal q går op i et produkt rs af to hele tal, og det er indbyrdes primisk med det ene tal r, så må det gå op i det andet tal s (Sætning 46).

Ved at isolere $a_0 q^n$ i ligning (4.2) kan vi på helt samme måde se, at p må gå op i a_0. Vi har dermed vist følgende sætning:

Sætning 71 *Hvis en uforkortelig brøk p/q ($p, q \in \mathbb{Z}$) er rod i polynomiet $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ med heltallige koefficienter, så går p op i a_0, og q går op i a_n*

Hvis vi nu betragter det konkrete polynomium $P(x) = x^4 - 2x^3 - 3x^2 + 5x + 2$ så siger sætningen, at hvis den uforkortelige brøk p/q er rod i polynomiet, så vil p gå op i 2, og q vil gå op i 1. Det betyder at p er en af tallene 1, -1, 2, -2, og q er et af tallene 1, -1. Derfor må p/q være et af tallene 1, -1, 2, -2.

Hvad er det vi nu har vist? Har vi vist at de fire tal 1, -1, 2, -2 er rødder i polynomiet, eller at et af tallene er rod i polynomiet? Nej, vi har ikke vist nogen af delene. Det vi har vist er, at *hvis* polynomiet har rationale rødder, skal de findes blandt disse fire tal. Vi har med andre ord fundet fire kandidater til eksistensproblemet. Hermed er analysen slut.

Syntese.

Når vi har kandidaterne kan vi gå i gang med at bevise, at de opfylder det ønskede, eller rettere undersøge *om* de opfylder det ønskede. Det kalder man ofte *syntesen*.

I det forelagte problem skal vi bare indsætte de fire tal 1, -1, 2, -2 i ligningen, det ene efter det andet, og undersøge om ligningen er opfyldt. Vi ser let at $P(1) = 3$, $P(-1) = -3$, $P(2) = 0$ og $P(-2) = 12$. Altså har vi bevist, at der eksisterer en rational rod i polynomiet $P(x)$.

Entydigheden.

Hvis vi bare var faldet over kandidaten 2 ved at prøve os frem eller ved et inspireret gæt, ville vi også ved indsættelse have kunnet bevise, at den var en rod. Men analysen har faktisk givet os meget mere information. Vi sluttede jo fra analysen, at *hvis* p/q var en rational rod i polynomiet, så måtte p/q være et af de fire tal 1, -1, 2, -2. Vi ved altså, at der ikke eksisterer andre rationale rødder end disse fire. Da vi derefter ved simpel indsættelse i polynomiet konstaterede, at de tre tal 1, -1, -2 ikke var rødder, medens 2 var en rod, kan vi nu slutte at 2 er den *eneste* rationale rod i ligningen. Ud over at give os en kandidat til eksistensspørgsmålet har analysen altså givet os et entydighedsbevis. Og vi har endog fået fundet den entydigt eksisterende rod. Dermed har vi fået løst hele Problem 70.

4.3 Analyse og syntese

Lad os rekapitulere, hvad analyse-syntese metoden går ud på. Vi ønsker at vise, at der eksisterer et objekt med en bestemt egenskab. Vi antager at vi kender objektet og undersøger, hvad vi kan sige om det. Hvis det er et algebraisk objekt, giver vi det et navn som et bogstav (her kaldte vi det p/q), og vi regner med det som om det var kendt. Hvis det er et geometrisk objekt, tegner vi det ind på figuren og deducerer, som om det var kendt (vi laver en prøvefigur, se nedenfor). I heldige tilfælde kan vi slutte os til, at hvis objektet findes med de givne egenskaber, så er det entydigt bestemt. Hvis det sker har vi allerede vist entydigheden af objektet, i den forstand at hvis der overhovedet findes et objekt med de ønskede egenskaber, så er der kun et, nemlig det objekt (den kandidat), som kom ud af analysen. For at vise eksistensen skal vi bevise at (eller om) kandidaten har den ønskede egenskab. Hvis den har egenskaben, har vi vist eksistensen (og eftervisningen kaldes det syntetiske bevis eller syntesen); hvis den ikke har egenskaben, har vi vist, at der ikke eksisterer objekter med den givne egenskab.

I andre tilfælde (som i ovenstående eksempel) fører analysen ikke til en entydig karakterisation af det søgte objekt, men til en overskuelig mængde kandidater (i ovenstående eksempel fire kandidater). Man må da efterprøve, om kandidaterne har egenskaben. Hvis ingen har egenskaben kan vi slutte, at der ikke eksisterer løsninger til problemet. Hvis nogle af kandidaterne opfylder egenskaben, så udgør beviset herfor det syntetiske eksistensbevis. Desuden har vi vist, at der ikke findes andre løsninger på problemet. Hvis der kun er én af kandidaterne, som opfylder egenskaben, giver analysen entydigheden af løsningen. Hvis der er flere, fås ikke entydighed, men vi har fundet hele mængden af løsninger.

Øvelse 72 *I forrige kapitel beviste vi sætningen: Der findes et naturligt tal x så $x^2 = x$. Lav en analyse, og undersøg, om det fører til en entydig kandidat. Gennemfør syntesen, og formuler den sætning, du kan konkludere ud fra undersøgelsen.*

4.4 Ligningsløsning. Et eksempel på analyse - syntese

Når man løser ligninger benytter man analyse - syntese metoden. Lad os se på et eksempel:

Problem 73 *Find et kvadrat, som lagt sammen med sin side giver 2.*

Med andre ord: Vi ønsker at finde ud af, om der eksisterer sådanne kvadrater, og hvis der findes nogen, skal vi angive dem. Vi fortolker opgaven sådan, at det vi skal finde, er længden af det ønskede kvadrats side.

Hvis vi kun var interesseret i eksistensspørgsmålet kunne et bevis se således ud:

Da $1^2 + 1 = 2$ har et kvadrat med siden 1 den ønskede egenskab.

Men hvis vi også vil forklare, hvor ettallet kom fra (eller vi ikke kan gætte det), og vi også ønsker at undersøge om der findes andre løsninger, kan vi ty til en *analyse* af problemet:

Vi *antager* derfor, at vi kender et kvadrat, som har den givne egenskab, og vi kalder dens sidelængde x. Nu opererer vi med x som om det var et kendt tal. Da kvadratet antages at opfylde egenskaben i problemet, gælder der, at

$$x^2 + x = 2 \tag{4.4}$$

Hvis vi lægger $\frac{1}{4}$ til på begge sider, får vi, at x må opfylde at

$$x^2 + x + \frac{1}{4} = (x + \frac{1}{2})^2 = 2\frac{1}{4}, \tag{4.5}$$

og dermed, at

$$x + \frac{1}{2} = \pm\sqrt{2\frac{1}{4}} = \pm 1\frac{1}{2}, \tag{4.6}$$

så

$$x = 1 \text{ eller } x = -2. \tag{4.7}$$

Vi har nu fundet ud af at hvis x er siden i et kvadrat, som lagt sammen med sin side er lig med 2, så må x være et af tallene 1 eller -2. Det færdiggør analysen.

Nu skal vi vise om de to tal virkeligt opfylder betingelsen:

Da 1 er længden af siden i et kvadrat, og $1^2 + 1 = 2$, er 1 en løsning på problemet. Tallet -2 derimod, kan ikke være længde af et kvadrats side så -2 er ikke en løsning på problemet. Altså er 1 den eneste løsning af problemet. Vi har dermed vist en eksistens og entydighedssætning.

Faktisk er det netop dette vi gør når vi løser ligninger. Vi kalder den ubekendte x og regner med den, som om den var kendt. Efter en række omskrivninger af den oprindelige ligning (som er den egenskab x skal opfylde), kommer vi (forhåbentligt) frem til nogle værdier af x. Hvis vi vil løse $x^2 + x = 2$ kan vi ligesom ovenfor argumentere som følger:

$$\begin{aligned}
x^2 + x &= 2 && (4.8)\\
\Rightarrow\ & x^2 + x + \frac{1}{4} = 2\frac{1}{4} && (4.9)\\
\Rightarrow\ & (x + \frac{1}{2})^2 = 2\frac{1}{4} && (4.10)\\
\Rightarrow\ & x + \frac{1}{2} = \pm\sqrt{2\frac{1}{4}} = \pm 1\frac{1}{2} && (4.11)\\
\Rightarrow\ & x = 1 \text{ eller } x = -2 && (4.12)
\end{aligned}$$

Det vi så har argumenteret for er, at hvis x opfylder ligningen så er $x = 1$ eller $x = -2$. For at bevise at 1 og -2 virkeligt er løsninger, skal vi "gøre prøve", dvs. vi skal indsætte de to værdier i ligningen og se, om de opfylder den. Det

gør de, hvorfor vi kan slutte at 1 og -2 er løsninger til ligningen, og de er de eneste.

I den ovenstående deduktion (4.7) brugte vi kun medførepile. Det var derfor vi efter endt deduktion ikke kunne slutte baglæns til, at 1 og -2 faktisk også er løsninger til ligningen. Man kan også ved ligningsløsning vælge at slutte ensbetydende ved hver omskrivning. Faktisk er alle de åbne udsagn i kæden (4.10) ensbetydende, så vi i virkeligheden kan slutte således:$\Leftrightarrow$

$$\begin{aligned} x^2 + x &= 2 && (4.13)\\ &\Leftrightarrow\ x^2 + x + \frac{1}{4} = 2\frac{1}{4} && (4.14)\\ &\Leftrightarrow\ (x + \frac{1}{2})^2 = 2\frac{1}{4} && (4.15)\\ &\Leftrightarrow\ x + \frac{1}{2} = \pm\sqrt{2\frac{1}{4}} = \pm 1\frac{1}{2} && (4.16)\\ &\Leftrightarrow\ x = 1 \text{ eller } x = -2 && (4.17) \end{aligned}$$

Når man gør det, kan man naturligvis straks konkludere at 1 og -2 er løsninger og de er de eneste. Der er altså ikke grund til at lave en særskilt syntese eller prøve. Vi har sammenbygget analysen og syntesen.

Nogle gange kan det godt betale sig at slutte gennem ensbetydende udsagn, men andre gange er det for besværligt at holde rede på begge implikationerne samtidigt, og det kan være en kilde til fejl i argumentet. Så meget ofte er det lettere og mere sikkert at holde de to implikationsretninger separate, og altså lave analysen og syntesen hver for sig.

4.5 Ikke stringent analyse

Hvis man bare er interesseret i et rent eksistensudsagn, er det som sagt ligegyldigt, hvordan man kommer frem til en kandidat, bare det kan bevises, at en virker. Man kan stadig have glæde af at lave en analyse, men så behøver man ikke gøre sig umage for at sikre sig, at ens slutninger i analysen er helt stringente. Lad os se på et eksempel. I parenteser angiver jeg de overvejelser vi springer over:

Problem 74 *Vis at der eksisterer en løsning til differentialligningen*

$$f''(x) = -f(x) \tag{4.18}$$

på $\mathbb{R}$ *som opfylder*

$$f(0) = 0 \quad og \quad f'(0) = 1. \tag{4.19}$$

Analyse: Antag at f løser problemet og skriv den som en potensrække

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 \cdots + a_n x^n + \cdots \tag{4.20}$$

(Det er slet ikke sikkert, at en eventuel løsning kan skrives som en potensrække; men det bekymrer vi os ikke om. Vi bekymrer os heller ikke for meget om rækkens konvergens). Vi differentierer rækken to gange og får:

$$f'(x) = a_1 + 2a_2x + 3a_3x^2 \cdots + na_nx^{n-1} + \cdots \quad (4.21)$$

$$f''(x) = 2a_2 + 2 \cdot 3a_3x \cdots + (n-1)na_nx^{n--2} + \cdots \quad (4.22)$$

(Her differentierer vi ledvist uden at bekymre os om vi nu må gøre det). Differentialligningen kan altså skrives:

$$2a_2 + 2 \cdot 3a_3x \cdots + (n-1)na_nx^{n--2} + \cdots \quad (4.23)$$

$$= -(a_0 + a_1x + a_2x^2 + a_3x^3 \cdots + a_{n-2}x^{n-2} + \cdots). \quad (4.24)$$

Sammenlignes led til ens potenser af x (må vi det?), får vi ligningssystemet:

$$2a_2 = -a_0 \quad (4.25)$$

$$2 \cdot 3a_3 = -a_1 \ ... \quad (4.26)$$

$$(n-1)na_n = a_{n-2} \ ... \quad (4.27)$$

Men da vi forudsatte, at $f(0) = 0$ og $f'(0) = 1$ (4.19), ser vi fra (4.20, 4.21), at $a_0 = 0$ og $a_1 = 1$. Så fra (4.26) ser vi, at $a_2 = 0$, $a_3 = -\frac{1}{3!}$, $a_4 = 0$, $a_5 = \frac{1}{5!}$ osv. (Vi behøver ikke gå igennem et egentligt induktionsbevis). Det fører til følgende rækkeudvikling:

$$f(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots \quad (4.28)$$

Vi genkender denne række som potensrækken for $\sin x$ og har derfor fundet en kandidat til en løsning af problemet.

Så kan det egentlige bevis (syntesen) begynde. Beviset går simpelthen ud på at checke, at funktionen $\sin x$ opfylder differentialligningen (4.18) på hele $\mathbb{R}$ og begyndelsesværdibetingelserne (4.19) i 0. Det ses let at være tilfældet. Dermed har vi løst eksistensproblemet og fundet en løsning på problemet.

Selv om vi ikke var omhyggelige med vores argumenter i analysen, gav den os altså en brugbar kandidat, som vi kunne bruge i eksistensbeviset. Men netop fordi vi ikke var omhyggelige med analysen, kan vi ikke bruge den til at slutte, at vi har fundet den eneste løsning. Analysen havde jo ikke formen: Hvis f er en løsning, så må den være af formen (4.28). Den havde i stedet formen: Hvis f er en løsning, og forskellige andre ting er opfyldt (især at f kan rækkeudvikles i en potensrække som er konvergent overalt på $\mathbb{R}$), så er f af formen (4.28). Hvis vi havde holdt rede på hvilke ting vi havde forudsat undervejs i analysen, kunne vi have sluttet at $\sin x$ er den eneste løsning til problemet, som opfylder disse forudsætninger. Men da vi ikke holdt rede på det må vi nøjes med at bruge analysen til at give os en kandidat til løsningen.

4.6 Geometriske eksempler

I geometri kan man bruge analyseideen til at finde løsninger på konstruktionsproblemer. I Euklids Elementer finder man kun synteserne. Med andre ord:

Når Euklid skal konstruere en figur med givne egenskaber så giver han en konstruktion og beviser dernæst, at den derved fremkomne figur faktisk har de forlangte egenskaber. Han fortæller ikke, hvordan han har fundet på konstruktionen. For eksempel er elementernes første sætning en løsning af en konstruktionsopgave, nemlig at konstruere en ligesidet trekant på en given ret linje. Euklid giver en konstruktion af en trekant og et bevis for, at den faktisk er en ligesidet trekant på den givne side. I dette tilfælde er det nok ikke så svært at se, hvordan han har fået ideen til konstruktionen, men lad os alligevel lave en analyse, som fører til konstruktionen.

Analyse: Antag derfor at ΔABC er en ligesidet trekant på det givne linjestykke AB. Da $AB = AC$, ligger C på en cirkel med centrum i A og radius AB. Da $BA = BC$, ligger C også på en cirkel med centrum B og radius AB. Altså må C være et skæringspunkt mellem disse to cirkler. Dem er der to af. Analysen har dermed frembragt to kandidater for C og dermed for den ligesidede trekant på AB.

Euklids syntetiske bevis starter med ud fra AB at konstruere de to cirkler og deres skæringspunkter C, og fortsætter med et bevis for, at trekanterne bestemt ved disse punkter faktisk løser opgaven. Men hvor Euklids argument kun giver et konstruktivt eksistensargument, så giver anlysen kombineret med syntesen endvidere, at der kun er de to løsninger. Euklid omtaler faktisk kun én løsning. Det er fordi de to her fundne trekanter er kongruente, og derfor ens i Euklids forstand.

Lad os se på to mindre oplagt eksempler:

Problem 75 *Konstruer en trekant med en given grundlinje AB, en given højde h fra C og en given sidelængde $a = BC$.*

Analyse: Antag at ΔABC opfylder det ønskede. (Det hjælper at lave en tegning, en såkaldt prøvefigur). Da BC har den givne længde a, ligger C på en cirkel med centrum i B og radius a. Og da højden fra C har længden h, ligger C på en af de to linjer parallelle med AB i afstanden h. Hvis nu $a > h$, så skærer cirklen hver af parallellerne i to punkter. Det giver altså fire kandidater til C. Hvis $a = h$, så rører cirklen hver af de to paralleller i et punkt. Der er altså to kandidater til C. Hvis $a < h$ skærer cirklen ikke de to paralleller, så der er ingen kandidater til C. I det sidste tilfælde har analysen altså vist, at der ikke er nogen løsninger. I de to andre tilfælde har analysen givet os nogle kandidater.

For at udføre *syntesen* skal vi først vise, hvordan vi kan konstruere de ønskede skæringspunkter. Hvis vi skal følge Euklid, skal det ske med passer og lineal. Vi kan naturligvis konstruere cirklen med passer og lineal, og det er heller ikke svært at se, hvordan de to paralleller kan konstrueres med passer og lineal (konstruktionen kan findes i Euklids Elementer). At cirklen og parallellerne skærer hinanden i det ovenfor angivne antal punkter, kræver aksiomer, som ikke findes hos Euklid; men hvis vi godtager dette intuitivt plausible resultat, kan vi let fuldføre syntesen, dvs. bevise at trekanterne bestemt af AB og hvert af de konstruerede punkter C faktisk løser problemet.

Problem 76 *Konstruer en ligebenet trekant med given grundlinje AB og en given topvinkel $\angle C = v$ mellem $0°$ og $180°$*

Analyse: Antag at ΔABC har de ønskede egenskaber. Da trekanten er ligebenet, er vinklerne ved grundlinjen lige store (Euklid bog I sætning 5). Hvis vi kalder vinklerne ved grundlinjen u så er $v + 2u = 180°$, thi vinkelsummen i en trekant er $180°$ (Euklid I, 32). Altså er $u = 90° - \frac{1}{2}v$. Ud fra denne oplysning er det let at lave syntesen.

Syntese: Konstruktion: Afsæt linjer ud fra A og B som danner vinklen $90° - \frac{1}{2}v$ med linjestykket AB. Det kan man gøre med passer og lineal (Euklid I, 23). Hvert par af linjer på samme side af AB vil skære hinanden (parallelpostulatet). Kald skæringspunkterne C_1 og C_2. De to trekanter ABC_1 og ABC_2 er da løsninger på problem 5. Bevis: Begge trekanter har grundlinje AB, de har begge ens vinkler ved grundlinjen og er derfor ligebenede (Euklid I, 6), og topvinklen er i dem begge lig med $180° - 2(90° - \frac{1}{2}v) = v$. QED

Syntesen har dermed bevist hvordan man kan konstruere to trekanter, der løser problemet, og analysen har vist at det er de to eneste løsninger. Læg mærke til at vi i analysen brugte at i en ligebenet trekant er vinklerne ved grundlinjen lige store (Euklid I, 5), medens vi i syntesen brugte den omvendte sætning (Euklid I, 6), som siger, at i en trekant, hvor to af vinklerne er lige store er de modstående sider lige store. Dette er helt typisk.

4.7 Advarsler

Man skal passe meget på med ikke at forveksle analysen med syntesen, og man skal være opmærksom på at analysen kun leverer kandidater og eventuelt viser entydighed (eller n-tydighed). Der er berømte eksempler på, at selv gode matematikere har forvekslet en analyse med en syntese. Lad os illustrere det med et bevis for cirklens såkaldt "isoperimetriske" egenskab:

Cirklen er den lukkede kurve med given omkreds (iso=ens, perimeter=omkreds), som omslutter det største areal.

Jacob Steiner gav følgende bevis: Han antog at en kurve med den givne omkreds omsluttede det størst mulige areal og viste så, at kurven måtte være en cirkel. (Vi skal ikke gå ind på hans elegante og korrekte bevis for dette). Derfra sluttede han at cirklen havde den isoperimetriske egenskab. Som Dirichlet senere påpegede, kan man naturligvis ikke slutte sådan. Det eneste man kan slutte fra Steiners argument er en entydighed, nemlig at *hvis* der eksisterer en isoperimetrisk kurve, så må det være cirklen. Steiner havde faktisk kun lavet analysen og havde glemt syntesen. Det skyldes naturligvis, at han mente, at det var klart, at der eksisterer en isoperimetrisk kurve. Men det er faktisk noget der bør bevises.

Lad mig vise et eksempel på hvordan man kan komme til et helt absurd resultat, hvis man på Steinersk vis forveksler analysen med syntesen, idet man forudsætter eksistensen af det objekt man vil finde:

"Sætning". Tallet 1 er det største naturlige tal.

"Bevis". Antag at n er det største naturlige tal. Jeg vil nu vise ved kontraposition at $n = 1$. Antag nemlig at $n \neq 1$. Så er $n^2 > n$ hvorfor n ikke er det største naturlige tal. Derfor må n altså være lig med 1. Derfor er 1 det største naturlige tal. QED

Konklusionen er her så åbenlys tåbelig, at alle kan se, at der er noget galt med beviset. Det, der er galt, er at vi har forvekslet analysen med syntesen. Vi argumenterede helt rigtigt for at *hvis* n er det største naturlige tal må n være lig med 1. Det er en analyse. Det vi heraf kan slutte er at 1 er den eneste mulige kandidat, altså at *hvis* der et største naturligt tal, må det være tallet 1. Det går først galt når vi derfra slutter at så må 1 være det største naturlige tal. Det ville vi kunne gøre, hvis vi havde et bevis for, at der eksisterede et største naturligt tal. Men da dette naturligvis er usandt, kan vi ikke slutte fra analysens entydighedsudsagn til eksistensudsagnet. Steiner's bevis for cirklens isoperimetriske egenskab er af helt samme natur, blot er eksistensudsagnet i dette tilfælde intuitivt plausibelt (og faktisk også korrekt under passende antagelser). Steiner gav dog ikke eksistensbeviset.

Lad os se på endnu en forkert brug af en analyse:

Definition 77 *Følgen af* ***Fibonaccital***[1] a_n *defineres rekursivt ved at* $a_1 = a_2 = 1$ *og* $a_{n+2} = a_{n+1} + a_n$.

De første led i følgen er: 1,1,2,3,5,8,13,21....

Sætning 78 *Forholdet* a_{n+1}/a_n *mellem to på hinanden følgende Fibonaccital konvergerer mod* $\frac{1+\sqrt{5}}{2}$ *for* $n \longrightarrow \infty$

"Bevis": Da $a_{n+2} = a_{n+1} + a_n$, fås ved division med a_{n+1}at

$$\frac{a_{n+2}}{a_{n+1}} = 1 + \frac{a_n}{a_{n+1}}. \tag{4.29}$$

Hvis grænseværdien af forholdet a_{n+1}/a_n kaldes A, ses af ovenstående ligning (ved brug af regnereglerne for grænseværdier), at

$$A = 1 + \frac{1}{A} \tag{4.30}$$

eller

$$A^2 = A + 1. \tag{4.31}$$

Denne ligning har rødderne $A = \frac{1\pm\sqrt{5}}{2}$. Da det er klart at grænseværdien ikke kan være negativ, må den være lig med $\frac{1+\sqrt{5}}{2}$ (det gyldne snit). QED

Kommentar: Resultatet er faktisk korrekt, men argumentet er ikke korrekt. I argumentet har vi jo taget for givet, at følgen a_{n+1}/a_n konvergerer. Vi har således i virkeligheden ikke lavet et bevis men en analyse. Vi har nemlig set

[1] Efter Leonardo Fibonacci af Pisa (ca. 1170-1240)

at *hvis* følgen a_{n+1}/a_n konvergerer, må grænseværdien være $\frac{1+\sqrt{5}}{2}$. Men vi har ikke vist konvergensen.

Igen kan vi illustrere hvor galt det kan gå, hvis man bruger et lignende argument på en anden følge:

"Sætning". Fibonaccifølgen a_n konvergerer mod 0 for $n \longrightarrow \infty$.

"Bevis". Hvis grænseværdien kaldes a, fås fra den definerende ligning $a_{n+2} = a_{n+1} + a_n$ at $a = a + a$, hvoraf vi slutter at $a = 0$.

"Sætningen" og dens "bevis" er naturligvis forkerte, og faktisk divergerer a_n mod ∞ for $n \longrightarrow \infty$.

4.8 Terminologi

I matematik betyder ordet analyse faktisk to meget forskellige ting. Den betydning vi her har set på går tilbage til antikke græske matematikere og filosoffer. Når man taler om matematisk analyse i dag mener man dog normalt den gren af matematikken, som er udsprunget af differential- og integralregningen og som indeholder emner som uendelige rækker, differentialligninger, topologiske vektorrum mm.

4.9 Problem- og sætningsanalyse

Alle analyser, vi har set på indtil nu, er analyser af problemer. Analyse-metoden er dog også nyttig når man skal bevise sætninger. Som forklaret i Afsnit 2.2 er et bevis for en sætning jo en kæde af udsagn, som ender med denne sætning, og hvori ethvert udsagn enten er et aksiom i teorien, eller en allerede bevist sætning eller fremgår ved en gyldig slutning af de tidligere udsagn i kæden. Beviser for sætninger starter altså med aksiomerne, eller allerede beviste sætninger og foregår ved successive gyldige slutninger derfra. Hvis man skal bevise en bestemt sætning er problemet naturligvis, at det sjældent er klart, hvilke aksiomer og tidligere sætninger, man skal ty til, og hvilke logiske slutninger, man skal lave for at nå frem til sætningen. Blind brug af tidligere sætninger og slutningsregler fører sjældent (aldrig) til målet. Derimod kan man få ideen til beviset ved at starte med den sætning, man skal bevise, og så prøve at finde nogle udsagn, hvorfra sætningen kan udledes. Hvis disse er aksiomer eller allerede beviste sætninger er vi færdige. Hvis ikke, prøver vi at finde andre udsagn, hvorfra de kan udledes og så videre. Hvis man er heldig, vil man til sidst ende med nogle udsagn, som man ved er sande, fordi de enten er aksiomer eller allerede beviste sætninger.

For at tydeliggøre denne ide, kan vi se på en sætning s, som kan bevises ved følgende simple kæde af udsagn startende fra aksiomet a.

$$a \Rightarrow p_1 \Rightarrow p_2 \Rightarrow p_3 \Rightarrow p_4 \Rightarrow s \tag{4.32}$$

Hvis vi bliver bedt om at finde et bevis for s, kan vi gå frem på følgende måde: Jeg vil bevise s. Hvis jeg bare kunne bevise p_4, ville jeg være færdig, thi s følger

fra p_4. Men for at bevise p_4, behøver jeg bare bevise p_3, thi p_3 medfører p_4. Og da $p_2 \Rightarrow p_3$, vil jeg være færdig, hvis jeg bare kunne vise p_2. Men p_2 følger fra p_1, som er sand, da den følger fra aksiomet a. På denne måde vil vi altså have trævlet beviset op bagfra. Jeg vil kalde dette en analyse, fordi vi begynder med det vi vil vise og ender med det vi ved er sandt. Men da vi hele tiden har argumenteret med pilene i den rigtige retning, giver analysen faktisk et korrekt bevis. Man vil ofte foretrække at præsentere det færdige bevis ved at starte med a og så arbejde sig til højre i kæden af udsagn. Men det skyldes kun æstetiske hensyn og ønsket om at gøre beviset mere overskueligt. Derved adskiller denne type analyser sig fra de ovenfor behandlede problemanalyser, hvor analysen og syntesen var helt adskildte.

Man kunne naturligvis også lave en bevisanalyse ved at gå gennem en kæde af konsekvenser af sætningen indtil man når et aksiom eller noget bevist:

$$a \Leftarrow p_1 \Leftarrow p_2 \Leftarrow p_3 \Leftarrow p_4 \Leftarrow s, \tag{4.33}$$

men så skal man ligesom i problemanalysen efterfølgende lave en separat syntese, som sikrer at man faktisk kan gå den modsatte vej fra a til s.

I de fleste tilfælde er strukturen af et bevis for en sætning mere kompleks end en simpel kæde som (4.32), og i mange tilfælde har man faktisk også en fornemmelse af hvilke sande sætninger eller aksiomer, man skal begynde med. Det gælder især i matematikundervisningen hvor man bliver præsenteret for en velformuleret sætning, som man skal bevise. I dette tilfælde har sætningen altid formen: Hvis a, b, c, d så e. I beviset skal man så antage at a, b, c, d er sande og skal så bevise at e er sand. I dette tilfælde er det klart at udsagnene a, b, c, d skal indgå i beviset (måske sammen med andre aksiomer og sætninger i den teori sætningen er en del af). I sådanne tilfælde vil man normalt prøve at finde beviset ved at starte fra begge ender og se, om man kan få argumentet til at mødes på midten. Mere præcist, man vil prøve at finde konsekvenser af a, b, c, d, som går i retning af e, og man vil prøve at finde udsagn som medfører e og som ligger "nærmere" ved a, b, c, d. Forhåbentligt kan man så få processen til at mødes på midten. Vi skal senere se på nogle eksempler.

Chapter 5

Induktionsbeviser

5.1 Simpel induktion

Ved beviser for sætninger, der involverer et vilkårligt naturligt tal, kan man ofte benytte en særlig bevisteknik kaldet induktion. Lad os som eksempel betragte følgende sætning:

Sætning 79 *For ethvert $n \in \mathbb{N}$ gælder, at summen af de første n ulige tal er lig med n^2, altså:*

$$1 + 3 + 5 + \cdots + (2n - 1) = n^2 \tag{5.1}$$

Man kan naturligvis prøve sætningen af for de første værdier af n:

$$\begin{aligned} 1 &= 1^2 && (5.2)\\ 1 + 3 &= 4 = 2^2 && (5.3)\\ 1 + 3 + 5 &= 9 = 3^2 && (5.4)\\ &\text{osv.} && (5.5) \end{aligned}$$

Men uanset hvor mange værdier af n vi tester sætningen for, vil det ikke kunne gøre det ud for et bevis. Højst en sandsynliggørelse.

Den generelle metode til at bevise en universel sætning, som skal gælde om alle $n \in \mathbb{N}$, er at antage at n er et vilkårligt naturligt tal, og så bruge generelle sætninger om naturlige tal til at vise sætningen for n. For eksempel vil et bevis for den generelle gyldighed af

$$(n + 1)^2 = n^2 + 2n + 1 \tag{5.6}$$

forløbe således:

Lad $n \in \mathbb{N}$. Da følger af regnereglerne for naturlige tal, at

$$\begin{aligned} (n + 1)^2 &= (n + 1)(n + 1) && (5.7)\\ &= n^2 + n + n + 1 = n^2 + 2n + 1. && (5.8) \end{aligned}$$

Det er imidlertid svært at se, hvordan man på denne måde direkte skulle kunne udlede (5.1). Hvis man derimod først har fået udledt sætningen for et bestemt n, så er det let at se at sætningen også er sand for det næste naturlige tal $n+1$. Hvis vi ved at $1+3+5+\cdots+(2n-1)=n^2$, så kan vi nemlig argumentere som følger:

$$(n+1)^2 = n^2+2n+1 = 1+3+5+\cdots+(2n-1)+(2n+1)\,, \tag{5.9}$$

hvorfor sætningen er sand for det efterfølgende naturlige tal $n+1$. Men når først vi har indset, hvordan vi på denne måde kan komme fra n til $n+1$, kan vi jo argumentere for sætningen på følgende vis:

Først indses at (5.1) er sand for $n=1$. Det har vi checket i (5.2). Men så må (5.1) gælde for det efterfølgende naturlige tal, altså for $n=2$, hvorfor det gælder for $n=3$, hvorfor det gælder for $n=4$, osv. Ligesom i (5.5) slutter dette argument med osv. men nu har vi vished for, at vi faktisk kan fortsætte argumentet skridt for skridt lige så længe vi vil, hvorved vi vil kunne nå et hvilket som helst naturligt tal.

Generelt kan vi bevise en sætning for alle værdier af $n \in \mathbb{N}$ ved at bevise

1. at sætningen er sand for $n=1$

2. at *hvis* sætningen er sand for n lig en bestemt værdi m, *så* er den sand for $n=m+1$

Vi kan formulere dette princip som en sætning:

Sætning 80 *(Princippet om **simpel induktion**). Lad $p(x)$ være et prædikat, hvor den frie variabel x kan løbe over de naturlige tal $\mathbb{N}$.*

Såfremt $p(x)$ har følgende 2 egenskaber:

1. *$p(1)$ er sand,*

2. *for hvert $m \in \mathbb{N}$, kan man af $p(m)$ slutte $p(m+1)$,*

da gælder $p(n)$ for alle $n \in \mathbb{N}$.

Når et bevis gennemføres efter dette princip, kalder man det et **induktionsbevis.** Punkt 1. kaldes **induktionsstarten**, og punkt 2. kaldes **induktionsskridtet**. I punkt 2. antager man altså $p(m)$, og konkluderer da $p(m+1)$. Derfor kaldes $p(m)$ for **induktionsantagelsen**.

Lad os gennemføre induktionsbeviset for sætning (79):

Bevis. for sætning (79): Lad $p(n)$ betegne følgende prædikat for $n \in \mathbb{N}$

$$p(n) : 1+3+5+\cdots+(2n-1)=n^2. \tag{5.10}$$

1. Induktionsstarten: Da $1=1^2$ er $p(1)$ sand.

2. Induktionsskridtet: Antag at $p(m)$ er sand, altså at

$$1 + 3 + 5 + \cdots + (2m - 1) = m^2. \tag{5.11}$$

Da viser følgende udregning at $p(m + 1)$ er sand:

$$(m + 1)^2 = m^2 + 2m + 1 = 1 + 3 + 5 + \cdots + (2m - 1) + (2m + 1)\,. \tag{5.12}$$

I denne udregning brugte vi induktionsantagelsen (5.11) i andet lighedstegn.

Af princippet om simpel induktion følger at $p(n)$ er sand for alle $n \in \mathbb{N}$, og dermed er sætningen bevist. ■

Lad os gennemføre et par andre eksempler på induktionsbeviser.

Sætning 81 *En $(n + 2)$-kant har vinkelsum lig $n \cdot 180°$.*

Bevis. Vi beviser sætningen ved induktion efter n:

1. Induktionsstarten: Det forudsættes bekendt, at vinkelsummen i en trekant er 180°. Et bevis findes i Euklids Elementer bog I.

2. Induktionsskridtet: Antag, at for en bestemt værdi af $m \in \mathbb{N}$ er vinkelsummen i enhver $(m+2)$-kant lig $m \cdot 180°$. Betragt en vilkårlig $((m + 1) + 2)$-kant, altså en $(m + 3)$-kant. Ved at tegne en passende valgt diagonal, kan vi dele denne $(m+3)$-kant i en trekant og en $(m+2)$-kant. Vinkelsummen i $(m+3)$-kanten er da summen af vinkelsummen i trekanten og vinkelsummen i $(m + 2)$-kanten, altså lig $180° + m \cdot 180° = (m + 1) \cdot 180°$. Hermed er påstanden vist for $n = (m + 1)$.

Ifølge princippet om simpel induktion er sætningen dermed bevist for alle $n \in \mathbb{N}$. ■

Sætning 82 *For ethvert $n \in \mathbb{N}$ går 6 op i $(n^3 - n)$.*

Bevis. Vi beviser sætningen ved induktion efter n.

1. *Induktionsstarten*: Sætningen er sand for $n = 1$, da $\left(1^3 - 1\right) = 0$ og $6 \mid 0$.

2. *Induktionsskridtet*: Antag at $6 \mid \left(m^3 - m\right)$ for en bestemt værdi af m. Vi skal vise at $6 \mid \left((m + 1)^3 - (m + 1)\right)$. Nu gælder at

$$\begin{aligned} (m + 1)^3 - (m + 1) &= m^3 + 3m^2 + 3m + 1 - m - 1 && (5.13)\\ &= (m^3 - m) + 3m\,(m + 1)\,. && (5.14) \end{aligned}$$

Ifølge induktionsantagelsen går 6 op i det første led, og da et af tallene m og $m + 1$ er lige må $m\,(m + 1)$ indeholde en faktor 2, så 6 også går op i sidste led. Derfor må 6 gå op i summen. Vi har altså vist at $6 \mid \left(m^3 - m\right)$.

Ifølge princippet om simpel induktion er sætningen dermed sand for alle $n \in \mathbb{N}$. ■

Induktionsskridtet kan også formuleres $p(m) \Rightarrow p(m+1)$. Her er det værd at huske, at dette udsagn *ikke* betyder: "*da* $p(m)$ er sand, er $p(m+1)$ også sand" men derimod "*hvis* $p(m)$ er sand, så er $p(m+1)$ også sand". Induktionsskridtet alene beviser altså ikke at $p(m+1)$ er sand. Først når Induktionsstarten (altså sandheden af $p(1)$) er bevist, kan man successivt bruge induktionsskridtet til at slutte: Da $p(1)$ er sand er $p(2)$ sand. Da $p(2)$ er sand er $p(3)$ sand, osv.

Hvis $q(n)$ betegner prædikatet

$$q(n) : 1 + 3 + 5 + \cdots + (2n-1) = n^2 + 7 \tag{5.15}$$

(sammenlign med (5.10), så kan vi bevise $q(m) \Rightarrow q(m+1)$, ganske som vi viste induktionsskridtet i beviset for sætning (79). Men det viser intet om, hvorvidt $q(n)$ er sand for nogle eller alle $n \in \mathbb{N}$. Først hvis vi kunne vise, at $q(1)$ var sand, ville vi kunne slutte at q ville være sand for alle $n \in \mathbb{N}$. Men vi kan naturligvis ikke vise $q(1)$ (overvej), og faktisk er $q(n)$ falsk for alle $n \in \mathbb{N}$.

Moralen er at man skal huske at bevise induktionsstarten.

Man skal også være meget opmærksom på om induktionsskridtet virkelig gælder for alle $m \in \mathbb{N}$. Ellers kan man "bevise" "sætninger" som den følgende:

"**Sætning**": Alle kaniner har samme farve.

"**Bevis**": Det er klart, at der er et endeligt antal kaniner i verden. Sætningen er altså bevist, hvis vi for ethvert $n \in \mathbb{N}$ kan vise, at i en mængde af n kaniner, har alle kaninerne samme farve. Vi fører beviset herfor ved induktion efter n.

Induktionsstarten: I en mængde med kun 1 kanin, er det klart at alle kaniner har samme farve.

Induktionsskridtet: Antag nu at sætningen er sand for enhver mængde med m kaniner. Lad K være en mængde med $m+1$ kaniner. Tag en kanin k_1 væk fra K. Ifølge induktionsantagelsen har de resterende m kaniner samme farve (kald den F). Sæt nu k_1 tilbage, og borttag en anden kanin k_2, som jo har farven F. Da har de resterende kaniner (igen ifølge induktionsantagelsen) alle samme farve, og da alle kaninerne på nær måske k_1 har farven F, må alle kaninerne have farven F. Derfor har alle kaninerne i K samme farve.

Ifølge princippet om simpel induktion har kaninerne i enhver endelig mængde af kaniner samme farve.

Øvelse 83 *Erfaring og sund fornuft fortæller os at der må være noget galt med ovenstående bevis. Find fejlen.*

5.2 Fuldstændig induktion

Nogle gange er det ikke nok at vide, at $p(m)$ er sand, når man skal vise at $p(m+1)$ er sand. Men det kan være, at man kan slutte $p(m+1)$, når man ved at $p(n)$ er sand for alle foregående naturlige tal, altså for $1, 2, 3, \cdots, m$. Og hvis man kan vise at $p(1)$ er sand, kan man skridtvis slutte som følger: Da $p(1)$ er sand er $p(2)$ sand; da $p(1)$ og $p(2)$ er sande, er $p(3)$ sand; da $p(1)$, $p(2)$ og $p(3)$

er sande er $p(4)$ sand; osv. Dermed kan vi bevise at $p(n)$ er sand for alle $n \in \mathbb{N}$. Vi formulerer denne bevisstrategi som en sætning:

Sætning 84 *(Princippet om **fuldstændig induktion**). Lad $p(x)$ være et prædikat, hvor den frie variabel x kan løbe over de naturlige tal $\mathbb{N}$.*

Såfremt $p(x)$ har følgende 2 egenskaber:

1. *$p(1)$ er sand,*
2. *for hvert $m \in \mathbb{N}$ kan man af $p(1)$, $p(2)$, ..., $p(m)$ slutte $p(m+1)$,*

da gælder $p(n)$ for alle $n \in \mathbb{N}$.

Vi illustrerer brugen af fuldstændig induktion med nedenstående sætning.

Sætning 85 *Ethvert naturligt tal $n > 1$ er et produkt af primtal.*

Bevis. Bemærk først, at når vi taler om "et produkt af primtal", så er det underforstået at dette produkt gerne må bestå af kun en faktor; med andre ord: Et primtal anses for i sig selv at være et "produkt af primtal".

Betragtes prædikatet $p(n)$ defineret ved:

$$n = 1 \vee (n \text{ er et produkt af primtal}), \tag{5.16}$$

hvor den frie variabel n tillades at løbe over de naturlige tal $\mathbb{N}$, så er vores opgave at vise, at $p(n)$ er sand for alle $n \in \mathbb{N}$.

Vi viser dette ved fuldstændig induktion efter n.

Induktionsstarten: Da $1 = 1$ er $p(1)$ sand.

Induktionsskridtet: Vi lader nu $m \in \mathbb{N}$ være vilkårlig, og antager at udsagnene $p(1)$, $p(2)$, ..., $p(m)$ alle er sande. Vi skal da vise $p(m+1)$. Da $m + 1 > 1$ skal vi vise, at $m + 1$ er et produkt af primtal. Vi splitter beviset herfor op i to tilfælde:

1. $m + 1$ er et primtal: I så fald er $m + 1$ ifølge ovenstående konvention et produkt af primtal, og vi er færdige.
2. $m+1$ ikke er et primtal: Da $m+1 > 1$, må $m+1$ i så fald være sammensat (se Definition 40). Der findes altså $a, b \in \mathbb{N}$ så:

$$m + 1 = a \cdot b, \tag{5.17}$$

og så $1 < a, b < m + 1$. Da $a, b < m + 1$, gælder $a, b \leq m$, så udsagnene $p(a)$ og $p(b)$ forekommer begge i listen $p(1)$, $p(2)$, ..., $p(m)$. På grund af induktionsantagelsen ved vi altså at både $p(a)$ og $p(b)$ er sande. Idet $a > 1$ og $b > 1$, betyder det, at såvel a og b er et produkt af primtal. Det samme gælder da om $m + 1 = a \cdot b$.

Altså er $p(m + 1)$ sand.

Sætningen følger nu fra princippet om fuldstændig induktion. ■

5.3 Induktionsaksiomet

Vi har præsenteret induktionsprincipperne som oplagte følger af sund fornuft. Formelt er de følger af aksiomerne for de naturlige tal. Disse kan formuleres som følger:

Definition 86 ***Peano's aksiomssystem for de naturlige tal***[1]

De naturlige tal er en mængde $\mathbb{N}$ *udstyret med en efterfølgerfunktion* $S : \mathbb{N} \longrightarrow \mathbb{N}$*, hvorom det gælder:*

1. $1 \in \mathbb{N}$.
2. *For ethvert* $n \in \mathbb{N} : 1 \neq S(n)$.
3. *For ethvert* $m, n \in \mathbb{N} : m \neq n \Rightarrow S(m) \neq S(n)$.
4. ***Induktionsaksiomet****: Hvis det om en delmængde* $A \subseteq \mathbb{N}$ *gælder, at* $1 \in A$*, og* $m \in A \Rightarrow S(m) \in A$*, så gælder, at* $A = \mathbb{N}$.

Fra dette aksiomssystem kan man opbygge hele aritmetikken for naturlige tal. Det vil vi dog ikke gøre i denne bog. Vi vil nøjes med at bevise, at princippерne om simpel og fuldstændig induktion er konsekvenser af aksiomerne, specielt induktionsaksiomet.

Vi bemærker at operationen addition (+) indføres på en måde så at $S(n) = n + 1$ for alle $n \in \mathbb{N}$.

Intuitivt siger de første tre aksiomer, at man kan starte med 1 og successivt danne *nye* naturlige tal ved at tage efterfølgeren (altså lægge en til). Induktionsaksiomet siger, at alle naturlige tal kan nås på denne måde. Hvor de første aksiomer kan opfattes som generatorer af naturlige tal, er induktionsaksiomet et aksiom, der sikrer, at der ikke er flere naturlige tal, end de tal som genereres ved hjælp at de første tre aksiomer.

Intuitivt er det klart, at det netop er induktionsaksiomet, som får induktionsbeviser til at virke. Når vi ved at $p(1)$ og $p(n) \Rightarrow p(n+1)$, kan vi jo successivt slutte at p er sand for alle de successive efterfølgere af 1, og induktionsaksiomet siger at der ikke er andre naturlige tal end disse.

Lad os formalisere denne idé:

Bevis. af sætning (80) (Simpel induktion): Lad $p(n)$ være et prædikat, hvor den frie variabel kan løbe over de naturlige tal. Antag endvidere at $p(1)$ er sand og at $p(m) \Rightarrow p(m+1)$. Vi skal da vise at $p(n)$ er sand for alle $n \in \mathbb{N}$.

Betragt sandhedsmængden for p altså

$$A = \{n \in \mathbb{N} \mid p(n) \text{ er sand}\}. \tag{5.18}$$

Da $p(1)$ er sand gælder, at $1 \in A$.

Da $p(m) \Rightarrow p(m+1)$, gælder det at $m \in A \Rightarrow S(m) \in A$.

[1] Opkaldt efter den Italienske matematiker Giuseppe Peano (1858-1932)

Mængden A opfylder altså forudsætningerne i induktionsaksiomet, hvorfor vi af aksiomet kan slutte at $A = \mathbb{N}$. Men det betyder at $p(n)$ er sand for alle $n \in \mathbb{N}$.

Dermed har vi vist princippet om simpel induktion. ■

For at se at induktionsaksiomet er nødvendigt for at kunne bruge princippet om simpel induktion, kan vi betragte følgende eksempel:

Eksempel 87 *Lad*

$$A = \left\{ x \in \mathbb{R} \mid x = 1 - \frac{1}{n} \text{ for et } n \in \mathbb{N} \right\} \tag{5.19}$$

og

$$B = \left\{ x \in \mathbb{R} \mid x = 2 - \frac{1}{n} \text{ for et } n \in \mathbb{N} \right\} \tag{5.20}$$

og definer $C = A \cup B$. Definer efterfølgerfunktionen $S : C \longrightarrow C$ ved

$$S(1 - \frac{1}{n}) = 1 - \frac{1}{n+1}, \tag{5.21}$$

$$S(2 - \frac{1}{n}) = 2 - \frac{1}{n+1} \tag{5.22}$$

for alle $n \in \mathbb{N}$. Da opfylder C med denne efterfølgerfunktion de tre første aksiomer i Peanos aksiomssystem men ikke induktionsaksiomet. Der gælder jo at $A \subseteq C$ og $0 \in A$ og $n \in A \Rightarrow S(n) \in A$, men $A \neq C$.

Hvis $p(x)$ er et prædikat, hvor den frie variabel kan løbe over C og vi ved at $p(1 - \frac{1}{1})$ er sand, og at $p(x) \Rightarrow p(S(x))$ så er det intuitivt klart at $p(x)$ er sand for alle $x \in A$, men det er også klart, at man ikke kan slutte at $p(x)$ er sand for $x \in B$. Vi kan jo ikke ved successivt at tage efterfølgeren, nå fra $1 - \frac{1}{1}$ til et element i B.

Vi vil dernæst bevise princippet om fuldstændig induktion ud fra princippet om simpel induktion:

Bevis. af sætning (84) (Fuldstændig induktion): Lad $p(n)$ være et prædikat, hvor den frie variabel kan løbe over de naturlige tal. Antag endvidere at $p(1)$ er sand og at $p(1) \wedge p(2) \wedge \ldots \wedge p(m) \Rightarrow p(m+1)$. Vi skal da vise at $p(n)$ er sand for alle $n \in \mathbb{N}$.

Betragt hertil følgende prædikat i den frie variabel n :

$$q(n) : (\forall k \in \mathbb{N} : k \leq n \Rightarrow p(k)). \tag{5.23}$$

Den frie variabel n kan her antage værdier i $\mathbb{N}$.[2]

Vi påstår nu, at vi kan vise $q(n)$ for alle $n \in \mathbb{N}$ via simpel induktion. Er dette gjort, følger $p(n)$ for alle $n \in \mathbb{N}$: For hvis vi har $q(n)$ for et $n \in \mathbb{N}$, er implikationen $k \leq n \Rightarrow p(k)$ sand for ethvert $k \in \mathbb{N}$. Idet $n \leq n$ kan vi derfor slutte $p(n)$.

[2] Vi kan også skrive $q(n)$ som $p(1) \wedge p(2) \wedge \ldots \wedge p(n)$.

Induktionsstarten: $q(1)$ er udsagnet: $\forall k \in \mathbb{N} : k \leq 1 \Rightarrow p(k)$. Dette udsagn er sandt, thi det eneste naturlige tal med $k \leq 1$ er tallet 1, og vi har antaget at $p(1)$ er sand.

Induktionsskridtet: Antag nu at $q(m)$ er sand for et $m \in \mathbb{N}$. Da udsagnet $k \leq m$ er sandt for $k = 1, 2, \ldots, m$, følger derfor $p(k)$ for $k = 1, 2, \ldots, m$. På grund af vores antagelse om $p(x)$, kan vi heraf slutte $p(m+1)$. Men da er implikationen

$$k \leq m + 1 \Rightarrow p(k) \tag{5.24}$$

sand, d.v.s. vi har sluttet $q(m+1)$.

Fra princippet om simpel induktion kan vi nu slutte at $q(n)$ er sand for alle $n \in \mathbb{N}$. ■

Man kan omvendt vise, at princippet om simpel induktion følger af princippet om fuldstændig induktion. Beviset overlades til læseren. De to principper er altså ækvivalente.

Bemærkning 88 *Vær opmærksom på, at det ikke er alle sætninger om naturlige tal, som mest hensigtsmæssigt bevises ved et induktionsbevis. Mange sætninger bevises som sædvanlige universelle sætninger, jvf. beviset i starten af dette afsnit for at* $(n+1)^2 = n^2 + 2n + 1$. *Det kan kun betale sig at bruge et induktionsbevis, hvis beviset for* $p(m+1)$ *simplificeres ved at antage, at* $p(m)$ *(eller* $p(1) \wedge p(2) \wedge \ldots \wedge p(m)$*) er sand. Hvis du opdager, at du slet ikke har brugt induktionsantagelsen i beviset for* $p(m+1)$*, så har du jo et universelt bevis for* $p(m+1)$.

Bemærkning 89 *I filosofisk og dagligdags tale betyder induktion noget andet end i matematik. Induktion i filosofisk forstand betyder en slutning fra det specielle til det generelle, som når man ud fra enkeltobservationer slutter sig til en generel lovmæssighed. For eksempel har jeg observeret, at solen står op hver dag i mit liv, og jeg slutter derfra den generelle lovmæssighed, at solen står op hver dag. Ligeså har man observeret at de kendte planeter bevæger sig om solen i ellipser, og slutter derfra, at alle planeter (også de eventuelt uopdagede) vil bevæge sig om solen i ellipser. Disse slutninger er induktive. I modsætning hertil står deduktioner, som er slutninger fra det generelle til det specielle. For eksempel, når man fra gravitationsloven udleder, at planeter bevæger sig i ellipser, er der tale om en deduktion.*

Naturvidenskaber benytter induktion, hvorimod matematik kun bruger deduktive slutninger. Selv matematisk induktion er en type deduktion, i ordets filosofiske betydning.

5.4 Rekursion

Hvis $a \in \mathbb{R}$, defineres a^n for ethvert $n \in \mathbb{N}$ ved følgende to regler:

1. $a^1 = a$,
2. $\forall n \in \mathbb{N} : a^{n+1} = a \cdot a^n$.

Ideen i definitionen er at vi starter med at definere a^1 og derfra successivt (rekursivt) bestemmer a^2, a^3, osv. ud fra den foregående værdi, ved hjælp af den anden regel. Man kalder en sådan definition for en rekursiv definition.

Ved første øjekast kan det synes ret oplagt, at man på denne måde kan definere a^n for alle $n \in \mathbb{N}$. Men ligesom i tilfældet med induktionsprincippet bør man bevise, at vi har formuleret en holdbar definition. Det bygger faktisk på følgende sætning af Dedekind (1888):

Sætning 90 ***Rekursionssætningen****: Lad der være givet en mængde A, et element $a \in A$, og en afbildning $f : A \longrightarrow A$.*

Da findes der netop en afbildning $\phi : \mathbb{N} \longrightarrow A$ med følgende egenskaber:

1. $\phi(1) = a$,
2. $\forall n \in \mathbb{N} : \phi(n+1) = f(\phi(n))$.

Det intuitive indhold i sætningen er at vi successivt definerer

$$\begin{aligned} \phi(1) &: = a && (5.25)\\ \phi(2) &: = f(\phi(1)) = f(a) && (5.26)\\ \phi(3) &: = f(\phi(2)) = f(f(a)) && (5.27)\\ &\quad \text{osv.} && (5.28) \end{aligned}$$

Vi skal ikke gennemgå det subtile formelle bevis for denne sætning, blot nævne at det bygger på induktionsaksiomet.

Lad os se, hvordan rekursionssætningen kan begrunde den ovennævnte definition af a^n:

Lad $A = \mathbb{R}$ og $a \in \mathbb{R}$ vilkårlig, og $f(x) := a \cdot x$. Rekursionssætningen siger da at der findes netop en afbildning $\phi : \mathbb{N} \longrightarrow \mathbb{R}$, så

1. $\phi(1) = a$,
2. $\forall n \in \mathbb{N} : \phi(n+1) = a \cdot \phi(n)$.

Man definerer da $a^n = \phi(n)$.

Ofte vil man i en rekursiv definition af en funktion benytte andre end den umiddelbare forgænger ved definitionen af $\phi(n+1)$. Man kan formulere og bevise en generalisation af rekursionssætningen, der siger at dette er tilladt og entydigt bestemt. Vi vil dog ikke formulere denne generalisering, da den er af noget teknisk natur.

Definition 77 af Fibonaccitallene er et eksempel på en sådan mere generel rekursiv definition.

Rekursive definitioner forekommer gang på gang i matematikken, specielt i kombinatorik og talteori.

5.5 Opgaver

1. Brug et induktionsbevis til at bevise, at for ethvert $n \in \mathbb{N}$ gælder:

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}. \tag{5.29}$$

2. Brug fuldstændig induktion til at bevise, at ethvert naturligt tal er en sum af forskellige potenser af to. (Overvej, hvorfor argumentet fejler, hvis du prøver at vise at ethvert naturligt tal kan skrives som en sum af forskellige potenser af tre).

3. Giv et induktionsbevis for at 4 går op i $5^n - 1$ for alle $n \in \mathbb{N}$.

4. Gæt en formel for

$$a_n = \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{n(n+1)}, \tag{5.30}$$

og bevis din formodning.

5. Bevis at følgende formler er korrekte for alle $n \in \mathbb{N}$:

$$2 + 5 + 8 + \cdots + (3n - 1) = \frac{n(3n+1)}{2} \tag{5.31}$$

$$1 + 5 + 9 + \cdots + (4n - 3) = n(2n - 1) \tag{5.32}$$

$$1^3 + 2^3 + 3^3 + \cdots + n^3 = \left(\frac{n(n+1)}{2}\right)^2. \tag{5.33}$$

6. Lad $f(x) = \ln x$. Udregn de første afledede af f, og opstil en formodning om en formel for $\frac{d^n f}{dx^n}$. Bevis din formodning ved induktion.

7. a. Bevis, at enhver ikke tom delmængde af de naturlige tal har et mindste element (velordningsprincippet). Vink: Antag at $A \subseteq \mathbb{N}$ og at A ikke har et mindste element. Brug da induktionsaksiomet på $\complement A$ til at vise, at A er tom.

8. Lad $x \neq 1$ være et reelt tal. Definer rekursivt betydningen af

$$1 + x + x^2 + \cdots + x^n, \tag{5.34}$$

og bevis ved induktion at

$$1 + x + x^2 + \cdots + x^n = \frac{x^{n+1} - 1}{x - 1}. \tag{5.35}$$

9. Lad $k \in \mathbb{N}$ og lad $A \subseteq \mathbb{N}$, og antag at det gælder at

1. $k \in A$

2. $m \in A \Rightarrow (m + 1) \in A$.

Vis at $\{n \in \mathbb{N} \mid k \leq n\} \subseteq A$.

10. Formuler en sætning, svarende til princippet om simpel induktion, men hvor induktionen ikke starter ved 1, men ved et vilkårligt $k \in \mathbb{N}$. Brug resultatet i opgave 9 til at bevise din sætning.

11. Betragt polynomier

$$f(x) = a_n x^n + \ldots + a_1 x + a_0 \tag{5.36}$$

med rationale koefficienter $a_0, \ldots a_n$. Hvis $a_n \neq 0$, så kaldes n for f's grad og betegnes med $\deg f$. Således er $\deg f$ kun defineret, hvis $f \neq 0$.

Et polynomium f af grad ≥ 1 kaldes reducibelt, hvis der findes en spaltning $f = g \cdot h$, hvor g, h er polynomier med grader strengt mindre end $\deg f$. Er f ikke reducibel, kaldes f irreducibel.

Vis ved fuldstændig induktion, at ethvert polynomium er et produkt af irreducible polynomier

Chapter 6

Mængdelære

6.1 Hvad er en mængde?

En mængde er det mest basale begreb i matematikken. Man definerer alle andre matematiske begreber inden for mængdelæren. Men hvordan definerer man da mængder? Man definerer dem ud fra en række aksiomer. Det foretrukne aksiomssystem kaldes Zermelo-Fraenkels aksiomssystem eller ZFC hvor C står for det omdiskuterede udvalgsaksiom (axiom of Choice). I dette kursus vil vi nøjes med en mere naiv tilgang til mængder, idet vi betragter dem som samlinger af ting, hvor "ting" skal tages i vid betydning. Tingene, som udgør mængden kaldes dens *elementer.*

Man angiver ofte mængder med store bogstaver og elementerne med små bogstaver. Man kan angive en mængde ved at opskrive en liste af dens elementer i en tuborg-parentes. For eksempel betyder $M = \{1, 2, 7\}$ mængden bestående af de tre tal 1, 2 og 7.. Elementernes rækkefølge er ligegyldig. Altså $\{1, 2, 7\} = \{2, 7, 1\}$. Mængder kan have andre mængder som elementer. For eksempel har mængden $\{4, \{2, 3\}\}$ to elementer nemlig 4 og $\{2, 3\}$, medens tallene 2 og 3 *ikke* er element i mængden $\{4, \{2, 3\}\}$. At x er element i M skrives $x \in M$ og siges ofte: "x ligger i M". Altså har vi $4 \in \{4, \{2, 3\}\}$, og $\{2, 3\} \in \{4, \{2, 3\}\}$, hvorimod $2 \notin \{4, \{2, 3\}\}$. Her betyder $\notin$ naturligvis "ikke element i" altså: $x \notin M \Leftrightarrow \neg(x \in M)$.

En mængde er entydigt karakteriseret ved sine elementer. Det betyder at to mængder er ens, hvis de har samme elementer: Eller skrevet formelt:

Definition 91 *Lad A og B være mængder. Da er $A = B$ hvis*

$$x \in A \Leftrightarrow x \in B. \tag{6.1}$$

Mængder kan godt have ét element f.eks. $\{1\}$. Det er dog vigtigt at sondre mellem tallet 1 og mængden $\{1\}$, som har det ene element 1. Ja, man tillader endog at en mængde slet ingen elementer har. Eksistensen af en sådan mængde er et aksiom. Vi vil nu bevise entydigheden:

Sætning 92 *Der er præcist én mængde uden nogen elementer.*

Bevis. Lad A og B være to mængder uden elementer. Vi skal da vise at $A = B$. Ifølge definition (91) skal vi altså vise at

$$x \in A \Leftrightarrow x \in B. \tag{6.2}$$

Men dette udsagn er sandt, da både $x \in A$ og $x \in B$ er falsk for alle x (jvf. Definition 93). ■

Bemærk at dette bevis følger den strategi for entydighedsbeviser, som blev beskrevet i forbindelse med formel (2.37).

Definition 93 *Den entydige mængde uden elementer kaldes den* ***tomme mængde*** *og betegnes med* $\emptyset$.

Bemærkning 94 *Der er forskel på* $\emptyset$ *og* $\{\emptyset\}$. *Mængden* $\emptyset$ *er den tomme mængde, og har ingen elementer. Mængden* $\{\emptyset\}$ *er derimod mængden af den tomme mængde. Den har ét element, nemlig* $\emptyset$.

Øvelse 95 *Hvilke elementer har mængderne:* $\{\emptyset, \{\emptyset\}\}$ *og* $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$*?*

Når man opgiver en mængde ved at angive alle dens elementer i en tuborgparamtes, siger man at mængden er givet på elementform. Denne form kan strengt taget kun bruges til at angive endelige mængder, og i praksis kun mængder med få elementer; men man bruger den også nogle gange til at betegne mængder med uendeligt mange elementer. For eksempel kan man skrive mængden af lige tal på formen $\{2, 4, 6, ...\}$, medens $\{2, 4, 6, ...100\}$ betegner de lige tal mindre eller lig 100. Det er klart at denne måde at betegne mængder kun kan bruges, når der ikke kan opstå tvivl om hvad "..." står for.

Notation 96 *Følgende notation bruges om forskellige talmængder:*

$\mathbb{N}$ *betegner mængden af de naturlige tal altså mængden* $\{1, 2, 3, ...\}$.

$\mathbb{Z}$ *betegner mængden af de hele tal altså mængden* $\{...-3, -2, -1, 0, 1, 2, 3, ...\}$.

$\mathbb{Q}$ *betegner mængden af rationale tal altså mængden af brøker* $\frac{a}{b}$ *hvor* $a, b \in \mathbb{Z}$ *og* $b \neq 0$.

$\mathbb{R}$ *betegner mængden af reelle tal.*

$\mathbb{C}$ *betegner mængden af komplekse tal*

$\mathbb{Z}_+$, $\mathbb{Q}_+$, $\mathbb{R}_+$ *betegner henholdsvist de positive hele tal, de positive rationale tal og de positive reelle tal.*

$\mathbb{Z}_-$, $\mathbb{Q}_-$, $\mathbb{R}_-$ *betegner henholdsvist de negative hele tal, de negative rationale tal og de negative reelle tal*

Sandhedsmængder. Hvis $p(x)$ er et prædikat om elementerne i en mængde M, kan man betragte mængden bestående af de af M's elementer, som gør prædikatet sandt. Denne mængde kaldes p's sandhedsmængde og betegnes symbolsk ved

$$\{x \in M \mid p(x)\}. \tag{6.3}$$

Man siger "mængden af de x i M for hvilke $p(x)$".

Eksempel 97 $\mathbb{Z}_+ = \{x \in \mathbb{Z} \mid x > 0\}$

Eksempel 98 $\mathbb{Q} = \{x \in \mathbb{R} \mid \exists a, b \in \mathbb{Z} : (b \neq 0) \wedge \left(x = \frac{a}{b}\right)\} = \{\frac{a}{b} \mid (a, b \in \mathbb{Z}) \wedge (b \neq 0)\}$

Hvis det er helt klart hvilken grundmængde den frie variabel x tænkes at løbe over, tillader man sig at skrive $\{x \mid p(x)\}$. Men man skal være forsigtig med denne skrivemåde. Mængdelærens aksiomer tillader ikke, at man danner mængden af al ting, som opfylder et bestemt prædikat. Vi skal senere se hvorfor.

Notation 99 ***Intervaller.*** *Man bruger følgende skrivemåde for intervaller af reelle tal (her er $a, b \in \mathbb{R}$):*

$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$ kaldes det lukkede interval fra a til b.
$]a, b[= \{x \in \mathbb{R} \mid a < x < b\}$ kaldes det åbne interval fra a til b.
$[a, b[= \{x \in \mathbb{R} \mid a \leq x < b\}$ og $]a, b] = \{x \in \mathbb{R} \mid a < x \leq b\}$ kaldes halvåbne intervaller.
$[a, \infty[= \{x \in \mathbb{R} \mid a \leq x\}$ og $]-\infty, a] = \{x \in \mathbb{R} \mid x \leq a\}$ kaldes lukkede.
$]a, \infty[= \{x \in \mathbb{R} \mid a < x\}$ og $]-\infty, a[= \{x \in \mathbb{R} \mid x < a\}$ kaldes åbne.
I engelsksproget literatur er det mest almindeligt at betegne åbne intervaller med runde parenteser. F.eks. betegner (a, b) det åbne interval, som vi ovenfor har betegnet ved $]a, b[$.

6.2 Delmængder

Definition 100 *En mængde A kaldes en* ***delmængde*** *af en mængde B (eller at A er indeholdt i B), hvis alle elementerne i A også ligger i B. I så fald siger man også at A er indeholdt i B, eller at B indeholder A. Man skriver da $A \subseteq B$ eller $B \supseteq A$.*
Med andre ord $A \subseteq B$, hvis

$$x \in A \Rightarrow x \in B. \tag{6.4}$$

Bevisstrategi. Når man skal vise at $A \subseteq B$, skal man altså vise at et vilkårligt element i A ligger i B. Beviset begynder derfor med ordene: "Lad $x \in A$" og fortsætter med at vise, at vi ud fra $x \in A$ kan slutte at $x \in B$. Antag nu at mængderne er givet som sandhedsmængder for to prædikater: $A = \{x \in M \mid p(x)\}$ og $B = \{x \in M \mid q(x)\}$. Beviset for $A \subseteq B$ vil da forløbe således: "Lad $x \in A$. Så vil $p(x)$ være opfyldt" så argumenteres for at $p(x) \Rightarrow q(x)$ hvorefter der sluttes: "Altså er $q(x)$ sand hvorfor $x \in B$".

Eksempel 101 *Vis, at $\{x \in \mathbb{R} \mid x^2 < 2\} \subseteq \{x \in \mathbb{R} \mid x < 5\}$.*

Bevis. Antag at $x \in \{x \in \mathbb{R} \mid x^2 < 2\}$, altså at $x^2 < 2$. Da gælder, at $x^2 < 4$, hvorfor $-2 < x < 2$. Heraf sluttes, at $x < 5$, altså at $x \in \{x \in \mathbb{R} \mid x < 5\}$. ■
Vi kan præcisere ovenstående bevisstrategi til en sætning:

Sætning 102 *Udsagnet $p(x) \Rightarrow q(x)$ er ensbetydende med udsagnet $\{x \in M \mid p(x)\} \subseteq \{x \in M \mid q(x)\}$*

Bevis. .

$$\begin{aligned} \{x &\in M \mid p(x)\} \subseteq \{x \in M \mid q(x)\} && (6.5)\\ &\Updownarrow && (6.6)\\ x &\in \{x \in M \mid p(x)\} \Rightarrow x \in \{x \in M \mid q(x)\} && (6.7)\\ &\Updownarrow && (6.8)\\ p(x) &\Rightarrow q(x) && (6.9) \end{aligned}$$

■

Sætning 103 *Den tomme mængde er en delmængde af enhver mængde. Altså hvis A er en mængde gælder*

$$\emptyset \subseteq A \tag{6.10}$$

Bevis. Vi skal vise at

$$x \in \emptyset \Rightarrow x \in A. \tag{6.11}$$

Men udsagnet $x \in \emptyset$ er falsk for alle x, hvorfor udsagnet $x \in \emptyset \Rightarrow x \in A$ er sandt (se bemærkning definition 8 og bemærkning 24). ■

Bemærk at når $A \subseteq B$ kan A og B godt være ens. Hvis vi vil udelukke det, skriver vi $A \subset B$:

Definition 104 *A kaldes en **ægte delmængde** af B hvis $A \subseteq B$ og $A \neq B$. Vi skriver da $A \subset B$.*

Sætning 105 *Lad A og B være mængder. Da er $A = B$ hvis og kun hvis $(A \subseteq B) \wedge (B \subseteq A)$.*

Bevis. Ifølge Definition 91 og 100 og Eksenpel 12 gælder følgende biimplikationer:

$$\begin{aligned} A &= B && (6.12)\\ &\Updownarrow && (6.13)\\ x &\in A \Leftrightarrow x \in B && (6.14)\\ &\Updownarrow && (6.15)\\ (x &\in A \Rightarrow x \in B) \wedge (x \in A \Leftarrow x \in B) && (6.16)\\ &\Updownarrow && (6.17)\\ (A &\subseteq B) \wedge (B \subseteq A) && (6.18) \end{aligned}$$

■

Bevisstrategi. Når man skal vise at to mængder A og B er lig med hinanden beviser man ofte først $A \subseteq B$ og dernæst $B \subseteq A$.

Sætning 106 *Mængden af punkter i planen, som ligger lige langt fra to givne punkter A og B, er midtnormalen til linjestykket AB.*[1]

Bevis. Vi viser først at mængden af punkter i planen, som ligger lige langt fra A og B, er indeholdt i midtnormalen. Lad derfor C være et punkt, som ligger lige langt fra A og B. Tegn trekant ABC. Halver linjestykket AB i D og tegn CD. Nu er siderne i ΔCAD og ΔCBD parvist lige store, hvorfor de to trekanter er kongruente (Euklid I.8). Men så er $\angle CDA = \angle CDB$, hvorfor de ifølge definitionen på en ret vinkel (Euklid Def. 10) må være rette. Men det betyder at linjen CD er midtnormalen til AB. Altså ligger C på denne midtnormal.

Dernæst viser vi omvendt, at ethvert punkt på AB's midtnormal ligger lige langt fra A og B. Antag altså at C ligger på AB's midtnormal, som skærer AB i midtpunktet, som vi kalder D. Tegn trekant ABC. Ifølge Euklid I.4 er ΔCDA kongruent med ΔCDB, thi $\angle CDA = \angle CDB$, og de to hosliggende sider er også parvist lige store. Men så er $CA = CB$, så C ligger lige langt fra A og B. ■

Bemærkning 107 *I geometri kaldes mængden af punkter, som opfylder en bestemt egenskab, for "det geometriske sted for de punkter, der opfylder egenskaben". Vi har altså bevist at det geometriske sted for de punkter i planen, som ligger lige langt fra to punkter, er midtnormalen til linjestykket mellem punkterne.*

Sætning 108 *Lad A, B og C være mængder. Hvis $A \subseteq B$ og $B \subseteq C$, så gælder $A \subseteq C$.*

Bevis. Antag at $A \subseteq B$ og $B \subseteq C$. Vi skal vise, at $A \subseteq C$, altså at $x \in A \Rightarrow x \in C$. Lad derfor $x \in A$. Vi skal da vise at $x \in C$.

Da $A \subseteq B$, kan vi af $x \in A$ slutte, at $x \in B$, og da $B \subseteq C$, slutter vi videre at $x \in C$. ■

Notation 109 *Hvis $A \subseteq B$ og $B \subseteq C$ tillader man sig derfor at skrive $A \subseteq B \subseteq C$. På lignende måde kan man skrive $A \subset B \subseteq C$, $A \subseteq B \subset C$ og $A \subset B \subset C$. Betydningen heraf er klar. Derimod skriver man ikke strenge hvor inklusionstegnene vender hver sin vej (f.eks. $A \supseteq B \subseteq C$).*

Venn-diagrammer. Man kan illustrere mængder i et såkaldt Venn-diagram. Figur 6.1 illustrerer situationen $A \subseteq B \subseteq C$. Punkterne inden for bollerne forestiller elementerne i mængden. Sætning 108 kan nærmest aflæses ud af figuren. En sådan figurinspektion kan dog formelt ikke gøre det ud for et bevis, men den kan give gode ideer til hvilke formodninger, man skal opstille. Venndiagrammer giver også en god intuition om situationen, så det anbefales, at du så vidt muligt laver diagrammer, der illustrerer de følgende sætninger om mængder.

[1] Midtnormalen til et linjestykke er en ret linje, som står vinkelret på midten af linjestykket.

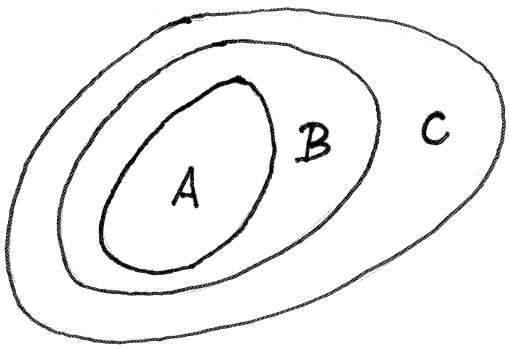

Figure 6.1: Venn-diagram, som illustrerer: $A \subseteq B \subseteq C$

6.3 Fællesmængde og foreningsmængde

Definition 110 *Lad A og B være to mængder. Mængden af de elementer, som ligger i både A og B kaldes for* ***fællesmængden*** *for A og B. Den betegnes $A \cap B$. Med andre ord*

$$A \cap B = \{x \mid (x \in A) \wedge (x \in B)\} \tag{6.19}$$

eller

$$x \in A \cap B \Leftrightarrow (x \in A) \wedge (x \in B) \tag{6.20}$$

Bevisstrategi: Når man skal vise at et element ligger i fællesmængden for to mængder skal man altså vise at det ligger i begge de to mængder.

Eksempel 111 $\{a, b, c, d, e, f\} \cap \{d, e, f, g, h, i\} = \{d, e, f\}$.

Eksempel 112 $\mathbb{Z} \cap \mathbb{R}_+ = \mathbb{Z}_+$.

Eksempel 113 *Lad AB være et linjestykke med midtpunkt D, og lad DC betegne dens midtnormal. Da er $AB \cap DC = D$*

På Venn-diagrammet i Figur 6.2 er fællesmængden for A og B krydsskraveret

Lad A, B og C være mængder. Da gælder

$$A \cap B = B \cap A \tag{6.21}$$

og

$$(A \cap B) \cap C = A \cap (B \cap C). \tag{6.22}$$

Bevis. Følger af de tilsvarende logiske regler for " $\wedge$ " ■

På grund af (6.22) giver det ikke anledning til misforståelser at skrive $A \cap B \cap C$ og tilsvarende for flere mængder.

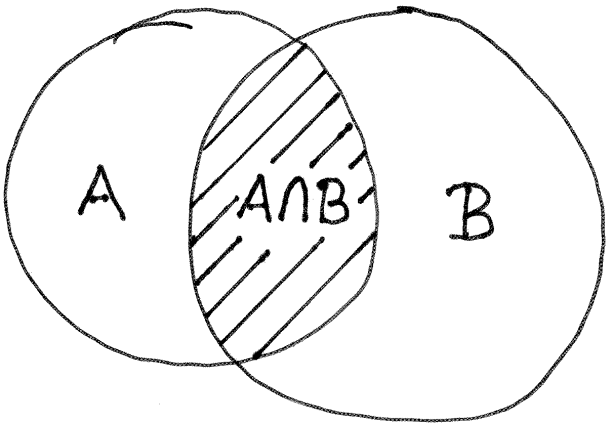

Figure 6.2: Fællesmængden $A \cap B$

Sætning 114 *Lad A og B være mængder. Da gælder:*

$$A \cap A = A, \quad A \cap B \subseteq A, \text{ og } A \cap \emptyset = \emptyset \tag{6.23}$$

Definition 115 *Lad A og B være mængder. Hvis $A \cap B = \emptyset$, siges A og B at være **disjunkte.***

Definition 116 *Lad A og B være to mængder. Mængden af elementer, som ligger i enten A eller B kaldes **foreningsmængden** af (eller for) A og B. Den betegnes med $A \cup B$. Med andre ord*

$$A \cup B = \{x \mid (x \in A) \vee (x \in B)\} \tag{6.24}$$

eller

$$x \in A \cup B \Leftrightarrow (x \in A) \vee (x \in B) \tag{6.25}$$

Bevisstrategi: Når man skal vise at et element ligger i foreningsmængden af to mængder skal man altså vise at det ligger i mindst en af de to mængder.

Eksempel 117 $\{a, b, c, d, e, f\} \cup \{d, e, f, g, h, i\} = \{a, b, c, d, e, f, g, h, i\}$.

Eksempel 118 $(\mathbb{Z}_+ \cup \mathbb{Z}_-) \cup \{0\} = \mathbb{Z}$

På Venn-diagrammet i Figur 6.3 repræsenterer hele det skraverede område foreningsmængden af A og B.

Sætning 119 *Lad A, B og C være mængder. Da gælder*

$$A \cup B = B \cup A \tag{6.26}$$

og

$$(A \cup B) \cup C = A \cup (B \cup C). \tag{6.27}$$

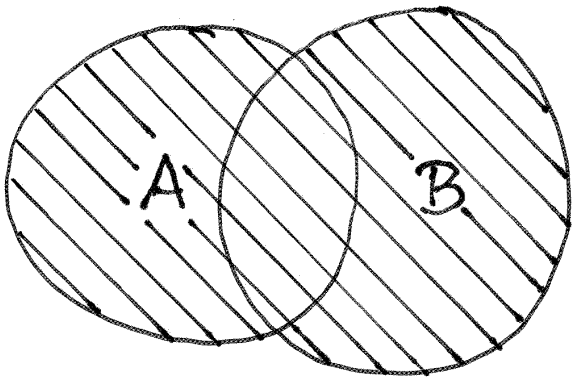

Figure 6.3: Foreningsmængden $A \cup B$

Bevis. Følger af de tilsvarende logiske regler for " $\vee$ " . ■

På grund af (6.27) giver det ikke anledning til misforståelser at skrive $A \cup B \cup C$ og tilsvarende for flere mængder.

Sætning 120 *Lad A og B være mængder. Da gælder:*

$$A \cup A = A, \quad A \cup B \supseteq A, \text{ og } A \cup \emptyset = A \tag{6.28}$$

Man kan endog danne foreningsmængde og fællesmængde af en hel (eventuelt uendelig) familie af mængder.

6.3.1 Familier af mængder.

Eksempel 121 *Betragt intervallerne $I_1 = [0,1]$, $I_2 = \left[0, \frac{1}{2}\right]$, $I_3 = \left[0, \frac{1}{3}\right]$, ..., $I_n = \left[0, \frac{1}{n}\right]$,.... De udgør en familie af delmængder af $\mathbb{R}$. De er indiceret ved de naturlige tal, i den forstand, at der til hvert naturligt tal svarer netop et af intervallerne (til n svarer I_n).*

Eksempel 122 *Betragt enhedscirklerne i planen med centrum i et punkt på x-aksen. Lad C_t betegne den enhedscirkel, hvis centrum har koordinaterne $(t, 0)$. Da udgør alle C_t 'erne en familie af cirkler indiceret ved det reelle index t.*

Definition 123 *Mere generelt, hvis Λ er en mængde (af indices), og der til ethvert index $\alpha \in \Lambda$ svarer en mængde A_α, da siger vi at A_α'erne udgør en familie af mængder, der er indiceret ved Λ. Familien betegnes ved $\{A_\alpha\}_{\alpha \in \Lambda}$.*

Definition 124 *Lad $\{A_\alpha\}_{\alpha \in \Lambda}$ være en familie af mængder. Fællesmængden for alle familiens mængder ("fællesmængden for A_α'erne for α i Λ") betegnes*

med $\bigcap_{\alpha\in\Lambda} A_\alpha$ *og defineres formelt ved:*

$$\bigcap_{\alpha\in\Lambda} A_\alpha = \{x \mid \forall\alpha \in \Lambda : x \in A_\alpha\}\,. \tag{6.29}$$

Med andre ord $x \in \bigcap_{\alpha\in\Lambda} A_\alpha$*, hvis* $x \in A_\alpha$ *for alle* $\alpha \in \Lambda$*. I tilfælde af at indexmængden er* $\{1, 2, 3, ..., m\}$ *eller* $\mathbb{N}$*, skriver man også* $\bigcap_{n=1}^{m}$ *og* $\bigcap_{n=1}^{\infty}$*.*

Bemærkning 125 *Ved negation af definitionen ses at* $x \notin \bigcap_{\alpha\in\Lambda} A_\alpha$*, hvis og kun hvis* $\exists\alpha \in \Lambda : x \notin A_\alpha$

Eksempel 126 *Betragt familien af intervaller* $\{I_n\}_{n\in\mathbb{N}}$ *defineret i eksempel 121. Vi vil vise at* $\bigcap_{n\in\mathbb{N}} I_n = \{0\}$*.*

Bevis. $\supseteq$: Det er let at vise at $\bigcap_{n\in\mathbb{N}} I_n \supseteq \{0\}$. Vi skal vise at ethvert element i $\{0\}$ ligger i $\bigcap_{n\in\mathbb{N}} I_n$ altså i I_n for alle $n \in \mathbb{N}$. Men det eneste element i $\{0\}$ er 0, og $0 \in I_n = \left[0, \frac{1}{n}\right]$ for alle $n \in \mathbb{N}$.

$\subseteq$: For at vise at $\bigcap_{n\in\mathbb{N}} I_n \subseteq \{0\}$ skal vi vise at

$$x \in \bigcap_{n\in\mathbb{N}} I_n \Rightarrow x \in \{0\}\,, \tag{6.30}$$

eller med andre ord at

$$x \in \bigcap_{n\in\mathbb{N}} I_n \Rightarrow x = 0. \tag{6.31}$$

Det vil vi vise ved kontraposition, så vi vil altså vise at

$$x \neq 0 \Rightarrow x \notin \bigcap_{n\in\mathbb{N}} I_n \tag{6.32}$$

Antag altså at $x \neq 0$. Det betyder at $x < 0$ eller $x > 0$. Vi behandler de to tilfælde hver for sig og skal altså (iflg. bemærkning 125) vise, at under hver af de to antagelser eksisterer der et $n \in \mathbb{N}$, så $x \notin \bigcap_{n\in\mathbb{N}} I_n$.

Antag altså først at $x < 0$. Da gælder for alle $n \in \mathbb{N}$, at $x \notin I_n$, og så meget mere eksisterer der da et $n \in \mathbb{N}$, så $x \notin I_n$.

Antag dernæst at $x > 0$. Bestem da $n \in \mathbb{N}$, så $n > \frac{1}{x}$. At dette er muligt, vil vi senere bevise stringent. Men så er $x > \frac{1}{n}$, hvoraf følger at $x \notin I_n = \left[0, \frac{1}{n}\right]$. Altså har vi vist, at der findes et $n \in \mathbb{N}$ så $x \notin I_n$. ■

Bemærkning 127 *I dette bevis har vi brugt en stor del af det maskineri, vi har bygget op tidligere. Prøv at lokaliser alle de tidligere definitioner og bevisstrategier, vi har haft i gang.*

Definition 128 *Lad $\{A_\alpha\}_{\alpha\in\Lambda}$ være en familie af mængder. Foreningsmængden af alle familiens mængder ("foreningsmængden af A_α'erne for α i Λ) betegnes med $\bigcup_{\alpha\in\Lambda} A_\alpha$ og defineres formelt ved:*

$$\bigcup_{\alpha\in\Lambda} A_\alpha = \{x \mid \exists\alpha \in \Lambda : x \in A_\alpha\} \tag{6.33}$$

Med andre ord $x \in \bigcup_{\alpha\in\Lambda} A_\alpha$, hvis der findes et $\alpha \in \Lambda$, så $x \in A_\alpha$. I tilfælde af at indexmængden er $\{1, 2, 3, ..., m\}$ eller $\mathbb{N}$ skriver man også $\bigcup_{n=1}^{m}$ og $\bigcup_{n=1}^{\infty}$.

Bemærkning 129 *Ved negation af definitionen ses at $x \notin \bigcup_{\alpha\in\Lambda} A_\alpha$, hvis og kun hvis $\forall\alpha \in \Lambda; x \notin A_\alpha$.*

Eksempel 130 *Betragt familien af intervaller defineret ved $J_n = \left[0, 1 - \frac{1}{n}\right]$ for $n \in \mathbb{N}$. Vi vil vise at $\bigcup_{n=1}^{\infty} J_n = [0, 1[$.*

Bevis. $\subseteq$: Antag først at $x \in \bigcup_{n=1}^{\infty} J_n$. Vi skal vise, at $x \in [0, 1[$. Når $x \in \bigcup_{n=1}^{\infty} J_n$ betyder det pr. definition, at der findes et $n \in \mathbb{N}$, så $x \in J_n$. Men da $J_n = \left[0, 1 - \frac{1}{n}\right] \subseteq [0, 1[$, ses heraf, at $x \in [0, 1[$.

$\supseteq$: Antag dernæst at $x \in [0, 1[$. Vi skal vise at $x \in \bigcup_{n=1}^{\infty} J_n$, altså at $\exists n \in \mathbb{N} :$ $x \in J_n$. Når $x \in [0, 1[$, gælder specielt at $x < 1$ eller at $0 < 1 - x$, så vi kan bestemme et $n \in \mathbb{N}$, så $\frac{1}{1-x} < n$. Men da er $\frac{1}{n} < 1 - x$ eller $x < 1 - \frac{1}{n}$. Da vi endvidere havde antaget, at $0 < x$, følger det, at $x \in J_n = \left[0, 1 - \frac{1}{n}\right]$. ■

Bemærk at foreningsmængden af en uendelig familie af lukkede intervaller ikke behøver at være lukket.

Bemærkning 131 *I sidste del af beviset bestemte jeg pludseligt et n så $\frac{1}{1-x} < n$. Ved første blik kunne det se ud som en kanin, der blev trukket op af hatten. Men naturligvis er denne ide fremkommet ved en analyse, hvor vi begynder med det vi vil vise: Vi vil bestemme n så $x \in J_n$. Da vi jo har antaget at $x \in [0, 1[$ skal vi bare sørge for, at $x < 1 - \frac{1}{n}$. Men det betyder, at $\frac{1}{n} < 1 - x$, eller at $\frac{1}{1-x} < n$ (her bruges at $1 - x > 0$). Og det kan netop lade sig gøre, fordi $x \neq 1$. Syntesen, som blev præsenteret i beviset, er bare analysen kørt baglæns.*

Bemærkning 132 *Vi kunne også i beviserne for eksemplerne 126 og 130 have brugt at $\frac{1}{n} \to 0$ for $n \to \infty$. Men da det ikke ville have simplificeret beviset, er det bedre at undgå brugen af dette resultat.*

Sætning 133 *Lad $\{A_\alpha\}_{\alpha\in\Lambda}$ være en familie af mængder, og lad A være en mængde.*

a. Hvis $A \subseteq A_\alpha$ for alle $\alpha \in \Lambda$, da gælder at $A \subseteq \bigcap_{\alpha\in\Lambda} A_\alpha$.

b. Hvis $A_\alpha \subseteq A$ for alle $\alpha \in \Lambda$, da gælder at $\bigcup_{\alpha\in\Lambda} A_\alpha \subseteq A$.

Bevis. Overlades til læseren. ■

6.4 Mængdedifferens og komplementærmængde

Definition 134 *Lad A og B være mængder. Mængden af de elementer i A, som ikke er element i B, kaldes mængdedifferensen mellem A og B. Den betegnes med $A \setminus B$. Med andre ord:*

$$A \setminus B = \{x \mid (x \in A) \wedge (x \notin B)\} \tag{6.34}$$

eller

$$x \in A \setminus B \Leftrightarrow (x \in A) \wedge (x \notin B) \tag{6.35}$$

Eksempel 135 $\{a, b, c, d, e, f\} \setminus \{d, e, f, g, h, i\} = \{a, b, c\}$.

Eksempel 136 $\mathbb{Z} \setminus \mathbb{Z}_- = \mathbb{Z}_+ \cup \{0\}$.

På Venn-diagrammet i Figur 6.4 er $A \setminus B$ skraveret.

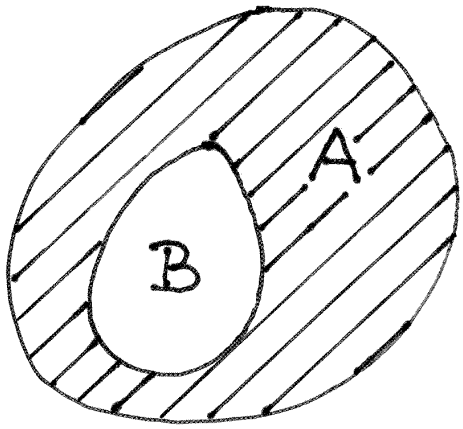

Figure 6.4: Mængdedifferensen $A \setminus B$

Sætning 137 *Lad A og B være mængder. Da gælder:*

$$A \setminus B \subseteq A, \quad A \setminus A = \emptyset \quad og \quad A \setminus \emptyset = A \tag{6.36}$$

Øvelse 138 *Bevis at $A \setminus B$, $B \setminus A$ og $A \cap B$ er disjunkte mængder hvis foreningsmængde er lig med $A \cup B$.*

Det er ofte naturligt at betragte mængder, som alle er delmængder af en fast mængde, som vi da kalder **grundmængden**. For eksempel, når man arbejder med reel analyse er talmængderne alle delmængder af grundmængden $\mathbb{R}$, og i

plangeometrien er punktmængderne alle delmængder af planen, der altså spiller rollen af grundmængde.

Lad os nu se på en sådan situation, hvor alle betragtede mængder er delmængder af en grundmængde, som vi kalder U (Vi bruger bogstavet U, fordi grundmængden kaldes "the universe" på engelsk).

Definition 139 *Lad A være en delmængde af en grundmængde U. Da kaldes $U \setminus A$ for **komplementærmængden** til A, og den betegnes med $\complement A$.*

$$\complement A = U \setminus A = \{x \in U \mid x \notin A\} \tag{6.37}$$

Eksempel 140 *Inden for $\mathbb{R}$ gælder:*

$$\complement\, [a, \infty[\;=\;]-\infty, a[\tag{6.38}$$
$$\complement\,]a, b[\;=\;]-\infty, a] \cup [b, \infty[\tag{6.39}$$
$$\complement \mathbb{R}_+ \;=\; \mathbb{R}_- \cup \{0\}\,. \tag{6.40}$$

Sætning 141 *Inden for grundmængden U gælder:*

$$\complement U = \emptyset \quad og \quad \complement \emptyset = U. \tag{6.41}$$

og for en vilkårlig mængde A

$$\complement\complement A = A \tag{6.42}$$

Bemærkning 142 *Når man tager komplementærmængden til en mængde, er det vigtigt at vide, hvad grundmængden er. Man kunne måske tro at man kunne tage en helt universel grundmængde, altså mængden af al ting, så $\complement A$ kunne betyde alt der ikke er element i A. Det viser sig dog at det fører til problemer (se afsnit 6.8)*

6.5 Mængdealgebra

Vi har i det foregående indført en række operationer på mængder, som minder om regneoperationerne på de reelle tal. Foreningsmængde $\cup$ svarer på en ret oplagt måde til addition $+$, fællesmængde $\cap$ svarer, på en mindre oplagt måde, til multiplikation $\cdot$, og mængdedifferens $\setminus$ svarer til differens $-$. På samme måde som regneoperationerne på $\mathbb{R}$ opfylder visse regneregler, så er der lignende regneregler for mængdeoperationerne. Dem skal vi udlede i dette afsnit.

Vi starter med to regneregler, som svarer til den distributive lov, som jo siger at

$$a \cdot (b + c) = a \cdot b + a \cdot c \tag{6.43}$$

Sætning 143 ***De distributive love.*** *Lad A, B og C være mængder. Da gælder*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \tag{6.44}$$

og

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C). \tag{6.45}$$

Bevis. Lad os vise den sidste identitet (6.45). Den første vises analogt.

For vilkårligt x gælder følgende biimplikationer:

$$\begin{aligned} x &\in A \cup (B \cap C) && (6.46)\\ &\Leftrightarrow (x \in A) \vee (x \in B \cap C) && (6.47)\\ &\Leftrightarrow (x \in A) \vee ((x \in B) \wedge (x \in C) && (6.48)\\ &\Leftrightarrow ((x \in A) \vee (x \in B)) \wedge ((x \in A) \vee (x \in C)) && (6.49)\\ &\Leftrightarrow (x \in A \cup B) \wedge (x \in A \cup C) && (6.50)\\ &\Leftrightarrow x \in (A \cup B) \cap (A \cup C). && (6.51) \end{aligned}$$

(Ved overgang fra (6.48) til (6.49) brugte vi den logiske regel 1.15). Heraf følger (6.45). ■

Bemærk at det her gav et mere overskueligt bevis at opskrive en række biimplikationer, end det ville have været at vise de to inklusioner $A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C)$ og $A \cup (B \cap C) \supseteq (A \cup B) \cap (A \cup C)$ hver for sig.

Bemærkning 144 *Der er ikke helt analogi mellem de algebraiske operationer og mængdeoperationerne. For mængdeoperationerne gælder begge de distributive love (6.44) og(6.45). For de algebraiske regneoperationer gælder kun (6.43), hvorimod udsagnet $a + (b \cdot c) = (a + b)(a + c)$ ikke generelt er sandt.*

Øvelse 145 *Tegn Venn-diagrammer der illustrerer de distributive love.*

De distributive love kan generaliseres til familier af mængder:

Sætning 146 ***De distributive love.*** *Lad A være en mængde og $\{B_\alpha\}_{\alpha\in\Lambda}$ en familie af mængder. Da gælder:*

$$A \cap (\bigcup_{\alpha\in\Lambda} B_\alpha) = \bigcup_{\alpha\in\Lambda} (A \cap B_\alpha) \tag{6.52}$$

og

$$A \cup (\bigcap_{\alpha\in\Lambda} B_\alpha) = \bigcap_{\alpha\in\Lambda} (A \cup B_\alpha) \tag{6.53}$$

Bevis. Denne gang viser vi den første identitet (6.52). Den sidste vises analogt.

For vilkårligt x gælder der følgende biimplikationer:

$$\begin{aligned} x &\in A \cap (\bigcup_{\alpha\in\Lambda} B_\alpha) && (6.54)\\ &\Leftrightarrow (x \in A) \wedge (x \in \bigcup_{\alpha\in\Lambda} B_\alpha) && (6.55)\\ &\Leftrightarrow (x \in A) \wedge (\exists \alpha \in \Lambda : x \in B_\alpha) && (6.56)\\ &\Leftrightarrow \exists \alpha \in \Lambda : (x \in A) \wedge (x \in B_\alpha) && (6.57)\\ &\Leftrightarrow \exists \alpha \in \Lambda : x \in A \cap B_\alpha && (6.58)\\ &\Leftrightarrow x \in \bigcup_{\alpha\in\Lambda} (A \cap B_\alpha) && (6.59) \end{aligned}$$

(Ved overgang fra (6.56) til (6.57) brugtes en variant af formel (1.35)) Heraf følger (6.52). ■

Sætning 147 ***De Morgans love***[2]*. Lad X være en mængde og lad $\{B_\alpha\}_{\alpha\in\Lambda}$ være en familie af delmængder af X. Da gælder:*

$$X \setminus (\bigcup_{\alpha\in\Lambda} B_\alpha) = \bigcap_{\alpha\in\Lambda} (X \setminus B_\alpha) \tag{6.60}$$

og

$$X \setminus \bigcap_{\alpha\in\Lambda} B_\alpha = \bigcup_{\alpha\in\Lambda} (X \setminus B_\alpha) \tag{6.61}$$

Bevis. Vi beviser kun (6.61). Beviset for (6.60) er analogt.

For vilkårligt x har vi følgende biimplikationer:

$$\begin{aligned}
x \quad &\in \quad X \setminus \bigcap_{\alpha\in\Lambda} B_\alpha && (6.62)\\
&\Leftrightarrow \quad (x \in X) \wedge (\neg(x \in \bigcap_{\alpha\in\Lambda} B_\alpha)) && (6.63)\\
&\Leftrightarrow \quad (x \in X) \wedge (\neg(\forall\alpha \in \Lambda : x \in B_\alpha)) && (6.64)\\
&\Leftrightarrow \quad (x \in X) \wedge (\exists\alpha \in \Lambda : \neg(x \in B_\alpha)) && (6.65)\\
&\Leftrightarrow \quad \exists\alpha \in \Lambda : (x \in X) \wedge (\neg(x \in B_\alpha)) && (6.66)\\
&\Leftrightarrow \quad \exists\alpha \in \Lambda : (x \in X \setminus B_\alpha) && (6.67)\\
&\Leftrightarrow \quad x \in \bigcup_{\alpha\in\Lambda} (X \setminus B_\alpha) && (6.68)
\end{aligned}$$

Heraf følger (6.61) ■

Hvis vi specielt lader X være grundmængden U, så kan De Morgans love skrives på formen:

Sætning 148 *Lad $\{B_\alpha\}_{\alpha\in\Lambda}$ være en familie af delmængder af grundmængden U. Da gælder:*

$$\complement(\bigcup_{\alpha\in\Lambda} B_\alpha) = \bigcap_{\alpha\in\Lambda} (\complement B_\alpha) \tag{6.69}$$

og

$$\complement(\bigcap_{\alpha\in\Lambda} B_\alpha) = \bigcup_{\alpha\in\Lambda} (\complement B_\alpha) \tag{6.70}$$

Specielt hvis familien består af to delmængder, kan De Morgans love formuleres som følger:

Sætning 149 *Lad X, A og B være mængder. Da gælder:*

$$X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B) \tag{6.71}$$

og

$$X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B) \tag{6.72}$$

Øvelse 150 *Tegn et Venn-diagram, der illustrerer disse specialtilfælde af De Morgans love.*

[2] Efter Augustus de Morgan 1806-1871

Bemærkning 151 *Ovenfor blev der brugt to forskellige måder at præsentere to resultater, hvoraf den ene er en generalisering af den anden. Ved præsentation af de distributive love formulerede jeg først det specielle tilfælde i sætning 143 og generaliserede det derefter i sætning 146. Ved præsentation af De Morgans love beviste jeg straks det generelle resultat (Sætning 147), hvorefter de specielle resultater i sætning 149 faldt ud som specialtilfælde. Den første metode kan have pædagogiske fordele, medens den anden er den korteste og matematisk mest elegante.*

Sætning 152 *Lad A, B og C være mængder. Da gælder:*

$$B \setminus (B \setminus A) = A \cap B \tag{6.73}$$

og

$$(A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B) \tag{6.74}$$

Øvelse 153 *Illustrer disse identiteter i et Venn-diagram og bevis dem formelt.*

6.6 Produktmængde

Definition 154 *Et **ordnet par** (a, b) er et par i en bestemt rækkefølge. Det betyder at to ordnede par er lig hinanden, hvis og kun hvis de indeholder de samme objekter i samme rækkefølge:*

$$(a_1, b_1) = (a_2, b_2) \Leftrightarrow (a_1 = a_2) \wedge (b_1 = b_2). \tag{6.75}$$

Bemærkning 155 *Det er vigtigt skældne mellem mængden $\{a, b\}$ og det ordnede par (a, b). Hvis $a \neq b$ er $(a, b) \neq (b, a)$ medens $\{a, b\} = \{b, a\}$. Desuden er der mening i at tale om talparet (a, a), hvorimod $\{a, a\}$ blot er en besværlig måde at skrive $\{a\}$.*[3]

Definition 156 *Lad A og B være to givne mængder. Vi betragter da ordnede par (a, b), hvor a på førstepladsen er et element fra A og b på andenpladsen er et element fra B. Mængden af sådanne par kaldes **produktmængden** (eller det cartesiske produkt) af A og B og betegnes med $A \times B$.*

$$A \times B = \{(a, b) \mid (a \in A) \wedge (b \in B)\} \tag{6.76}$$

Eksempel 157 *Lad $A = \{1, 2\}$ og $B = \{a, b, c\}$. Da består $A \times B$ af følgende ordnede par:*

$$\begin{array}{ccc} c & (1,c) & (2,c) \\ b & (1,b) & (2,b) \\ a & (1,a) & (2,a) \\ & 1 & 2 \end{array} \tag{6.77}$$

Altså $A \times B = \{(1, a), (1, b), (1, c), (2, a), (2, b), (2, c)\}$

[3] Det er dog muligt at definere et ordnet par (a, b) udelukkende i mængdeteoretiske termer. Man definerer det da som $\{\{a\}, \{a, b\}\}$

Notation 158 *Man kan naturligvis lade $A = B$ i definitionen af produktmængden. I så fald skriver man ofte A^2 i stedet for $A \times A$.*

Eksempel 159 *Bestem A^2 og B^2 når A og B har samme betydning som i eksempel 157*

Eksempel 160 *$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ består af reelle talpar (a, b) hvor $a, b \in \mathbb{R}$. Hvis man i planen indlægger to på hinanden vinkelrette tallinjer, kan man på velkendt og entydig vis tilordne et bestemt talpar fra $\mathbb{R}^2$ til ethvert punkt i planen.*

Øvelse 161 *Opskriv følgende delmængde af $\mathbb{R}^2$: $[1, 4] \times]-1, 3[$ på formen $\{(x, y) \mid ....\}$ og illustrer mængden i talplanen.*

Definition 162 *Produktmængden $A \times B \times C$ af tre givne mængder A, B og C defineres tilsvarende ved*

$$A \times B \times C = \{(a, b, c) \mid (a \in A) \wedge (b \in B) \wedge (c \in C)\}. \tag{6.78}$$

og så videre for flere mængder.

Sætning 163 *Lad A, B og C være mængder. Da gælder*

$$\begin{aligned}
A \times (B \cap C) &= (A \times B) \cap (A \times C), && (6.79)\\
A \times (B \cup C) &= (A \times B) \cup (A \times C), && (6.80)\\
(A \cap B) \times C &= (A \times C) \cap (B \times C), && (6.81)\\
(A \cup B) \times C &= (A \times C) \cup (B \times C), && (6.82)\\
A \times (B \setminus C) &= (A \times B) \setminus (A \times C), && (6.83)\\
(A \setminus B) \times C &= (A \times C) \setminus (B \times C), && (6.84)\\
A &\subseteq B \Rightarrow A \times C \subseteq B \times C, && (6.85)\\
A &\subseteq B \Rightarrow C \times A \subseteq C \times B && (6.86)
\end{aligned}$$

Bevis. Vi viser kun (6.80). Resten overlades til læseren.

For et vilkårligt talpar (x, y) har vi følgende biimplikationer:

$$\begin{aligned}
(x, y) &\in A \times (B \cup C) && (6.87)\\
&\Leftrightarrow (x \in A) \wedge (y \in B \cup C) && (6.88)\\
&\Leftrightarrow (x \in A) \wedge ((y \in B) \vee (y \in C)) && (6.89)\\
&\Leftrightarrow ((x \in A) \wedge (y \in B)) \vee ((x \in A) \wedge (y \in C)) && (6.90)\\
&\Leftrightarrow ((x, y) \in A \times B) \vee ((x, y) \in A \times C) && (6.91)\\
&\Leftrightarrow (x, y) \in (A \times B) \cup (A \times C) && (6.92)
\end{aligned}$$

Heraf følger (6.80). ■

6.7 Potensmængden

Som vi allerede har bemærket, kan mængder selv være elementer i andre mængder. En familie af mængder $\{B_\alpha\}_{\alpha\in\Lambda}$ er et eksempel på en mængde af mængder. Man betragter ofte en mængde af delmængder af en given mængde. Mængden der består af alle delmængderne af en given mængde kaldes dens potensmængde.

Definition 164 *Lad A være en mængde. Mængden af alle A's delmængder kaldes A 's potensmængde og betegnes med $P(A)$.*

Eksempel 165 *Betragt mængden $\{1,2,3\}$. Den har følgende delmængder: $\emptyset$, $\{1\}$, $\{2\}$, $\{3\}$, $\{1,2\}$, $\{1,3\}$, $\{2,3\}$, $\{1,2,3\}$. Altså er*

$$P(\{1,2,3\}) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\} \tag{6.93}$$

Eksempel 166 $P(\emptyset) = \{\emptyset\}$

Sætning 167 *Hvis A og B er mængder og $A \subseteq B$ da er $P(A) \subseteq P(B)$.*

Bevis. Antag at $A \subseteq B$, og antag at $X \in P(A)$. Vi skal da vise, at $X \in P(B)$. Udsagnet $X \in P(A)$ betyder pr. definition, at $X \subseteq A$. Da endvidere $A \subseteq B$, ved vi fra Sætning 108, at $X \subseteq B$, hvorfor $X \in P(B)$. ■

Sætning 168 *Lad A og B være mængder. Da gælder:*

$$P(A \cap B) = P(A) \cap P(B) \tag{6.94}$$

og

$$P(A \cup B) \supseteq P(A) \cup P(B) \tag{6.95}$$

Bevis. Overlades til læseren. ■

Øvelse 169 *Vis ved et modeksempel at inklusionen $P(A \cup B) \subseteq P(A) \cup P(B)$ ikke er sand generelt. Overvej hvad A og B skal opfylde for at $P(A \cup B) \subseteq P(A) \cup P(B)$.*

Bemærkning 170 *De to udsagn $A \subseteq B$ og $A \in P(B)$ er ækvivalente. Udsagnet $A \subseteq B$ er dog begrebsmæsigt simplere og bør derfor normalt foretrækkes frem for $A \in P(B)$. Man bør generelt formulere sig på den simplest mulige måde. Af samme grund vil det ofte være lettere forståeligt, hvis man skriver "en familie af delmængder af A" i stedet for "en delmængde af $P(A)$".*

6.8 Russels paradoks

Et paradoks er i daglig tale en selvmodsigende situation. Man siger, at der er et paradoks (eller en indre modstrid) i en matematisk teori, hvis man kan udlede en sætning p og også kan udlede dens negation $\neg p$. En sådan situation kan man ikke acceptere i matematikken. For ikke alene betyder det, at vi ikke

ved, hvilken af de to alternativer p og $\neg p$ vi skal regne for sand (og logikkens regler kræver at præcist et af de to udsagn er sandt); det betyder faktisk også, at vi kan bevise ethvert udsagn q (og dets negation) i teorien. Husk nemlig at man kan bevise et udsagn q ved at vise at $\neg q \Rightarrow (p \wedge \neg p)$ (bevis ved modstrid). Men hvis både p og $\neg p$ er sætninger i teorien er $(p \wedge \neg p)$ sand, hvorfor udsagnet $\neg q \Rightarrow (p \wedge \neg p)$ er sandt. Altså er det vilkårlige udsagn q bevist ved modstrid. Altså hvis en teori indeholder ét paradoks, så er enhver sætning paradoksal i den forstand at både den og dens negation er sand, eller sagt anderledes: ethvert udsagn i teorien er både sandt og falsk. En sådan teori er både paradoksal og uinteressant.

Vi har flere gange i det foregående advaret mod at opfatte "alting" som en mængde. For eksempel insisterede vi på, at man skulle have en grundmængde for at kunne danne komplementærmængder, og vi understregede nødvendigheden af, at den frie variabel i et prædikat var begrænset til en given mængde, så vi ikke kan tale om alt i hele verden, som opfylder prædikatet. Man kan heller ikke tale om mængden af alle mængder, men kun om mængden af mængder, som er delmængder af samme givne mængde (potensmængden). Vi skal her se, hvordan der kan opstå paradokser, hvis man ikke tager disse forholdsregler. Så lad os lige for en stund glemme forholdsreglerne og formulere Russels paradoks[4].

Der er i verden nogle mængder, som indeholder sig selv som element, for eksempel mængden af mængder, og mængden af de mængder, som kan beskrives på dansk med under 20 ord. Denne sidstnævnte mængde indeholder jo sig selv, for jeg har lige beskrevet den med under 20 ord. Der er naturligvis også mængder, som ikke indeholder sig selv som element. Alle de mængder vi har set på i denne bog har således ikke haft sig selv som element. Nu kan vi så se på mængden M af alle de mængder, som ikke har sig selv som element, altså

$$M = \{X \mid X \notin X\} \tag{6.96}$$

Vi spørger så, om M har sig selv som element eller ej? For at tydeliggøre argumentet kalder vi prædikatet $X \notin X$ for $p(X)$. Altså er $M = \{X \mid p(X)\}$

Antag først at $M \in M$, altså at $M \in \{X \mid X \notin X\} = \{X \mid p(X)\}$. Ifølge definitionen på sandhedsmængden $\{X \mid p(X)\}$ betyder det, at $p(M)$ er sand altså at $M \notin M$.

Antag dernæst at $M \notin M$. Da $M = \{X \mid X \notin X\} = \{X \mid p(X)\}$, betyder det, at $p(M)$ er falsk, altså at $\neg p(M)$ er sand. Men det betyder, at $\neg(M \notin M)$, altså at $M \in M$.

Vi har altså vist at $M \in M \Rightarrow (\ \neg(M \in M))$ og $(\ \neg(M \in M)) \Rightarrow M \in M$. Men da et udsagn enten er sandt eller falsk, må enten $M \in M$.eller $(\ \neg(M \in M))$ være sandt. I begge tilfælde har vi udledt

$$M \in M \wedge (\neg(M \in M)) \tag{6.97}$$

altså et paradoks.

[4] Efter matematikeren, filosoffen og fredsforkæmperen Bertrand Russell 1872 - 1970

Der er altså indbygget et paradoks i mængdelæren, med mindre man forbyder nogle af de mængder, som indgår i Russels paradoks.

Russel's paradoks understreger vigtigheden af at formulere et sæt aksiomer for mængdelæren, altså regler for hvordan man må danne mængder og operere med dem. Det mest almindeligt accepterede aksiomssystem for mængdelæren er det såkaldte Zermelo-Fraenkel aksiomssystem eller ZFC. Vi skal ikke opstille dette aksiomssystem her, men blot understrege at vores tidligere forbud mod at tale om mængden af alt, er en følge af dette aksiomssystem. Inden for en mængdelære beskrevet ved ZFC kan Russels paradoks og lignende paradokser ikke opstå. Der kan dog ikke gives et bevis for, at der slet ikke kan opstå paradokser[5] (Gödels sætning).

6.9 Opgaver

1. Skriv følgende mængder på elementform:

   1. $\left\{x \in \mathbb{R} \mid 4x^2 - 4x - 3 = 0\right\}$
   2. $\left\{x \in \mathbb{Z} \mid 4x^2 - 4x - 3 = 0\right\}$
   3. $\{x \in \mathbb{N} \mid x \text{ går op i } 12\}$

2. Bevis, at

   1. $\{1, 2\} \subseteq \left\{x \in \mathbb{R} \mid x^2 + 2x \geq 3\right\}$
   2. $\{1, 2\} \subseteq \left\{x \in \mathbb{R} \mid x^4 - x^3 - x^2 - 5x + 6 = 0\right\}$
   3. $\left\{x \in \mathbb{R} \mid x^2 + 3x + 4 \leq 7\right\} \subseteq \left\{x \in \mathbb{R} \mid x^2 + 3x + 4 \leq 8\right\}$

3. Tegn Venn-diagrammer der illustrerer følgende situationer:

   1. $A \subseteq B \subseteq C$
   2. $A \subseteq C$, $B \subseteq C$, og $A \cap B = \emptyset$
   3. $A \cap B \subseteq C$, men hverken A eller B er delmængder af C
   4. $A \cap B \neq \emptyset$, $A \cap C \neq \emptyset$ og $B \cap C \neq \emptyset$, og mængderne $A \cap B$, $A \cap C$ og $B \cap C$ er parvist disjunkte

4. Lad A og B være delmængder af en given grundmængde U.

   1. Vis, at $B \subseteq A$ hvis og kun hvis

$$A \cup \complement B = U \tag{6.98}$$

[5] Det kan i hvert fald ikke bevises inden for mængdelæren selv eller inden for en simplere matematisk teori.

2. Vis, at A og B er disjunkte, hvis og kun hvis

$$\complement A \cup \complement B = U \tag{6.99}$$

Overbevis dig først om rigtigheden af udsagnene ved at tegne Venn-diagrammer og giv derefter formelle beviser.

5. Vis, at

$$\bigcap_{n=1}^{\infty}]-\frac{1}{n}, 1+\frac{1}{n}[= [0,1] \tag{6.100}$$

og

$$\bigcup_{n=1}^{\infty}]-\frac{1}{n}, 1+\frac{1}{n}[= [-1,2] \tag{6.101}$$

6. Bestem

$$\bigcap_{n=1}^{\infty} [-1, 1-\frac{1}{n}] \tag{6.102}$$

og

$$\bigcup_{n=1}^{\infty} [-1, 1-\frac{1}{n}] \tag{6.103}$$

7. Gælder der følgende distributive love?

$$A \cup (B \setminus C) = (A \cup B) \setminus (A \cup C), \tag{6.104}$$

$$A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C), \tag{6.105}$$

og

$$\complement(A \setminus B) = (\complement A) \setminus (\complement B)? \tag{6.106}$$

Angiv et modeksempel eller et bevis i hvert tilfælde.

8. Lad mængderne A og B være defineret ved:

$$A = \{10, 20, 30, 40\}, \tag{6.107}$$

$$B = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\} \tag{6.108}$$

Tegn et skema med dobbelt indgang, hvis rubrikker svarer til $A \times B$.

Marker mængderne

$$K = \{(x, y) \in A \times B \mid x + y \text{ er delelig med } 5\} \tag{6.109}$$

og

$$K' = \{(x, y) \in A \times B \mid x + y \text{ er delelig med } 7 \text{ og } x + y \text{ er et kvadrattal}\} \tag{6.110}$$

9. Lad A og B være mængder. Gælder det generelt at

$$P(A \setminus B) = P(A) \setminus P(B)? \tag{6.111}$$

Giv bevis eller modeksempel.

Chapter 7

Relationer

Eksempel 171 *For at give et indtryk af hvad man mener med en relation, skal vi give nogle eksempler:*

1. **Eksempel 172** *(a) "Person p har besøgt land l" er en relation mellem mængden af personer og mængden af lande.*

   (b) "p_1 har samme efternavn som p_2" er en relation på mængden af mennesker.

   (c) "Trekant T har areal A" er en relation mellem mængden af trekanter og $\mathbb{R}_+$.

   (d) "$x \leq y$" er en relation på $\mathbb{R}$.

   (e) "$x = y$" er en relation på enhver mængde A.

   (f) "n går op i m", som også skrives "$n \mid m$", er en relation på $\mathbb{N}$.

   (g) "$A \subseteq B$" er en relation på $P(U)$.

   (h) "$a \equiv b \pmod{n}$" (for givet $n \in \mathbb{N}$) er en relation på $\mathbb{Z}$. Her betyder $a \equiv b \pmod{n}$ at $n \mid (a - b)$. Man siger: a er ækvivalent med b modulo n.

Mere generelt definerer et prædikat $p(x, y)$ i de to frie variable $x \in A$ og $y \in B$ en relation mellem A og B, idet vi siger at x er relateret til y, hvis $p(x, y)$ er sand. Vi kan identificere en relation mellem A og B med en delmængde R af $A \times B$ nemlig mængden af de ordnede par (x, y), for hvilke x er relateret til y. Denne mængde er sandhedsmængden for $p(x, y)$ og kaldes også relationens graf.

Betragt for eksempel relationen $\leq$ mellem $\mathbb{R}$ og $\mathbb{R}$. Vi kan på sædvanlig vis illustrere $\mathbb{R}^2$ ved den cartesiske plan. Grafen for relationen $\leq$ er da den skraverede mængde på figuren 7.1

Den ovenstående beskrivelse indfanger den intuitive betydning af en relation. For at føre begrebet tilbage til de fundamentale begreber i mængdelæren vælger man dog at *definere* en relation ud fra grafen, altså som en delmængde af en produktmængde:

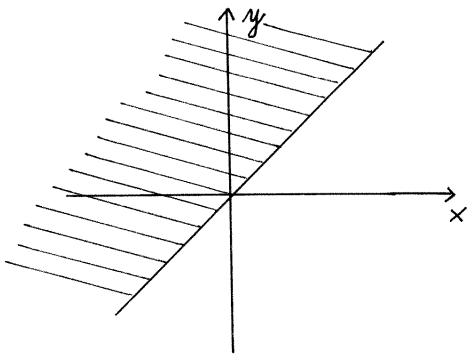

Figure 7.1: Grafen for relationen $\leq$

Definition 173 *Lad A og B være mængder. En delmængde R af $A \times B$ kaldes en **relation** mellem A og B. A kaldes **primærmængden** og B kaldes **sekundærmængden**. En delmængde af $A \times A$ kaldes en relation på A.*

Hvis $(x, y) \in R$ siger man at x er relateret til y, og man skriver xRy.

Bemærkning 174 *Selv om vi således formelt definerer en relation som en mængde, bruger vi ikke mængdelærens sprog, når vi taler om relationer. For eksempel vil man ikke skrive $(< \cup =) = \leq$. Man vil i stedet skrive: $((x < y) \vee (x = y)) \Leftrightarrow x \leq y$.*

Notation 175 *Ofte vælger man andre betegnelser for relationer end xRy. I eksempel 4 ovenfor skriver man naturligvis $x \leq y$. Notationen $x \sim y$ bruges også ofte for en generel relation, men i denne bog skal vi fortrinsvis bruge denne notation om de såkaldte ækvivalensrelationer (se nedenfor).*

Definition 176 *En relation R på en mængde A siges at være*

- ***refleksiv**, hvis $\forall x \in A : xRx$*
- ***symmetrisk**, hvis $xRy \Rightarrow yRx$*
- ***antisymmetrisk**, hvis $((xRy) \wedge (yRx)) \Rightarrow x = y$*
- ***transitiv**, hvis $(xRy) \wedge (yRz) \Rightarrow (xRz)$*

Eksempel 177 *Lad os undersøge om eksemplerne i 171. er refleksive, symmetriske, antisymmetriske og/eller transitive. Disse egenskaber er kun defineret for relationer, hvor primærmængden er lig sekundærmængden, hvilket udelukker eksempel a og c.*

Relationerne i eksempel b og h er refleksive, symmetriske og transitive.

Relationen i eksempel d,f og g er refleksive, antisymmetriske og transitive.

Relationen i eksempel e er refleksiv, symmetrisk, antisymmetrisk og transitiv.

Øvelse 178 *Bevis at det forholder sig som påstået i eksempel 177. Det eneste vanskelige tilfælde er relationen h.*

Eksempel 179 *I Euklids Elementer betyder = ikke identitet, men derimod ligestorhed. For eksempel siger Euklid at linjestykket AB er lig med linjestykket CD, hvis de er lige store. Euklids første almindelige begreb siger at lighedsrelationen er transitiv.*

Vi skal i det følgende betragte to særligt vigtige slags relationer, nemlig ækvivalensrelationer og ordningsrelationer

7.1 Ækvivalensrelationer

I matematik er man ofte i den situation, at forskellige objekter kan anses for ens i en vis forstand. For eksempel anser Euklid to linjestykker for ens, hvis de er lige lange. Ligeså, når man angiver vinkelmål i grader, vil man anse en vinkel på 270° for "den samme" vinkel som en på $-90°$. Mere generelt vil to vinkler på $a°$og $b°$ anses for "den samme" vinkel hvis $a-b$ er delelig med 360, eller sagt anderledes, at $a \equiv b \pmod{360}$. Man kan sige, at de to vinkler er ækvivalente. Denne ide om ækvivalens indfanges af begrebet en ækvivalensrelation.

Der er tradition for at man betegner ækvivalensrelationer med $\sim$ i stedet for R. Derfor skal vi i dette afsnit skrive $x \sim y$ i stedet for xRy.

Definition 180 *En relation $\sim$ på en mængde A kaldes en* ***ækvivalensrelation****, hvis den er refleksiv, symmetrisk og transitiv, altså hvis*

- *Refleksivitet:* $\forall x \in A : x \sim x$
- *Symmetri:* $x \sim y \Rightarrow y \sim x$
- *Transitivitet:* $(x \sim y) \wedge (y \sim z) \Rightarrow x \sim z$

Eksempel 181 *I eksempel 171 er b,e og h de eneste ækvivalensrelationer.*

Øvelse 182 *Hvilke af nedenstående relationer er ækvivalensrelationer:*

1. *Relationen på $\mathbb{R}$ defineret ved: $x \cdot y \geq 0$*
2. *Relationen $\neq$ defineret på $\mathbb{R}$.*
3. *Relationen på mængden af rette linjer i planen defineret ved $l \sim m$, hvis (l er parallel med m eller $l = m$).*
4. *Relationen "l står vinkelret på m" defineret på mængden af rette linjer i planen.*
5. *Relationen på mængden af rette linjer i planen defineret ved $l \sim m$, hvis l og m har et fælles punkt.*

6. *Relationen på mængden af orienterede linjestykker $\overline{AB}$ (pile fra A til B) i planen defineret ved $\overline{AB} \sim \overline{CD}$ hvis $\overline{AB}$ og $\overline{CD}$ er ensrettede og lige lange.*

En ækvivalensrelation på en mængde deler mængden op i delmængder bestående af elementer, som er indbyrdes ækvivalente. For eksempel deler ækvivalensrelationen b i eksempel 171 mennesker op i delmængder bestående af personer med samme efternavn. Ækvivalensrelationen "samme køn som" på mængden af mennesker deler menneskeheden op i kvinder og mænd. Ækvivalensrelationen "går på hold med" deler mængden af MatM studerende op i hold. Lad os gøre denne ide mere præcis:

Definition 183 *Lad $\sim$ være en ækvivalensrelation på mængden A. Hvis $a \in A$, betegner $[a]$ mængden af de elementer i A, som er ækvivalente med a, med andre ord*

$$[a] = \{x \in A \mid x \sim a\}. \tag{7.1}$$

*$[a]$ kaldes den til a hørende **ækvivalensklasse**. Hvis man eksplicit vil angive at ækvivalensklassen $[a]$ hører til relationen $\sim$ kan man skrive $[a]_\sim$*

Mængden af ækvivalensklasser betegnes $A/\sim$

Eksempel 184 *Betragt ækvivalensrelationen "samme køn" på mængden af mennesker. Da er [Mette Hansen] lig mængden af kvinder.*

Eksempel 185 *Den i Øvelse 182 eksempel 6 definerede relation er en ækvivalensrelation. Dens ækvivalensklasser kaldes vektorer i planen. Ækvivalensklassen $\left[\overline{AB}\right]$ betegnes også $\overrightarrow{AB}$.*

Eksempel 186 *Betragt ækvivalensrelationen "$a \equiv b \pmod n$" på $\mathbb{Z}$ beskrevet i 171.h. Lad os først se på tilfældet $n = 3$. For at bestemme ækvivalensklassen $[0]$ hørende til 0 skal vi altså bestemme alle hele tal a, så $a \sim 0$, dvs. så $3 \mid (a - 0)$ eller $3 \mid a$. Ækvivalensklassen $[0]$ består altså af alle de hele tal, som er delelige med 3, dvs*

$$[0] = \{..., -6, -3, 0, 3, 6, ...\} \tag{7.2}$$

Elementerne i $[1]$ er de hele tal a, som er ækvivalente med 1, altså de $a \in \mathbb{Z}$, hvor $3 \mid (a - 1)$. Men det betyder, at $\exists k \in \mathbb{Z} : 3k = (a - 1)$, eller at $a = 3k + 1$, hvad der betyder at a har rest 1 ved division med 3. Altså har vi

$$[1] = \{..., -5, -2, 1, 4, 7, ...\} \tag{7.3}$$

På samme måde indses at $[2]$ er mængden af de hele tal, som har rest 2 ved division med 3, altså

$$[2] = \{..., -7, -4, -1, 2.4, ...\} \tag{7.4}$$

*Disse ækvivalensklasser kaldes **restklasser** modulo 3. Hvis man eksplicit vil angive, at der er tale om restklasser modulo 3, kan man skrive restklasserne på formen $[a]_3$*

Sætning 187 *Lad $\sim$ være en ækvivalensrelation på mængden A, og lad $a, b \in A$. Da gælder at*

$$[a] = [b]\,, \tag{7.5}$$

hvis og kun hvis

$$a \sim b. \tag{7.6}$$

Bevis. Først vises, at $a \sim b \Rightarrow [a] = [b]$. Antag altså, at $a \sim b$. Vi skal vise, at $[a] = [b]$. Først viser vi, at $[a] \subseteq [b]$. Lad derfor $c \in [a]$. Det betyder pr. definition 183 at $c \sim a$. Og da vi havde antaget, at $a \sim b$ følger af transitiviteten, at $c \sim b$, hvoraf vi slutter, at $c \in [b]$. Altså gælder $[a] \subseteq [b]$. Dernæst skal det vises, at $[b] \subseteq [a]$. Det kan gøres helt som ovenfor. Man kan også bemærke at symmetrien medfører at $b \sim a$ hvorved inklusionen $[b] \subseteq [a]$ følger, af det netop gennemførte argument.

Dernæst vises at $[a] = [b] \Rightarrow a \sim b$. Antag altså, at $[a] = [b]$, og vi vil vise at $a \sim b$. Da relationen er refleksiv, ved vi, at $a \sim a$. Det betyder, at $a \in [a]$. Da $[a] = [b]$, har vi også, at $a \in [b]$. Men det betyder pr. definition, at $a \sim b$. ■

Korollar 188 *Lad $\sim$ være en ækvivalensrelation på mængden A, og lad $a, b \in A$. Da gælder at*

$$a \in [b] \Leftrightarrow [a] = [b] \tag{7.7}$$

Bevis. Under de angivne forudsætninger gælder følgende biimplikationer:

$$a \in [b] \Leftrightarrow a \sim b \Leftrightarrow [a] = [b] \tag{7.8}$$

Den første biimplikation skyldes definitionen af $[b]$, og den sidste er sætning 187. ■

Det følger af korollaret, at hvis a ligger i en ækvivalensklasse, så kan denne ækvivalensklasse betegnes med $[a]$. Vi kalder derfor ethvert element i en ækvivalensklasse for en **repræsentant** for ækvivalensklassen.

Bemærkning 189 *Ethvert tal har rest $0, 1$ eller 2 ved division med 3. Det betyder at et vilkårligt $a \in \mathbb{Z}$ er element i netop en af restklasserne $[0]$, $[1]$ og $[2]$ modulo 3, og da er $[a]$ ifølge korollaret netop lig denne restklasse. Der er altså netop de tre restklasser $[0]$, $[1]$ og $[2]$ modulo 3, og $[-3] = [0] = [3] = [6] = ...$, $[-2] = [1] = [4] = ...$, $[-1] = [2] = [5] = ...$ Mængden af de tre ækvivalensklasser kaldes $\mathbb{Z}/3$ (udtales Z modulo 3) (Man bruger også betegnelsen $\mathbb{Z}_3$ og $\mathbb{Z}/\mathbb{Z}_3$ for denne mængde).*

Bemærkning 190 *For ethvert naturligt tal n defineres $\mathbb{Z}/n$ på samme måde som mængden af de n restklasser modulo n*

Eksempel 191 *Mængden af restklasser $\mathbb{Z}/24$ er god at bruge, når man angiver klokkeslæt. To timer efter klokken 23 plejer vi ikke at sige, at klokken er 25, men at den er 1. Men $1 \equiv 25$ (modulo 24) så de to klokkeslæt er ækvivalente. Sagt anderledes: $[25] = [1]$ i $\mathbb{Z}/24$. Det er altså praktisk at opfatte (hele) klokkeslæt som restklasser i $\mathbb{Z}/24$ snarere end som hele tal.*

Eksempel 192 *På samme måde er det praktisk at opfatte vinkelmål angivet i grader, som restklasser modulo 360.*

For at få en god forståelse af ækvivalensklasserne for en ækvivalensrelation indfører vi begrebet en klasseinddeling af en mængde

Definition 193 *En familie Ω af delmængder af en mængde M kaldes en* ***klassedeling*** *(eller en partition) af M, hvis mængderne i Ω er parvist disjunkte, og deres foreningsmængde er hele M, med andre ord hvis*

1. *For alle mængder A og B i Ω gælder enten, at $A = B$ eller at $A \cap B = \emptyset$.*
2. $\bigcup_{A \in \Omega} A = M$

Eksempel 194

1. *Opdelingen af eleverne i en skole i klasser er en klassedeling*
2. *Inddelingen af dyr i arter er en klassedeling.*
3. *Inddelingen af mennesker i mænd og kvinder er en klassedeling.*

Vi vil nu vise to sætninger, som til sammen siger at ækvivalensrelationer og klassedelinger er to sider af samme sag. Sagt med andre ord: En ækvivalensrelation giver anledning til en klassedeling, og en klassedeling giver anledning til en ækvivalensrelation. Disse to sætninger er hovedsætningerne i dette afsnit.

Sætning 195 *Lad $\sim$ være en ækvivalensrelation på en mængde M. Da er familien af ækvivalensklasser $M/\sim$ en klassedeling af M.*

Bevis. Lad $\sim$ være en ækvivalensrelation på mængden M. Vi skal da bevise at familien af ækvivalensklasser $M/\sim$ opfylder de to definerende egenskaber i 193. Da mængden af ækvivalensklasser består af mængderne $[a]$ for $a \in M$, skal vi altså vise

1. For $a, b \in M$, gælder enten $[a] = [b]$ eller $[a] \cap [b] = \emptyset$
2. $\bigcup_{a \in M} [a] = M$

Ad.1. Lad $a, b \in M$. Vi vil vise egenskab 1, ved at vise, at hvis $[a] \cap [b] \neq \dot{\emptyset}$, så er $[a] = [b]$. Antag altså, at $[a] \cap [b] \neq \emptyset$. Så findes der et $c \in [a] \cap [b]$. Da $c \in [a]$, følger af definition 183 at $c \sim a$, og da $c \in [b]$ gælder, at $c \sim b$. Da nu $\sim$ er en ækvivalensrelation, er den specielt symmetrisk, så vi kan slutte at $a \sim c$. Og da $\sim$ også er transitiv kan vi fra $a \sim c$ og $c \sim b$ slutte, at $a \sim b$. Sætning 187 fortæller da at $[a] = [b]$

Ad. 2. Da enhver ækvivalensklasse pr. definition er en delmængde af M, vil $\bigcup_{a \in M} [a] \subseteq M$. Vi skal altså vise at $\bigcup_{a \in M} [a] \supseteq M$. Lad derfor $a \in M$. Da $\sim$ er refleksiv er $a \sim a$, så $a \in [a]$ (iflg. sætning 187) hvorfor $a \in \bigcup_{a \in M} [a]$ ■

Bemærkning 196 *Bemærk at vi i ovenstående bevis brugte alle de tre definerende egenskaber ved en ækvivalensrelation.*

Eksempel 197 *Mængden af restklasser modulo n er altså en klassedeling af de naturlige tal.*

Definition 198 *Lad Ω være en klassedeling af en mængde M. Vi definerer da en relation $\sim_\Omega$ på M ved:*

$a \sim_\Omega b$, hvis der findes en mængde $A \in \Omega$, så $a, b \in A$.

Sætning 199 *Lad Ω være en klassedeling af en mængde M. Da er relationen $\sim_\Omega$ defineret ovenfor en ækvivalensrelation på M.*

Bevis. Antag at Ω er en klassedeling af M. Vi skal da vise, at $\sim_\Omega$ er refleksiv, symmetrisk og transitiv.

Refleksivitet: Lad $a \in M$. Da Ω er en klassedeling af M, gælder specielt at $\bigcup_{A \in \Omega} A = M$. Altså findes der et $A \in \Omega$, så $a \in A$ (altså så $a, a \in A$), hvorfor $a \sim_\Omega a$ ifølge definitionen af $\sim_\Omega$.

Symmetri: Lad $a, b \in M$, og antag, at $a \sim_\Omega b$. Da findes pr. definition en mængde $A \in \Omega$, så $a, b \in A$; men så gælder jo ligeså, at $b, a \in A$, hvorfor $b \sim a$.

Transitivitet: Lad $a, b, c \in M$, og antag at $a \sim_\Omega b$ og at $b \sim_\Omega c$. Vi skal bevise, at $a \sim_\Omega c$. Da $a \sim_\Omega b$, findes en mængde $A \in \Omega$, så $a, b \in A$, og da $b \sim_\Omega c$ findes en mængde $B \in \Omega$, så $b, c \in B$. Da $b \in A$ og $b \in B$ er $A \cap B \neq \emptyset$, men da Ω er en klassedeling, ved vi fra den første definerende egenskab, at når $A \cap B \neq \emptyset$, er $A = B$. Men så gælder, at $a \in A$ og $c \in A$, og da $A \in \Omega$ betyder det at $a \sim_\Omega c$. ■

Bemærkning 200 *Bemærk at vi i ovenstående bevis brugte begge de to definerende egenskaber ved en klassedeling.*

Eksempel 201 *Opdelingen af eleverne i en skole i klasser, giver anledning til en ækvivalensrelation mellem eleverne. To elever er relateret, hvis de går i samme klasse.*

Sætning 202

1. *Givet en ækvivalensrelation $\sim$ på en mængde M. Den giver anledning til en klassedeling $M/\sim$ af M. Denne klassedeling giver så igen anledning til en ækvivalensrelation $\sim_{M/\sim}$. Denne ækvivalensrelation er den samme som $\sim$.*

2. *Givet en klassedeling Ω af en mængde M. Den giver anledning til en ækvivalensrelation $\sim_\Omega$ på M. Denne ækvivalensrelation giver så igen anledning til en klassedeæing $M/\sim_\Omega$. Denne klassedeling er den samme som Ω.*

Øvelse 203 *Bevis sætning 202*

Når man regner med vinkler, kan man tillade sig at sige at $270° + 270° = 180°$, fordi $270 + 270 = 540$ og $540 \equiv 180 \pmod{360}$. Vi regner altså som om to vinkler er "ens" når de er ækvivalente modulo 360. Et matematisk mere tilfredsstillende synspunkt er at regne med restklasserne. Vi skal nu vise at det er muligt at indføre addition og multiplikation på mængden af restklasser modulo n.

Definition 204 *Lad $n \in \mathbb{N}$. Vi definerer da addition $+$, og multiplikation $\cdot$ af to restklasser $[a], [b] \in \mathbb{Z}/n$ ved*

$$[a] + [b] = [a + b] \tag{7.9}$$

$$[a] \cdot [b] = [a \cdot b] \tag{7.10}$$

Bemærkning 205 *Vi definerer altså addition og multiplikation af to restklasser ud fra addition og multiplikation af to repræsentanter for restklasserne. For at disse definitioner skal give mening må vi godtgøre, at resultatet ikke afhænger af valget af repræsentant. Vi skal altså vise følgende sætning:*

Sætning 206 *Lad $n \in \mathbb{N}$. Hvis $[a_1], [a_2], [b_1], [b_2] \in \mathbb{Z}/n$ og $[a_1] = [a_2]$ og $[b_1] = [b_2]$ da er $[a_1 + b_1] = [a_2 + b_2]$ og $[a_1 \cdot b_1] = [a_2 \cdot b_2]$.*

Bevis. Vi viser blot at addition er repræsentantuafhængig; multiplikation går på samme vis.

Antag altså at

$$[a_1] = [a_2] \text{ og } [b_1] = [b_2], \tag{7.11}$$

dvs. iflg. 187 at

$$a_1 \equiv a_2 \pmod{n} \text{ og } b_1 \equiv b_2 \pmod{n}. \tag{7.12}$$

Pr. definition 171h. betyder det, at

$$n \mid (a_1 - a_2) \text{ og } n \mid (b_1 - b_2), \tag{7.13}$$

hvorfor der findes hele tal m, l så

$$a_1 - a_2 = nm \text{ og } b_1 - b_2 = nl \tag{7.14}$$

eller

$$a_1 = a_2 + nm \text{ og } b_1 = b_2 + nl. \tag{7.15}$$

Men så ser vi, at

$$a_1 + b_1 = a_2 + nm + b_2 + nl = a_2 + b_2 + n(m + l). \tag{7.16}$$

Da nu $m + l$ er et helt tal, slutter vi heraf, at

$$a_1 + b_1 \equiv a_2 + b_2 \pmod{n}, \tag{7.17}$$

eller

$$[a_1 + b_1] = [a_2 + b_2]. \tag{7.18}$$

■

Eksempel 207 *Betragt restklasserne modulo 360. Der gælder* $[195] + [341] = [176]$ *(vis dette). Det betyder, at hvis et hjul først drejer* $195°$ *og dernæst* $341°$ *i samme retning, vil det indtage samme stilling, som hvis det bare havde drejet* $176°$*.*

Øvelse 208 *Hvad er klokken* 127 *timer efter kl.* 15*? Udtryk dine overvejelser som et regnestykke med restklasser.*

7.2 Ordningsrelationer

Det er normalt at bruge betegnelsen $\leq$ (læses "mindre eller lig (med)) for en ordningsrelation, også når der er tale om andre ordningsrelationer end den velkendte ordning på de reelle tal. Vi skal følge denne konvention i dette afsnit.

Definition 209 *En relation* $\leq$ *på en mængde* M *kaldes en (partiel)* ***ordningsrelation*** *(eller en (partiel) ordning), hvis den er refleksiv, antisymmetrisk og transitiv, d.v.s hvis den opfylder følgende tre betingelser:*

1. *Refleksivitet: For alle* $a \in M$ *gælder det at* $a \leq a$
2. *Antisymmetri: For alle* $a, b \in M$ *gælder det at* $(a \leq b) \wedge (b \leq a) \Rightarrow a = b$*.*
3. *Transitivitet: For alle* $a, b, c \in M$ *gælder at* $(a \leq b) \wedge (b \leq c) \Rightarrow a \leq c$

En mængde forsynet med en ordning kaldes en ordnet mængde. Den ordnede mængde M *forsynet med ordningen* $\leq$ *betegnes* $(M, \leq)$*.*

Øvelse 210 *Bevis følgende simple sætning: Lad* $(M, \leq)$ *være en partielt ordnet mængde og* $A \subseteq M$*. Da definerer* $\leq$ *også en partiel ordning på* A*.*

Eksempel 211
1. *Den almindelig kendte ordning* $\leq$ *af de reelle tal er en ordningsrelation.*
2. *Relationen* $\subseteq$ *er en ordningsrelation på potensmængden* $P(M)$ *af en given mængde* M*.*
3. *"Være efterkommer af" er en ordningsrelation på mængden af mennesker, hvis man siger at en person er efterkommer efter sig selv.*
4. *Relationen* $=$ *på en mængde* M *er en ordningsrelation. Den derved ordnede mængde* $(M, =)$ *kaldes den totalt uordnede mængde.*

Bemærkning 212 *I det første af disse eksempler kan to vilkårlige elementer* a *og* b *sammenlignes, i den forstand at enten er* $a \leq b$ *eller* $b \leq a$*. Noget tilsvarende gælder ikke i almindelighed i de sidste tre eksempler. Vi siger, at den første ordningsrelation er en total ordning.*

Definition 213 *En ordningsrelation* $\leq$ *på en mængde* M *kaldes en* ***total ordningsrelation*** *(eller en* ***total ordning****), hvis der for alle* $a, b \in M$ *gælder enten* $a \leq b$ *eller* $b \leq a$*. I så fald kaldes* $(M, \leq)$ *en totalt ordnet mængde.*

Bemærkning 214 *Man bruger også betegnelsen lineær ordningsrelation om en total ordningsrelation. Det antyder ideen om, at man kan tænke sig en totalt ordnet mængde $(M, \leq)$ anbragt langs en ret linje, således at $a \leq b$, hvis a ligger til venstre for b. Hvis en partielt ordnet mængde ikke er totalt ordnet, kan den ikke tænkes anbragt således.*

Eksempel 215 *De reelle tal med den sædvanlige ordning er en totalt ordnet mængde.*

Bemærkning 216 *Hvis A er en delmængde af en partielt ordnet mængde $(M, \leq)$, vil relationen $\leq$ anvendt på A også være en ordningsrelation. Derved bliver $(A, \leq)$ en ordnet mængde. Man siger at A arver ordensstrukturen fra $(M, \leq)$. Hvis $(M, \leq)$ er en totalt ordnet mængde, bliver enhver delmængde, udstyret med den nedarvede ordning, ligeledes totalt ordnet. For eksempel bliver alle delmængder af $\mathbb{R}$ totalt ordnede mængder, når de udstyres med den sædvanlige ordning. Talmængderne $\mathbb{N}$, $\mathbb{Z}$ og $\mathbb{Q}$ er altså totalt ordnede mængder, når de udstyres med den sædvanlige ordning.*

Øvelse 217
1. *Overvej om $(P(M), \subseteq)$ kan være totalt ordnet for passende valg af M.*
2. *Angiv en uendelig delmængde af $P(\mathbb{N})$, som er totalt ordnet ved relationen $\subseteq$.*

Definition 218 *Lad $(M, \leq)$ være en partielt ordnet mængde, og lad $a, b \in M$. Vi siger da at*

- *$a < b$ hvis $a \leq b$ og $a \neq b$*
- *$a \geq b$ hvis $b \leq a$*
- *$a > b$ hvis $b < a$*

Definition 219 *En partielt ordnet mængde $(M, \leq)$ siges at opfylde **trikotomiloven** (eller tredelingsloven) hvis der for vilkårlige elementer $a, b \in M$ gælder, at præcist et af udsagnene*

$$a < b,\ a = b,\ b < a \tag{7.19}$$

er sandt.

Sætning 220 *En partielt ordnet mængde $(M, \leq)$ er totalt ordnet, hvis og kun hvis den opfylder trikotomiloven.*

Bevis. Antag først, at $(A, \leq)$ er totalt ordnet. Vi skal vise, at tredelingsloven gælder. Lad derfor $a, b \in A$ være vilkårlige. Vi skal vise, at mindst et af udsagnene $a < b$, $a = b$ og $b < a$ er sandt. Dette kan vises ved at vise, at såfremt de to første er falske, da må det tredie være sandt (overveje).

Vi antager derfor $\sim (a < b)$ og $\sim (a = b)$ og skal da vise $b < a$.

Da $a < b$ per definition betyder $(a \leq b \wedge a \neq b)$, har vi:

$$\sim (a < b) \equiv \sim (a \leq b) \vee a = b.$$

Da vi har antaget $\sim (a < b)$ og $a \neq b$, kan vi således slutte $\sim (a \leq b)$. Da A er totalt ordnet, slutter vi heraf, at $b \leq a$. Da også $b \neq a$, har vi da per definition $b < a$.

Antag nu omvendt, at $(A, \leq)$ tilfredstiller tredelingsloven. Vi skal da vise, at A er totalt ordnet. Lad da $a, b \in A$ være vilkårlige. Vi skal vise, at mindst et af udsagnene $a \leq b$ og $b \leq a$ er sandt. Dette kan vises ved at antage, at $a \leq b$ er falsk og på grundlag heraf slutte, at $b \leq a$ gælder.

Vi antager derfor $\sim (a \leq b)$. Nu, da $a < b$ betyder $(a \leq b \wedge a \neq b)$, er implikationen $a < b \Rightarrow a \leq b$ altid sand; da vi har antaget, at $a \leq b$ er falsk, kan derfor $a < b$ ikke være sand. Tilsvarende kan $a = b$ ikke være sand, da i så fald $a \leq b$ også ville være det.

Vi har nu indset, at udsagnene $a < b$ og $a = b$ begge er falske. Da $(A, \leq)$ tilfredsstiller tredelingsloven, slutter vi, at $b < a$. Men da er også $b \leq a$ sand som ønsket ■

I resten af dette afsnit vil vi nøjes med at betragte totalt ordnede mængder. En del af de begreber og sætninger vi indfører, kan dog også generaliseres til partielt ordnede mængder.

Definition 221 *Et element a i en totalt ordnet mængde $(M, \leq)$ kaldes et* ***største element*** *(eller et* ***maximalt element*** *eller et* ***maximum****), hvis $x \leq a$ for alle $x \in M$. Maximum af mængden M betegnes med* $\max M$.

Eksempel 222 *Det højre endepunkt b er det største element i (maksimum af) det lukkede interval $[a, b]$ med den sædvanlige ordning.*

Eksempel 223 *Det åbne interval $]a, b[$ har intet største element.*

Bevis. Beviset er et modstridsbevis. Antag at c er det største element i $]a, b[$. Ifølge definitionen er c da et element i $]a, b[$, så $a < c < b$. Betragt tallet $d = \frac{b+c}{2}$ (midt mellem b og c). Da $c < b$ er

$$d = \frac{b+c}{2} < \frac{b+b}{2} = b, \tag{7.20}$$

og da $a < b$ og $a < c$, er

$$a = \frac{a+a}{2} < \frac{b+c}{2} = d. \tag{7.21}$$

Altså ligger d i intervallet $]a, b[$, og da c var antaget at være det største element i intervallet, må det gælde at $d \leq c$. Men da $c < b$ fås også

$$c = \frac{c+c}{2} < \frac{c+b}{2} = d. \tag{7.22}$$

Specielt gælder altså $c \leq d$, og da også $d \leq c$, slutter vi af antisymmetrien, at $c = d$. Men da $c < d$ ved vi specielt at $c \neq d$. Dermed har vi opnået en modstrid, og vi kan konkludere at $]a, b[$ ikke har et største element. ■

Øvelse 224 *Bevis at intervallet* $[a, \infty[$ *ikke har noget største element.*

Sætning 225 *Hvis en totalt ordnet mængde har et største element, så er det entydigt bestemt.*

Bevis. For at vise denne entydighedssætning benytter vi den strategi vi lagde i Bemærkning 67.

Antag at a og b begge er største elementer i en ordnet mængde $(M, \leq)$. Da $a \in M$, og b er et største element i M, gælder at

$$a \leq b. \tag{7.23}$$

Da $b \in M$, og a er et største element i M, gælder at

$$b \leq a. \tag{7.24}$$

Men da relationen $\leq$ er antisymmetrisk, betyder det at $a = b$. ■

Hvis en ordnet mængde har et største element, kan vi altså tale om *det* største element i mængden.

Definition 226 *Et element* a *i en totalt ordnet mængde* $(M, \leq)$ *kaldes et* ***mindste element*** *(eller et* ***minimalt element*** *eller et* ***minimum****), hvis* $a \leq x$ *for alle* $x \in M$*. Minimum for en mængde* M *betegnes med* $\min M$*.*

Øvelse 227 *Bevis at det venstre endepunkt* a *er minimum af det lukkede interval* $[a, b]$*, og bevis at de åbne intervaller* $]a, b[$ *og* $]-\infty, b[$ *ikke har noget minimum.*

Definition 228 *Lad* A *være en delmængde af en totalt ordnet mængde* $(M, \leq)$*, og lad* $x \in M$*.*

1. x *kaldes en* ***majorant*** *for* A *hvis* $a \leq x$ *for alle* $a \in A$*.*
2. x *kaldes en* ***minorant*** *for* A *hvis* $x \leq a$ *for alle* $a \in A$*.*

Hvis der findes en majorant for A *siges* A *at være* ***opadtil begrænset****. Hvis der findes en minorant for* A *siges* A *at være* ***nedadtil begrænset****. Hvis* A *er både opadtil og nedadtil begrænset, kaldes* A *for* ***begrænset***

Eksempel 229 *Tallet* 1000 *er en majorant for* $[0, 1]$*. Intervallet* $[0, 1]$ *er altså opadtil begrænset.*

Øvelse 230 *Giv eksempler på intervaller, som er/ikke er opadtil og nedadtil begrænsede. I de tilfælde de er begrænsede, giv da flere majoranter/minoranter.*

Bemærkning 231 *Betegnelserne majorant og minorant hedder "upper bound" og "lower bound" på engelsk. De er i flere bøger fejlagtigt blevet oversat til "øvre og nedre grænse" på dansk. Disse betegnelser kan imidlertid virke vildledende. Normalt tænker man sig jo en grænse som liggende lige der, hvor noget begynder*

eller ender. Den danske grænse ligger ved Kruså ikke ved Hamborg. Men i matematik kan en majorant altså ligge "langt fra mængden", bare den er større eller lig med alle elementerne i mængden. Det er derfor vi i disse noter har valgt betegnelserne majorant og minorant. Når der er tale om talmængder bruger man også betegnelserne ***overtal*** *og* ***undertal*** *for majorant og minorant.*

Bemærkning 232 *Vi har bemærket, at en majorant eller minorant for en delmængde af en ordnet mængde ikke behøver at ligge i delmængden, men det kan ske.*

Sætning 233 *Lad A være en delmængde A af en totalt ordnet mængde $(M, \leq)$.*

1. *Hvis A har et største element x, så er x også en majorant for A.*

2. *Hvis A har en majorant x, som ligger i A, da vil x være det største element i A.*

Bevis. Overlades til læseren ■

Bemærkning 234 *Vi har set at der (for eksempel i de reelle tal) er mængder, som ikke har et maximum eller et minimum. Hvis mængden er ubegrænset, er der ikke noget at gøre ved det. Men hvis mængden er begrænset, kan man somme tider (for eksempel i de reelle tal), finde en erstatning for maximum og minimum, det såkaldte supremum og infimum. Disse begreber vil vi nu indføre.*

Definition 235 *Lad A være en delmængde af en totalt ordnet mængde $(M, \leq)$. Et element b i M kaldes* ***supremum*** *for A hvis det er den mindste majorant for A, d.v.s opfylder følgende to kriterier:*

1. *b er en majorant for A, altså: $\forall a \in A : a \leq b$.*

2. *b er den mindste majorant for A, altså: Hvis x er en majorant for A, da er $b \leq x$.*

Hvis b er supremum for A skriver man: $b = \sup A$

Det kan være praktisk at omformulere det andet krav i definitionen af supremum.

Sætning 236 *Lad A være en delmængde af en totalt ordnet mængde $(M, \leq)$. Da er $b = \sup A$, hvis og kun hvis*

1. *b er en majorant for A, altså: $\forall a \in A : a \leq b$.*

2. *$\forall x < b \exists a \in A : x < a$.*

Bevis. Da formuleringen af punkt 1 er det samme som i definitionen af supremum, skal vi bare vise at punkt 2 i definitionen og i sætningen er ækvivalente under forudsætning af at b er en majorant for A:

$$\begin{aligned} & b \text{ er den mindste majorant for } A && (7.25)\\ \Leftrightarrow\ & \text{ethvert } x \in M \text{ som er mindre end } b \text{ er } \textit{ikke} \text{ en majorant for } A && (7.26)\\ \Leftrightarrow\ & \forall x < b \exists a \in A : x < a. && (7.27) \end{aligned}$$

For at indse den sidste omskrivning, bemærker vi at udsagnet: "x er en majorant for A" jo betyder: $\forall a \in A : a \leq x$, så ved negering ses at udsagnet "x er ikke majorant for A" betyder: $\exists a \in A : x < a$. ■

Eksempel 237 *Intervallet* $]1, 2[$ *har supremum* 2.

Bevis. Vi vil bevise at 2 opfylder de to krav i sætning 236, altså 1. at 2 er en majorant for $]1, 2[$, og 2. at hvis $x < 2$, da findes et $a \in]1, 2[$ så $x < a$

1. Ifølge definitionen af $]1, 2[$, gælder at et vilkårligt $a \in]1, 2[$ opfylder $a < 2$ og desto mere $a \leq 2$. Altså er 2 en majorant for $]1, 2[$.

2. Antag at $x < 2$. Betragt tallet $y = \frac{x+2}{2}$ midt mellem x og 2. Da $x < 2$ gælder det at
$$y = \frac{x+2}{2} < \frac{2+2}{2} = 2. \tag{7.28}$$
Hvis $1 < y$, vælges $a = y$. Hvis $y \leq 1$, vælges $a = 3/2$. I begge tilfælde gælder at $a \in]1, 2[$ (overvej dette), og
$$x = \frac{x+x}{2} < \frac{x+2}{2} = y \leq a. \tag{7.29}$$
Vi har altså fundet et element a i $]1, 2[$, som opfylder $x < a$

■

Øvelse 238 *Bevis, at hvis* $k \in \mathbb{R}$*, vil intervallet* $]-\infty, k[$ *have supremum* k*.*

Sætning 239 *Lad* A *være en delmængde af en totalt ordnet mængde* $(M, \leq)$*. Hvis* A *har et supremum, er det entydigt bestemt.*

Bevis. Antag at x og y begge er supremum for A. Da x er den mindste majorant, og y er en majorant, gælder $x \leq y$. Da y er en mindste majorant, og x er en majorant, gælder $y \leq x$. Da relationen $\leq$ er antisymmetrisk, kan vi fra $x \leq y$ og $y \leq x$ slutte, at $x = y$. ■

Sætning 240 *Lad* A *være en delmængde af en totalt ordnet mængde* $(M, \leq)$*. Hvis* A *har et største element* a*, da vil* a *også være supremum for* A*.*

Bevis. Antag at a er det største element i A. Ifølge sætning 233 er a en majorant for A. Vi skal altså bare vise, at det er den mindste. Antag altså at x er en majorant for A. Da $a \in A$ gælder da, at $a \leq x$. Altså er a den mindste majorant for A. ■

Bemærkning 241 *Et største element (maksimum) er altså også et supremum. Det omvendte gælder ikke altid. For eksempel har intervallet* $]1, 2[$ *supremum* 2*, men det har intet største element.*

Bemærkning 242 *Det er klart, at en opadtil ubegrænset mængde ikke har noget supremum. Den har jo ikke en gang nogen majorant. Men hvad med opadtil begrænsede mængder, Kan vi bevise at de har et supremum? Nej, det kan vi heller ikke generelt (se Eksempel 245). Men der er en klasse vigtige totalt ordnede mængder, bl.a. de reelle tal, hvor enhver ikke tom opadtil begrænset mængde har et supremum. Disse giver vi et særligt navn:*

Definition 243 *En totalt ordnet mængde siges at have* ***supremumsegenskaben****, hvis enhver ikke tom opadtil begrænset delmængde har et suppremum.*

Øvelse 244 *Overvej om de hele tal* $(\mathbb{Z}, \leq)$ *med den sædvanlige ordning har supremumsegenskaben.*

Eksempel 245 *De rationale tal med den sædvanlige ordning har ikke supremumsegenskaben.*

Bevis. Analyse: For at bevise dette skal vi angive en ikke tom delmængde af $\mathbb{Q}$, som er opadtil begrænset, men som ikke har et supremum. Hvordan skal vi gætte en kandidat? Vi kommer senere til at vise (eller postulere), at de reelle tal har suppremumsegenskaben. Enhver ikke tom opadtil begrænset delmængde i $\mathbb{Q}$, vil også være en ikke tom opadtil begrænset delmængde af $\mathbb{R}$. Mængden har altså et supremum i $\mathbb{R}$. Vi skal altså sørge for, at dette supremum ikke ligger i $\mathbb{Q}$, det skal altså være et irrationalt tal. Vi har vist at $\sqrt{2}$ er irrationalt (se Sætning 50), så hvis vi kan finde en mængde af rationale tal, som i $\mathbb{R}$ har supremum $\sqrt{2}$, har vi en god kandidat. En sådan mængde er $B = \left\{x \in \mathbb{Q} \mid x < \sqrt{2}\right\}$. Nu er det pænere at undlade at involvere de reelle tal i en diskussion af de rationale tals egenskaber. Derfor vil vi helst angive mængden uden at involvere $\sqrt{2}$. Det kan vi gøre ved i stedet at se på mængden $A = \left\{x \in \mathbb{Q} \mid x^2 < 2\right\}$. Denne mængde er ikke lig med B, idet den ikke indeholder tallene mindre end $-\sqrt{2}$, men dens supremumsegenskab er er naturligvis de samme. Efter denne analyse, kan vi gå over til det syntetiske bevis.

Beviset. Betragt mængden

$$A = \left\{x \in \mathbb{Q} \mid x^2 < 2\right\}. \tag{7.30}$$

Den er en delmængde af $\mathbb{Q}$, og da $0 \in A$ er den ikke tom. Da 2 er en majorant (overvej), er mængden endvidere opadtil begrænset. Vi vil bevise at A ikke har et supremum i $\mathbb{Q}$. Beviset føres ved modstrid.

Antag altså, at $b \in \mathbb{Q}$ er supremum for A. Da $1 \in A$ og 2 er en majorant for A, gælder det, at $1 \leq b \leq 2$. Da $\mathbb{Q}$ er totalt ordnet, gælder et af følgende udsagn: $b^2 < 2$, $2 < b^2$, $b^2 = 2$. Vi vil udelukke de to første muligheder.

1. Antag først at $b^2 < 2$. Vi vil vise, at dette er i modstrid med antagelsen om at b er en majorant for A. Det gør vi ved at finde et tal i A, som er større end b.[1] Da vi har antaget at $b^2 < 2$, vil $d = \frac{2-b^2}{5}$ være et positivt rationalt tal. Tallet $c = b + d$ er derfor et rationalt tal større end b. Men c ligger i A, thi, da $b > 1$ er $d < \frac{2-1}{5} = \frac{1}{5} < 1$, og da endvidere $1 \leq b \leq 2$, kan vi slutte som følger:

$$c^2 = (b+d)^2 = b^2 + 2bd + d^2 < b^2 + 2 \cdot 2d + 1d \tag{7.31}$$

$$= b^2 + 5d = b^2 + 5\left(\frac{2-b^2}{5}\right) = b^2 + 2 - b^2 = 2. \tag{7.32}$$

   Vi har altså fundet et element c af A, som er større end b. Dette strider mod at b er en majorant for A. Vi kan altså udelukke at $b^2 < 2$.

2. Antag dernæst at $2 < b^2$. Vi vil vise at det er i modstrid med, at b er den *mindste* majorant for A. Det gør vi ved at finde en majorant for A, som er mindre end b.[2] Da vi har antaget at $2 < b^2$, vil $d = \frac{b^2-2}{4}$ være et positivt rationalt tal. Tallet $c = b - d$ er derfor et rationalt tal mindre end b. Vi vil vise at c er en majorant for A. Da $b \leq 2$ gælder nemlig at

$$c^2 = (b-d)^2 = b^2 - 2bd + d^2 > b^2 - 2bd \tag{7.33}$$

$$\geq b^2 - 2 \cdot 2d = b^2 - 4\left(\frac{b^2-2}{4}\right) = 2. \tag{7.34}$$

   Så hvis $x > c$ vil $x^2 > c^2 > 2$ (her bruges at $c \geq 0$). hvorfor $x \notin A$. Kontraposition af dette udsagn giver at $x \in A \Rightarrow x \leq c$, som netop betyder at c er en majorant for A. Vi har altså fundet en majorant c, som er mindre end b, i modstrid med at b var antaget at være et supremum for A. Vi har altså udelukket at $2 < b^2$.

Da vi hverken har $b^2 < 2$ eller $2 < b^2$ må det altså gælde at $b^2 = 2$. Men det strider mod Sætning 50.

Vi har dermed bevist at mængden A ikke har et supremum i $\mathbb{Q}$, hvorfor $\mathbb{Q}$ ikke har supremumsegenskaben.

Helt analogt til supremum, defineres infimum: ■

Definition 246 *Lad A være en delmængde af en totalt ordnet mængde $(M, \leq)$. Et element b i M kaldes **infimum** for A, hvis det er den største minorant for A, d.v.s opfylder følgende to kriterier:*

[1] Her skal vi altså finde en kandidat. I det følgende trækkes denne op af hatten. Prøv selv at lave den analyse, som fører frem til kandidaten.

[2] Prøv selv at lave analysen.

1. *b er en minorant for A, altså: $\forall a \in A : b \leq a$.*

2. *b er den største minorant for A, altså: Hvis x er en minorant for A, da er $x \leq b$.*

Hvis b er infimum for A, skriver man $b = \inf A$

Sætning 247 *Lad A være en delmængde af en totalt ordnet mængde $(M, \leq)$. Da er $b = \inf A$ hvis og kun hvis*

1. *b er en minorant for A, altså: $\forall a \in A : b \leq a$.*

2. *$\forall x > b \exists a \in A : a < x$.*

Bevis. Overlades til læseren. ■

Sætning 248 *Lad A være en delmængde af en totalt ordnet mængde $(M, \leq)$. Hvis A har et mindste element a, da vil a også være infimum for A.*

Bevis. Bevis overlades til læseren. ■

Øvelse 249 *Afgør om følgende mængder har et infimum i $\mathbb{R}$, og bestem infimum, hvis det findes:*

1. *$]1, 2[$.*

2. *$[1, 2]$.*

3. *$]-\infty, 0[$.*

4. *$\left\{\frac{1}{n} \mid n \in \mathbb{N}\right\}$*

Definition 250 *En totalt ordnet mængde siges at have **infimumsegenskaben**, hvis enhver ikke tom nedadtil begrænset delmængde har et infimum.*

Sætning 251 *En totalt ordnet mængde har supremumsegenskaben, hvis og kun hvis den har infimumsegenskaben.*

Bevis. Lad $(M, \leq)$ være en totalt ordnet mængde. Vi vil vise, at hvis den har supremumsegenskaben, da har den også infimumsegenskaben. Det bevises helt analogt, at hvis $(M, \leq)$ har infimumsegenskaben, da har den også supremumsegenskaben.

Antag altså at $(M, \leq)$ har supremumsegenskaben. Vi vil da vise, at den har infimumsegenskaben. Lad derfor A være en ikke tom nedadtil begrænset delmængde af M. Vi skal vise at A har et infimum. Lad N_A betegne mængden af minoranter for A altså:

$$N_A = \{x \in M \mid x \text{ er en minorant for } A\}. \tag{7.35}$$

Ideen i beviset er nu at vise, at N_A er ikke tom og opadtil begrænset, hvorfor den iflg supremumsegenskaben har et supremum. Dette supremum vil vi så vise er et infimum for A.

Da A er antaget at være nedadtil begrænset, har A en minorant, så N_A er ikke tom. Da endvidere A er antaget at være ikke tom, kan vi vælge et $a \in A$. Hvis $x \in N_A$ er x en minorant for A, hvorfor $x \leq a$. Det betyder at a er en majorant for N_A. Altså er N_A ikke tom og opadtil begrænset, og da vi har antaget at $(M, \leq)$ har supremumsegenskaben, har N_A et supremum. Sæt $s = \sup N_A$. Vi vil vise, at s er infimum af A.

Først vises, at s er en minorant for A: Lad derfor a være et vilkårligt element i A. Vi så ovenfor, at a er en majorant for N_A. Da nu s er den mindste majorant for N_A, må den være mindre eller lig med a: $s \leq a$. Da dette gælder for alle $a \in A$ er s en minorant for A.

Dernæst vises, at s er den største minorant for A: Antag nemlig at z er en anden minorant for A. Da er pr. definition $z \in N_A$. Da $s = \sup N_A$, og s således specielt er en majorant for N_A, vil $z \leq s$. Altså er s en største minorant for A, hvorfor $s = \inf A$.

Vi konkluderer derfor at $(M, \leq)$ har infimumsegenskaben. ■

Øvelse 252 *Gennemfør den udeladte halvdel af ovenstående bevis, altså beviset for at hvis $(M, \leq)$ har infimumsegenskaben, da har den også supremumsegenskaben. Prøv at lære det ovenstående bevis så godt, at du kan gennemføre den manglende del med lukket bog. Det lærer du mere af end hvis du "oversætter" ovenstående bevis ord for ord med åben bog.*

7.3 Opgaver

1. På $\mathbb{R}$ defineres relationen xRy ved

$$xRy \Leftrightarrow xy > 0 \tag{7.36}$$

Undersøg, om relationen er refleksiv, symmetrisk, antisymmetrisk eller transitiv.

Besvar samme spørgsmål, når xRy har betydningen

1. $xy \geq 0$
2. $xy^2 > 0$
3. $xy^2 \geq 0$
4. $x^2y^2 > 0$
5. $x^2y^2 \geq 0$
6. $x^2 - y^2 \geq 0$

2. Lad R være en symmetrisk og transitiv relation på en mængde M. Hvori består fejlen i følgende "bevis" for, at R er refleksiv:

Hvis aRb, gælder det at bRa, da R er symmetrisk. Men da R er transitiv, følger det fra aRb og bRa, at aRa. Altså er R refleksiv.

Giv et eksempel på en relation, der er symmetrisk og transitiv, men ikke refleksiv.

3. Lad P være et givet punkt i planen. Definer relationen $\sim$ på mængden af punkter i planen ved:

$$A \sim B \Leftrightarrow |PA| = |PB| . \tag{7.37}$$

Vis, at $\sim$ er en ækvivalensrelation, og beskriv dens ækvivalensklasser.

4. Lad F betegne mængden af reelle funktioner definerede på et interval I. På F defineres relationerne $R_1, ..., R_4$ på følgende måde:

1. $fR_1g \dot{\Leftrightarrow} f - g$ er konstant på I
2. $fR_2g \dot{\Leftrightarrow} f - g$ er et førstegradspolynomium
3. $fR_3g \dot{\Leftrightarrow} f - g$ er et andengradspolynomium
4. $fR_4g \dot{\Leftrightarrow} (f - g)(x) \neq 0$ for højst endeligt mange x i I.

Undersøg hvilke af relationerne, der er ækvivalensrelationer.

5. Lad $M = \{1, 2, 3, 4, 5, 6\}$. Betragt følgende relation på M :

$$\sim \; = \left\{ \begin{array}{c} (1,1), (2,2), (3,3), (4,4), (5,5), (6,6), (1,2), \\ (1,4), (2,1), (2,4), (4,1)\,(4,2), (3,6), (6,3) \end{array} \right\}. \tag{7.38}$$

Bevis at $\sim$ er en ækvivalensrelation.

Bestem $[a]$ for ethvert element $a \in M$, og angiv $M \diagup \sim$

6. Vis at de følgende relationer på de angivne mængder er ækvivalensrelationer. Gør det i hvert tilfælde på to måder:

- Ved at bestemme ækvivalensklasserne og vise at de giver en klassedeling af mængden
- Ved at vise direkte, at $\sim$ er refleksiv, symmetrisk og transitiv.

1. På mængden af danskere defineres $A \sim B$, hvis A og B er født samme år.
2. På $\mathbb{Z}$ defineres $a \sim b$, hvis $|a| = |b|$.

7. Bevis at følgende relationer er ækvivalensrelationer på de angivne mængder, og beskriv deres ækvivalensklasser:

1. $M = \mathbb{R}$, $a \sim b \dot{\Leftrightarrow} a - b \in \mathbb{Z}$
2. $M = \mathbb{R}$, $a \sim b \dot{\Leftrightarrow} \lfloor a \rfloor = \lfloor b \rfloor$, hvor $\lfloor a \rfloor$ betegner heltalsdelen af a, dvs. det største hele tal mindre eller lig med a.
3. $M = \mathbb{R}^2$, $(x_1, y_1) \sim (x_2, y_2) \dot{\Leftrightarrow} x_1^2 + y_1^2 = x_2^2 + y_2^2$.

8. Lad Q betegne den følgende delmængde af $\mathbb{Z} \times \mathbb{Z}$:

$$Q = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid b \neq 0\}. \tag{7.39}$$

Definer relationen $\sim$ på Q ved

$$(a, b) \sim (c, d) \dot{\Leftrightarrow} ad = bc. \tag{7.40}$$

Bevis, at $\sim$ er en ækvivalensrelation, og angiv ækvivalensklassen $[(2, 3)]$ og mere generelt ækvivalensklassen $[(a, b)]$.

Prøv at give en beskrivelse af $Q/\sim$.

9. Bevis at relationen $n \mid m$ er en partiel ordning på $\mathbb{N}$. Er det en total ordning?

10. Nedenfor angives en række delmængder af $\mathbb{R}$ med den sædvanlige ordning. Afgør i hvert tilfælde om

- mængden er opadtil begrænset, nedadtil begrænset og begrænset,
- mængden har et største element og et mindste element,
- mængden har et supremum og et infimum.

1. $[2, 6]$
2. $]2, 6[$
3. $]-\infty, 24693]$
4. $\mathbb{Q}_+$

11. Betragt $\mathbb{R}$ med den sædvanlige ordning. Bevis i alle detaljer, at

1. $\sup[0, 1[= 1$
2. $\inf\left\{\frac{1}{n} \mid n \in \mathbb{N}\right\} = 0$

12. Lad $(M, \leq)$ være en totalt ordnet mængde, og antag at $\emptyset \neq A \subseteq B \subseteq M$. Antag endvidere, at de nedenstående infima og suprema eksisterer: Bevis da, at

$$\inf B \leq \inf A \leq \sup A \leq \sup B. \tag{7.41}$$

Giv et eksempel, hvor $\inf A = \inf B$ og $\sup B = \sup A$, selv om A er en ægte delmængde af B.

Gælder sætningen også, hvis vi ikke antager at A og B er ikke tomme?

13. Lad A og B være ikke tomme delmængder af $\mathbb{R}$ (udstyret med den sædvanlige ordning), hvorom det gælder at $a < b$ for alle $a \in A$ og $b \in B$.

1. Bevis i alle detaljer at $\sup A$ og $\inf B$ begge eksisterer, og at

$$\sup A \leq \inf B. \tag{7.42}$$

I beviset må du gerne bruge, at $\mathbb{R}$ har supremumsegenskaben.

2. Antag desuden, at $A \cup B = \mathbb{R}$. Bevis, at

$$\sup A = \inf B. \tag{7.43}$$

14. I definitionen af supremumsegenskaben krævede vi, at enhver *ikke tom* opadtil begrænset delmængde skulle have et supremum. Overvej, at det er vigtigt at begrænse kravet til ikke tomme delmængder. Betragt for eksempel $\mathbb{R}$ og delmængden $\emptyset$. Er $\emptyset$ opadtil begrænset? Har $\emptyset$ en mindste majorant?

Chapter 8

Afbildninger, funktioner

I dette kapitel skal vi præcisere det funktionsbegreb, I har brugt i gymnasiet og de indledende matematikkurser. Vi skal indføre grundlæggende begreber om funktioner og vise nogle sætninger om disse begreber.

I gymnasiet har I mødt reelle funktioner. En reel funktion er måske blevet indført som en forskrift eller en regel, som afbilder et reelt tal over i et andet reelt tal. For eksempel afbilder funktionen x^2 et reelt tal x over i det reelle tal x^2. Men hvad betyder det at afbilde, og hvad skal vi forstå ved en regel eller en forskrift? For præcisere disse begreber, vil vi føre dem tilbage til de basale begreber i mængdelæren. Desuden vil vi generalisere funktionsbegrebet til andre mængder end mængder af reelle tal.

Ideen til hvordan funktionsbegrebet skal præciseres og generaliseres kan vi få ved at bemærke, at en funktion kan opfattes som en relation. For eksempel kan funktionen x^2 opfattes som den relation R på $\mathbb{R}$, som er defineret ved at xRy hvis $x^2 = y$. Vi vil derfor definere en funktion som en speciel slags relation. Ved at betragte relationer mellem andre mængder end $\mathbb{R}$, vil vi kunne generalisere funktionsbegrebet til funktioner fra en vilkårlig mængde A ind i en vilkårlig anden mængde B. Og da vi jo formelt har defineret en relation mellem to mængder A og B som en delmængde af produktmængden $A \times B$, vil vi derved få defineret funktionsbegrebet i rent mængdeteoretiske termer, uden brug af uldne ord som "forskrift" eller "regel".

Det er ikke enhver relation, som kan opfattes som en funktion. For at definere en funktion som en bestemt slags relation, skal vi altså have fat i den eller de egenskaber, som udmærker de relationer, der definerer funktioner. Betragt derfor en relation xRy, som er defineret ud fra en reel funktion f ved at $xRy \Leftrightarrow y = f(x)$. Denne relation har to særlige egenskaber:

1. Ethvert reelt x har et reelt billede $y = f(x)$, eller udtrykt i relationssprog: For ethvert $x \in \mathbb{R}$ findes et $y \in \mathbb{R}$, så xRy.

2. Ethvert reelt x har kun ét reelt billede, altså der er kun ét reelt y så $y = f(x)$. Udtrykt i relationssprog betyder det, at for ethvert $x \in \mathbb{R}$

findes der højst ét $y \in \mathbb{R}$ så xRy. Det kan også udtrykkes således: Hvis xRy_1 og xRy_2, da er $y_1 = y_2$.

Efter denne analyse kan vi nu formulere den formelle definition af en funktion eller en afbildning:

Definition 253 *Lad A og B være ikke tomme mængder. En relation f mellem A og B kaldes en **afbildning** eller en **funktion** fra A ind i B (og vi skriver $f : A \longrightarrow B$), hvis følgende to krav er opfyldt:*

1. *For alle $x \in A$ findes et $y \in B$, så at xfy (eller $(x, y) \in f$).*
2. *Hvis xfy_1 og xfy_2 (eller $(x, y_1) \in f$ og $(x, y_2) \in f$), da gælder at $y_1 = y_2$.*

Øvelse 254 *Lad $A = \{1, 2, 3, 4\}$ og $B = \{a, b, c, d\}$. Afgør, om følgende delmængder af $A \times B$ er afbildninger fra A ind i B.*

1. $\{(1, a), (2, b), (3, c), (4, d)\}$.
2. $\{(1, a), (2, a), (3, b), (4, b)\}$
3. $\{(1, a), (1, b), (2, c), (2, d), (3, a), (4, d)\}$
4. $\{(1, a), (2, c), (3, b)\}$

Notation 255 *Hvis f er en afbildning fra A ind i B og $x \in A$, da findes altså et entydigt $y \in B$, så xfy. Dette element y betegnes med $f(x)$ (siges:"f af x"). I stedet for xfy eller $(x, y) \in f$ skriver man derfor sædvanligvis $y = f(x)$.*

Bemærkning 256 *Hvis A og B er mængder af reelle tal, kan vi illustrere afbildningen f ved at indtegne mængden i et retvinklet koordinatsystem i planen. At en delmængde C af $A \times B$ af planen er en funktion, betyder da geometrisk, at enhver lodret linje, der skærer x-aksen i et punkt af A, har præcist ét punkt fælles med C. Det betyder nemlig, at der til ethvert $x \in A$ findes ét $y \in B$ så $(x, y) \in C$.*

Eksempel 257 *I $[-1, 1] \times [-1, 1]$ betragtes de fire mængder:*

1. *Enhedscirkelskiven: $D = \left\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\right\}$*
2. *Enhedscirklen: $C = \left\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\right\}$*
3. *Den højre del af enhedscirklen: $C_h = \left\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \wedge x \geq 0\right\}$*
4. *Den øvre del af enhedscirklen: $C_o = \left\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \wedge y \geq 0\right\}$.*

Den sidste mængde "er" en afbildning fra $[-1, 1]$ ind i sig selv, da der til hvert x i $[-1, 1]$ findes netop et $y \in [-1, 1]$, så $(x, y) \in C_o$.

D er ikke en afbildning fra $[-1, 1]$ ind i sig selv, da der til alle x'er i $]-1, 1[$ svarer mere end et $y \in [-1, 1]$ så $(x, y) \in D$. Det samme gælder C. C_h er heller ikke en afbildning fra $[-1, 1]$ ind i sig selv, da der til negative x'er ikke svarer noget $y \in [-1, 1]$, så $(x, y) \in C_h$.

Eksempel 258 *Lad $\mathcal{C}$ betegne mængden af cirkler i planen Π. Lad f betegne følgende delmængde af $\mathcal{C} \times \Pi$: $f = \{(C, p) \in \mathcal{C} \times \Pi \mid p \text{ er centrum for } C\}$. Da er f en afbildning, og $f(C) =$ centrum for C. f er altså den afbildning, som afbilder en cirkel i sit centrum.*

Bemærkning 259 *I dansksproget litteratur bruges ordene funktion og afbildning ofte i flæng. Dog bruges ordet funktion fortrinsvist om en afbildning, som afbilder ind i de reelle eller komplekse tal. Vi vil i disse noter følge denne tradition. Afbildninger af planen eller rummet ind i sig selv kaldes også transformationer, og afbildninger mellem mængder af funktioner kaldes normalt operatorer.*

Selv om vi formelt har defineret en funktion eller afbildning $f : A \longrightarrow B$ som en relation mellem A og B, dvs. som en delmængde af $A \times B$, så er det nyttigt at bibeholde den intuitive idé af en funktion eller afbildning som en "maskine", som sender elementer x fra A over i elementer $f(x)$ i B. Den almindelige sprogbrug om funktioner og afbildninger afspejler bedre denne intuitive, men upræcise opfattelse af funktioner. Vi kalder for eksempel $f(x)$ for billedet (eller funktionsværdien) af x, og vi kan tale om at "anvende" funktionen f på x. Den delmængde af $A \times B$ som ifølge definitionen "er" funktionen f, kalder man sædvanligvis funktionens **graf**.

Det er derimod ikke hensigtsmæssigt at tænke på en funktion som en regneforskrift. Selv om reelle funktioner ofte er defineret ved at angive y som en formel i x, så kan mange (selv reelle) funktioner ikke angives på denne form.

Ofte illustrerer man afbildninger ved figurer i stil med Figur 8.1.

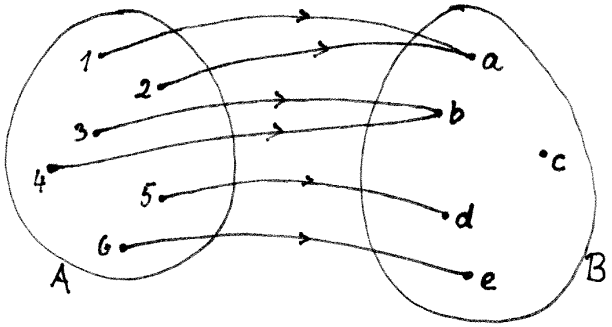

Figure 8.1: Afbildning af $A = \{1, 2, 3, 4, 5, 6\}$ ind i $B = \{a.b.c.d.e\}$

Lad f være en afbildning fra A ind i B. Da kaldes A for **definitionsmængden** og B kaldes **sekundærmængden** for f.

Definition 260 *Mængden af funktionsværdier $\{y \in B \mid \exists x \in A : y = f(x)\}$ kaldes afbildningens **billedmængde** eller **værdimængde**.*

Bemærkning 261 *I dansksproget litteratur hersker der en vis forvirring om betydningen af ordene billedmængde og værdimængde. I nogle fremstillinger*

(f.eks. tidligere noter til MatM) er det mængden B, som betegnes billedmængden eller værdimængden.

Bemærkning 262 *Når man skal definere en afbildning, skal man altså angive tre ting:*

1. *Definitionsmængden A.*
2. *Sekundærmængden B.*
3. *Funktionsværdien $f(x)$ for ethvert $x \in A$.*

Bemærkning 263 *To afbildninger f og g er lig med hinanden hvis*

1. *de har samme definitionsmængde,*
2. *de har samme sekundærmængde,*
3. *og for alle x i definitionsmængden gælder $f(x) = g(x)$*

Ofte sjusker man når man angiver funktioner. Det er således meget almindeligt at tale om funktionen x^2. Dette er upræcist, fordi man ikke har angivet definitionsmængden og sekundærmængden. Det vil dog ofte være klart fra sammenhængen, hvad disse mængder er (f.eks. $\mathbb{R}$ eller $\mathbb{C}$). Når man i reel analyse taler om funktionen $1/x$ eller andre funktioner, som ikke naturligt er defineret på hele $\mathbb{R}$, så underforstås det ofte, at man som definitionsmængde skal tage den største mængde, hvor funktionen er defineret, altså i dette tilfælde $\mathbb{R} \setminus \{0\}$.

Angivelsen af en funktion ved udtrykkene x^2 og $1/x$ illustrerer også en anden unøjagtighed, som man ofte tillader sig ved omtalen af funktioner. I den generelle definition skelner vi mellem funktionen f og dens værdi $f(x)$ i et givet element $x \in A$. Men ofte omtaler man funktionen som $f(x)$ i stedet for som f. Det er for eksempel tilfældet med funktionerne x^2 og $1/x$, som jo vanskeligt lader sig angive uden at skrive x'et. Hvis man vil angive disse funktioner med en notation, der tydeliggør, at der er tale om funktionen, og ikke dens værdi i punktet x, kan man skrive: "funktionen $x \longrightarrow x^2$", men det gøres sjældent.

8.1 Injektivitet og surjektivitet

Definition 264 *1. En afbildning $f : A \longrightarrow B$ kaldes **surjektiv** (udtales "syrjektiv"), hvis værdimængden er hele B, altså hvis ethvert y i B er en funktionsværdi, eller udtrykt formelt:*

$$\forall y \in B \exists x \in A : y = f(x) \tag{8.1}$$

*Man siger da at f afbilder A **på** B.*

*2. En afbildning $f : A \longrightarrow B$ kaldes **injektiv (eller en-entydig)**, hvis ethvert y i B højst er billede af ét x i A, dvs. hvis*

$$f(x_1) = f(x_2) \Rightarrow x_1 = x_2 \tag{8.2}$$

3. *En afbildning $f : A \longrightarrow B$ kaldes* ***bijektiv,*** *hvis den er både surjektiv og injektiv, altså hvis ethvert $y \in B$ er billede af præcist ét $x \in A$. En bijektiv afbildning kaldes også en* ***bijektion****.*

Bemærkning 265 *Et bevis for at en afbildning er surjektiv er et eksistensbevis. Et bevis for at en afbildning er injektiv er et entydighedsbevis.*

Øvelse 266 *Hvilke af afbildningerne i 254 er surjektive og hvilke er injektive?*

Øvelse 267 *Tegn figurer med boller og pile, som illustrerer afbildninger som er surjektive og injektive, og afbildninger, som ikke er surjektive og injektive. Forklar, hvordan man på figurerne kan se om en afbildning er injektiv, og hvordan man ser at den er surjektiv.*

Bemærkning 268 *Hvis en funktion $f : A \longrightarrow B$ mellem to reelle talmængder repræsenteres ved sin graf i et retvinklet koordinatsystem i planen, da er den surjektiv, hvis enhver vandret linje gennem et punkt i B på y-aksen skærer grafen mindst en gang. Den er injektiv, hvis enhver vandret linje højst skærer grafen i ét punkt med x-koordinat i A..*

Eksempel 269 *Overvej at følgende påstande er sande:*

1. *Funktionen* $\sin : \mathbb{R} \longrightarrow \mathbb{R}$ *er hverken surjektiv eller injektiv.*
2. *Funktionen* $\sin : \mathbb{R} \longrightarrow [-1, 1]$ *er surjektiv men ikke injektiv.*
3. *Funktionen* $\sin : [-\pi/2, \pi/2] \longrightarrow \mathbb{R}$ *er injektiv men ikke surjektiv.*
4. *Funktionen* $\sin : [-\pi/2, \pi/2] \longrightarrow [-1, 1]$ *er bijektiv.*

Bemærkning 270 *Når man skal afgøre om en afbildning er surjektiv og injektiv, er det altså vigtigt at specificere definitionsmængden og sekundærmængden præcist. Generelt kan en vilkårlig funktion gøres surjektiv ved at indskrænke sekundærmængden til billedmængden.*

Eksempel 271 *Funktionen $f : \mathbb{R} \longrightarrow \mathbb{R}$ defineret ved $f(x) = 2x + 7$ er injektiv.*

Bevis. Der gælder følgende implikationer:

$$\begin{aligned} f(x_1) &= f(x_2) && (8.3)\\ &\Leftrightarrow 2x_1 - 7 = 2x_2 - 7 && (8.4)\\ &\Leftrightarrow 2x_1 = 2x_2 && (8.5)\\ &\Leftrightarrow x_1 = x_2 && (8.6) \end{aligned}$$

■

Eksempel 272 *Funktionen $f : \mathbb{R} \longrightarrow \mathbb{R}$ defineret ved $f(x) = x^3 - x$ er ikke injektiv.*

Bevis. For at vise at funktionen ikke er injektiv skal vi finde to forskellige x'er, som afbildes i den samme værdi. Her kan vi betragte tallene 0 og 1. Der gælder jo at $0 \neq 1$ men $f(0) = f(1) = 0$. ■

8.2 Billeder og urbilleder

Definition 273 *Lad A, B være mængder, og $f : A \longrightarrow B$ en afbildning. Lad endvidere $T \subseteq A$. Delmængden $f(T)$ af B defineret ved*

$$f(T) = \{y \in B \mid \exists t \in T : y = f(t)\} \tag{8.7}$$

kaldes for ***billedet*** *af T under f.*

Bemærkning 274 *Billedet af definitionsmængden $f(A)$ er altså billedmængden.*

Eksempel 275 *Betragt funktionen $f : \mathbb{R} \longrightarrow \mathbb{R}$ defineret ved $f(x) = x^2$. Da er $f(]-2, 10[) = [0, 100[$.*

Sætning 276 *Lad A og B være mængder, og lad $f : A \to B$ være en afbildning.*
Lad $\{T_\alpha\}_{\alpha\in\Lambda}$ være en familie af delmængder af A.
Da gælder følgende:

1. $f(\bigcup_{\alpha\in\Lambda} T_\alpha) = \bigcup_{\alpha\in\Lambda} f(T_\alpha)$.
2. $f(\bigcap_{\alpha\in\Lambda} T_\alpha) \subseteq \bigcap_{\alpha\in\Lambda} f(T_\alpha)$.

Bevis. *Bevis for 1.* Lad $y \in f(\bigcup_{\alpha\in\Lambda} T_\alpha)$. Der findes da $x \in \bigcup_{\alpha\in\Lambda} T_\alpha$, så $y = f(x)$. Da $x \in \bigcup_{\alpha\in\Lambda} T_\alpha$, findes $\alpha \in \Lambda$, så $x \in T_\alpha$. Da $y = f(x)$, er dermed $y \in f(T_\alpha)$.

Vi har altså påvist eksistensen af et $\alpha \in \Lambda$, så $y \in f(T_\alpha)$, hvormed vi har vist, at $y \in \bigcup_{\alpha\in\Lambda} f(T_\alpha)$.

Altså gælder $f(\bigcup_{\alpha\in\Lambda} T_\alpha) \subseteq \bigcup_{\alpha\in\Lambda} f(T_\alpha)$. Den omvendte inklusion vises, idet man ser, at ovenstående ræsonnement kan "vendes om" (overvej dette).

Bevis for 2. Lad $y \in f(\bigcap_{\alpha\in\Lambda} T_\alpha)$. der findes da $x \in \bigcap_{\alpha\in\Lambda} T_\alpha$, så $y = f(x)$. Da $x \in \bigcap_{\alpha\in\Lambda} T_\alpha$, gælder for ethvert $\alpha \in \Lambda$, at $x \in T_\alpha$. Da $y = f(x)$, haves dermed $y \in f(T_\alpha)$, for ethvert $\alpha \in \Lambda$. Med andre ord gælder $y \in \bigcap_{\alpha\in\Lambda} f(T_\alpha)$. ■

Bemærkning 277 *Som øvelse bør man overveje, hvorfor beviset for (2) i den foregående sætning ikke også kan 'vendes om'. Man bør også konstruere et modeksempel til påstanden om, at den omvendte inklusion i (2) skulle være alment gyldig (det er den nemlig ikke).*

Bemærkning 278 *Sætning 276 er faktisk et resultat af Sætning 32 (1.27) og (1.29) og den manglende inklusion er et resultat af bemærkning 33. Hvis vi bruger implikationerne i Sætning 32, kan beviset for sætning 276 nemlig føres i næsten ren symbolsk form som følger:*

Bevis. *Bevis for sætning 276, 1. For ethvert $y \in B$* gælder følgende biimplikationer:

$$
\begin{aligned}
& y \in f(\bigcup_{\alpha\in\Lambda} T_\alpha) \\
& \Updownarrow \\
& \exists x \in A \ : \ x \in \bigcup_{\alpha\in\Lambda} T_\alpha \ \wedge \ y = f(x) \\
& \Updownarrow \\
& \exists x \in A \ \exists \alpha \in \Lambda \ : \ x \in T_\alpha \ \wedge \ y = f(x) \\
& \Updownarrow \\
& \exists \alpha \in \Lambda \ \exists x \in A \ : \ x \in T_\alpha \ \wedge \ y = f(x) \\
& \Updownarrow \\
& \exists \alpha \in \Lambda \ : \ y \in f(T_\alpha) \\
& \Updownarrow \\
& y \in \bigcup_{\alpha\in\Lambda} f(T_\alpha) \ ,
\end{aligned}
$$

hvoraf påstanden følger.

Bevis for sætning 276, 2. For ethvert $y \in B$ har vi følgende implikationer:

$$
\begin{aligned}
& y \in f(\bigcap_{\alpha\in\Lambda} T_\alpha) \\
& \Updownarrow \\
& \exists x \in A \ : \ x \in \bigcap_{\alpha\in\Lambda} T_\alpha \ \wedge \ y = f(x) \\
& \Updownarrow \\
& \exists x \in A \ \forall \alpha \in \Lambda \ : \ x \in T_\alpha \ \wedge \ y = f(x) \\
& \Downarrow \\
& \forall \alpha \in \Lambda \ \exists x \in A \ : \ x \in T_\alpha \ \wedge \ y = f(x) \\
& \Updownarrow \\
& \forall \alpha \in \Lambda \ : \ y \in f(T_\alpha) \\
& \Updownarrow \\
& y \in \bigcap_{\alpha\in\Lambda} f(T_\alpha) \ ,
\end{aligned}
$$

og det ønskede følger. ■

Definition 279 *Lad A, B være mængder, og $f : A \longrightarrow B$ en afbildning. Lad endvidere $S \subseteq B$. Delmængden $f^{-1}(S)$ af A defineret ved*

$$f^{-1}(S) = \{x \in A \mid f(x) \in C\} \tag{8.8}$$

kaldes ***urbilledet*** *(eller originalmængden) af S under f.*

Øvelse 280 *Tegn en figur med boller og pile til at illustrere urbilledet af en mængde.*

Det følger direkte fra definitionen af urbilledet at elementerne i $f^{-1}(S)$ kan karakteriseres som følger:

Sætning 281 *Lad A, B være mængder og $f : A \longrightarrow B$ en afbildning. Lad endvidere $S \subseteq B$. Da gælder:*

$$x \in f^{-1}(S) \Leftrightarrow f(x) \in S \tag{8.9}$$

Eksempel 282 *Betragt funktionen $f : \mathbb{R} \longrightarrow \mathbb{R}$ defineret ved $f(x) = x^2$. Da er*

1. $f^{-1}(\{4\}) = \{-2, 2\}$,
2. $f^{-1}([4, 9]) = [2, 3] \cup [-3, -2]$
3. $f^{-1}(\{-4\}) = \emptyset$

Sætning 283 *Lad A og B være mængder, og lad $f : A \to B$ være en afbildning.*

Lad $\{S_\alpha\}_{\alpha\in\Lambda}$ være en familie af delmængder af B, og lad S være en delmængde af B.

Da gælder følgende:

1. $f^{-1}(\bigcup_{\alpha\in\Lambda} S_\alpha) = \bigcup_{\alpha\in\Lambda} f^{-1}(S_\alpha)$.
2. $f^{-1}(\bigcap_{\alpha\in\Lambda} S_\alpha) = \bigcap_{\alpha\in\Lambda} f^{-1}(S_\alpha)$.
3. $f^{-1}(B\backslash S) = A \setminus f^{-1}(S)$.

Bevis. *Bevis for 1: For ethvert $x \in A$* gælder følgende biimplikationer:

$$\begin{aligned}
x \quad &\in \quad f^{-1}(\bigcup_{\alpha\in\Lambda} S_\alpha) && (8.10)\\
&\Leftrightarrow \quad f(x) \in \bigcup_{\alpha\in\Lambda} S_\alpha && (8.11)\\
&\Leftrightarrow \quad \exists\alpha \in \Lambda : f(x) \in S_\alpha && (8.12)\\
&\Leftrightarrow \quad \exists\alpha \in \Lambda : x \in f^{-1}(S_\alpha) && (8.13)\\
&\Leftrightarrow \quad x \in \bigcup_{\alpha\in\Lambda} f^{-1}(S_\alpha), && (8.14)
\end{aligned}$$

hvilket viser påstanden.

Bevis for 2: analogt med beviset for 1.

Bevis for 3: For ethvert $x \in A$ gælder følgende biimplikationer:

$$\begin{aligned}
x \quad &\in \quad f^{-1}(B\backslash S) && (8.15)\\
&\Leftrightarrow \quad f(x) \in B\backslash S && (8.16)\\
&\Leftrightarrow \quad \neg f(x) \in S && (8.17)\\
&\Leftrightarrow \quad \neg x \in f^{-1}(S) && (8.18)\\
&\Leftrightarrow \quad x \in A \setminus f^{-1}(S), && (8.19)
\end{aligned}$$

hvilket viser påstanden. ■

8.3 Sammensætning af afbildninger, invers afbildning

Definition 284 *Lad A, B og C være mængder og lad $f : A \longrightarrow B$ og $\dot{g} : B \longrightarrow C$ være afbildninger. Da defineres den* ***sammensatte afbildning*** *$g \circ f : A \longrightarrow C$ ved:*

$$\text{For alle } x \in A \text{ er } g \circ f(x) = g(f(x)). \tag{8.20}$$

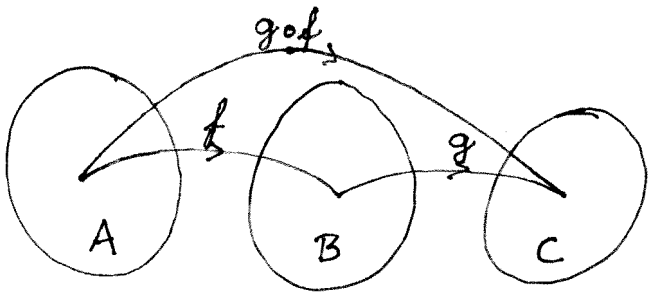

Figure 8.2: Den sammensatte funktion $g \circ f$

Bemærkning 285 *Funktionen $g \circ f$ fås altså ved først at anvende f og dernæst anvende g. Funktionerne skal altså anvendes i omvendt rækkefølge. Man har valgt denne skrivemåde fordi man derved får den simple definition: $g \circ f(x) = g(f(x))$.*

Eksempel 286 *Definer funktionerne $f : \mathbb{R} \longrightarrow \mathbb{R}$ og $g : \mathbb{R} \longrightarrow \mathbb{R}$ ved $f(x) = 2x + 7$ og $g(x) = x^2$. Da er funktionerne $g \circ f : \mathbb{R} \longrightarrow \mathbb{R}$ og $f \circ g : \mathbb{R} \longrightarrow \mathbb{R}$ beskrevet ved*

$$(g \circ f)(x) = g(2x + 7) = (2x + 7)^2 \tag{8.21}$$

og

$$(f \circ g)(x) = f(x^2) = 2x^2 + 7 \tag{8.22}$$

Øvelse 287 *Lad $A = \{1, 2, 3, 4\}$ og $B = \{a, b, c, d\}$. Definer $f : A \longrightarrow B$ og $\dot{g} : B \longrightarrow A$ ved deres grafer på følgende vis:*

$$f = \{(1, a), (2, a), (3, b), (4, b)\} \tag{8.23}$$
$$g = \{(a, 2), (b, 3), (c, 4), (d, 4)\} \tag{8.24}$$

Bestem graferne for $g \circ f$ og $f \circ g$.

Sætning 288 *Lad A, B og C være mængder og lad $f : A \longrightarrow B$ og $\dot{g} : B \longrightarrow C$ være afbildninger. Da gælder følgende:*

1. *Hvis f og g begge er surjektive, da er $g \circ f : A \longrightarrow C$ surjektiv.*
2. *Hvis f og g begge er injektive, da er $g \circ f : A \longrightarrow C$ injektiv.*
3. *Hvis f og g begge er bijektive, da er $g \circ f : A \longrightarrow C$ bijektiv.*

Bevis. *Bevis for 1*: Antag at f og g begge er surjektive, og lad $z \in C$. Da g er surjektiv, findes et $y \in B$ så $z = g(y)$. Da f er surjektiv, findes endvidere et $x \in A$ så $y = f(x)$. Alt i alt har vi altså

$$z = g(y) = g(f(x)) = g \circ f(x) \tag{8.25}$$

Vi har altså vist, at der til ethvert $z \in C$, findes et $x \in A$, så $z = g \circ f(x)$. Det betyder netop at $g \circ f$ er surjektiv.

Bevis for 2: Antag, at f og g begge er injektive. Lad $x_1, x_2 \in A$, og antag at $g \circ f(x_1) = g \circ f(x_2)$. Vi skal vise at $x_1 = x_2$.

1. Da g er injektiv, og $g(f(x_1)) = g(f(x_2))$, kan vi slutte at $f(x_1) = f(x_2)$, og da f er injektiv, kan vi herfra slutte at $x_1 = x_2$.

Bevis for 3: Følger af 1. og 2. ■

Der gælder en delvis omvending af denne sætning:

Sætning 289 *Lad A, B og C være mængder, og lad $f : A \longrightarrow B$ og $g : B \longrightarrow C$ være afbildninger. Da gælder følgende:*

1. *Hvis $g \circ f : A \longrightarrow C$ er surjektiv, da er g surjektiv.*
2. *Hvis $g \circ f : A \longrightarrow C$ er injektiv, da er f injektiv.*

Bevis. *Bevis for 1*: Antag at $g \circ f : A \longrightarrow C$ er surjektiv. Vi skal vise, at g er surjektiv. Lad derfor $z \in C$. Da $g \circ f$ er surjektiv, findes et $x \in A$, så $z = g \circ f(x) = g(f(x))$. Men så findes jo et $y \in B$, så $z = g(y)$, nemlig $y = f(x)$.

Bevis for 2: Antag at $g \circ f : A \longrightarrow C$ er injektiv. Vi skal vise, at f er injektiv. Lad $x_1, x_2 \in A$, og antag at $f(x_1) = f(x_2)$. Vi skal vise at $x_1 = x_2$.

Da $f(x_1) = f(x_2)$, gælder

$$g \circ f(x_1) = g(f(x_1)) = g\left(f(x_2)\right) = g \circ f(x_2), \tag{8.26}$$

og da $g \circ f$ var antaget injektiv, slutter vi heraf at $x_1 = x_2$. ■

Øvelse 290 *Angiv eksempler på situationer, hvor A, B og C er mængder, og hvor $f : A \longrightarrow B$ og $g : B \longrightarrow C$ er afbildninger, og hvor*

1. *g er surjektiv, men $g \circ f$ ikke er surjektiv.*
2. *f er injektiv, men $g \circ f$ ikke er injektiv.*

3. *$g \circ f$ er surjektiv men f er ikke surjektiv.*

4. *$g \circ f$ er injektiv, men g er ikke injektiv.*

Giv to slags eksempler: dels eksempler, illustreret med mængdeboller og pile, i tilfælde hvor mængderne er endelige, og dels eksempler hvor mængderne $A, B,$ og C alle er de reelle tal.

Sætning 291 *Lad A, B, C og D være mængder, og lad $f : A \longrightarrow B$ og $g : B \longrightarrow C$ og $h : C \longrightarrow D$ være afbildninger. Da gælder:*

$$h \circ (g \circ f) = (h \circ g) \circ f \tag{8.27}$$

Bevis. Overlades til læseren. ■

På grund af denne sætning kan vi tillade os at skrive $h \circ g \circ f$ i stedet for $h \circ (g \circ f)$ eller $(h \circ g) \circ f$. Lignende betragtninger gælder naturligvis ved sammensætning af flere end tre afbildninger.

Definition 292 *Lad A være en mængde. Med 1_A betegnes den **identiske afbildning** på A, dvs. afbildningen givet ved*

$$1_A(a) = a \text{ for alle } a \in A \tag{8.28}$$

Afbildningen 1_A er naturligvis bijektiv.

Definition 293 *Lad A og B være mængder og lad $f : A \to B$ være en afbildning. En afbildning $g : B \longrightarrow A$ kaldes*

1. *en **venstreinvers** til f, hvis $g \circ f = 1_A$,*

2. *en **højreinvers** til f, hvis $f \circ g = 1_B$,*

3. *en **invers** til f, hvis g er både højre og venstreinvers til f, altså hvis $g \circ f = 1_A$ og $f \circ g = 1_B$.*

Øvelse 294 *Overvej, om følgende funktioner har en højreinvers og en venstreinvers, og angiv en eller flere sådanne, hvis de findes:*

1. Funktionen $f : \mathbb{R} \longrightarrow \mathbb{R}$ defineret ved $f(x) = 4x - 3$.

2. Funktionen $f : \mathbb{R} \longrightarrow \mathbb{R}$ defineret ved $f(x) = x^2$.

3. Funktionen $f : \mathbb{R} \longrightarrow \mathbb{R}_+ \cup \{0\}$ defineret ved $f(x) = x^2$.

4. Funktionen $f : \mathbb{R} \longrightarrow [-1, 1]$ defineret ved $f(x) = \sin x$.

Sætning 295 *Lad A og B være mængder, og lad $f : A \to B$ være en afbildning. Hvis f har en invers afbildning, da er den entydigt bestemt.*

Bevis. Antag at $g : B \longrightarrow A$ og $h : B \longrightarrow A$ begge er inverse til f. Da gælder

$$g = g \circ 1_B = g \circ (f \circ h) = (g \circ f) \circ h = 1_A \circ h = h \tag{8.29}$$

■

Bemærkning 296 *Derimod er højre- og venstreinverser ikke nødvendigvis entydige. Hvis du ikke allerede har indset dette i forbindelse med Øvelse 294, bør du gøre det nu.*

Sætning 297 *Lad A og B være mængder, og lad $f : A \to B$ være en afbildning. Der gælder følgende:*

1. *f har en venstreinvers, hvis og kun hvis f er injektiv.*
2. *f har en højreinvers, hvis og kun hvis f er surjektiv.*
3. *f har en invers, hvis og kun hvis f er bijektiv.*

Bevis. *Bevis for 1*: Antag først, at f har en venstreinvers g. Da gælder pr. definition $g \circ f = 1_A$. Da 1_A er injektiv, følger det af sætning 289.2 at f er injektiv.

Antag omvendt, at f er injektiv. Vi skal da vise, at der findes en venstreinvers g til f. Vi skal med andre ord konstruere en afbildning $g : B \longrightarrow A$, så $g \circ f = 1_A$.

Bevis. Vi bemærker, at hvis b ligger i $f(A)$, findes netop ét $a \in A$ så $f(a) = b$. For da f er injektiv, findes der højst et sådant a, og da $b \in f(A)$ findes der mindst et sådant a. Vælger vi nu et eller andet element $a_0 \in A$, kan vi dermed definere en afbildning $g : B \longrightarrow A$ ved:

$$g(b) = \begin{cases} a, & \text{hvis} \quad b = f(a) \text{ for et } a \in A \\ a_0, & \text{hvis} \quad b \notin f(A) \end{cases} \tag{8.30}$$

Vi får da for $a \in A$, at $g(f(a)) = a$, altså at $g \circ f = 1_A$. ■

Bevis for 2: Hvis f har en højreinvers $g : B \longrightarrow A$, så gælder at $f \circ g = 1_B$, og da 1_B er surjektiv, følger det af 289, at f er surjektiv.

Antag omvendt at f er surjektiv. Vi skal da vise, at der findes en højreinvers g til f. Vi skal med andre ord konstruere en afbildning $g : B \longrightarrow A$, så $f \circ g = 1_B$.

Bevis. Når f er surjektiv betyder det at $f^{-1}(\{b\}) \neq \emptyset$ for ethvert $b \in B$. Vælg da for $b \in B$ et element $a_b \in f^{-1}(\{b\})$.[1] Ifølge sætning 281 gælder da, at $f(a_b) = b$ for ethvert $b \in B$. Defineres derfor en afbildning $g : B \longrightarrow A$ ved:

$$g(b) = a_b, \tag{8.31}$$

fås at

$$(f \circ g)(b) = f(g(b)) = f(a_b) = b \text{ for ethvert } b \in B, \tag{8.32}$$

dvs. at $f \circ g = 1_B$. ■

[1] Her bruger vi et omdiskuteret aksiom i mængdelæren, det såkaldte udvalgsaksiom.

Bevis for 3: Hvis f har en invers, er denne både en højre- og venstreinvers, så af 1. og 2. fås, at f er både surjektiv og injektiv, altså bijektiv.

Omvendt hvis f er bijektiv er den både surjektiv og injektiv og har derfor både en venstreinvers g og en højreinvers h som opfylder

$$g \circ f = 1_A \text{ og } f \circ h = 1_B \tag{8.33}$$

Hvis vi nu kan vise at der nødvendigvis må gælde $g = h$, så har vi bevist eksistensen af en afbildning, som er både højre- og venstreinvers, altså af en invers. Men at $g = h$ følger af følgende udregning:

$$g = g \circ 1_B = g \circ (f \circ h) = (g \circ f) \circ h = 1_A \circ h = h \tag{8.34}$$

■

Definition 298 *Hvis $f : A \to B$ er bijektiv, betegnes dens entydigt bestemte inverse afbildning med f^{-1}. Der gælder altså:*

$$f^{-1} \circ f = 1_A \text{ og } f \circ f^{-1} = 1_B \tag{8.35}$$

Man kalder også f's inverse for den ***omvendte afbildning****.*

Bemærkning 299 *Når $f : A \to B$ er en bijektiv afbildning, er f^{-1} karakteriseret ved, at der for alle $(x, y) \in A \times B$ gælder:*

$$x = f^{-1}(y) \Leftrightarrow f(x) = y. \tag{8.36}$$

Øvelse 300 *Lad $f : A \to B$ være en bijektiv afbildning. Bevis, at*

$$\left(f^{-1}\right)^{-1} = f \tag{8.37}$$

Øvelse 301 *Lad $f : A \to B$ og $g : B \longrightarrow C$ være bijektive afbildninger. Bevis, at*

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1} \tag{8.38}$$

Det er en del af opgaven at bevise, at de angivne funktioner eksisterer.

Bemærkning 302 *Når $f : A \to B$ er en afbildning, har vi brugt symbolet f^{-1} i to forskellige betydninger: 1. som betegnelsen for den inverse afbildning og. 2. i betegnelsen for urbilledet $f^{-1}(S)$ af en delmængde $S \subseteq B$. Disse to brug af symbolet f^{-1} må ikke forveksles. Man kan danne $f^{-1}(S)$, uanset om f er bijektiv eller ej, hvorimod man kun kan tale om den inverse afbildning f^{-1}, når f er bijektiv. Men hvis f er bijektiv, er der et potentielt problem: symbolet $f^{-1}(S)$ kan jo læses på to forskellige måder:*

1. *Som urbilledet af S under f.*
2. *Som billedet ved f^{-1} af S.*

Heldigvis bliver $f^{-1}(S)$ samme delmængde af A, uanset hvordan man opfatter symbolet.

8.4 Kardinalitet

Dette afsnit kradser kun lige lidt i overfladen af et stort og spændende emne. Vi nævner blot nogle sætninger uden bevis.

Definition 303 *To mængder siges at have samme* ***kardinalitet*** *eller* ***mægtighed****, hvis der findes en bijektion mellem dem.*

Sætning 304 *Relationen "samme kardinalitet" er en ækvivalensrelation mellem mængder.*

Sætning 305 *Hvis* $m, n \in \mathbb{N}$ *og* $m \neq n$*, så har* $\{1, 2, 3, ..., n\}$ *og* $\{1, 2, 3, ..., m\}$ *ikke samme kardinalitet.*

Definition 306 *Lad* M *være en mængde.*

Hvis der findes et $n \in \mathbb{N}$*, og en bijektion mellem* M *og* $\{1, 2, 3, ..., n\}$*, siges* M *at være* ***endelig****, og* n *kaldes antallet af* M*'s elementer.*

Hvis M *ikke er endelig, kaldes den* ***uendelig****.*

Sætning 307 *To endelige mængder har samme kardinalitet, hvis og kun hvis de har samme antal elementer.*

En endelig og en uendelig mængde har ikke samme kardinalitet.

Sætning 308 *Hvis* A *er en endelig mængde, og* $B \subseteq A$*, så er* B *også endelig.*

Sætning 309 $\mathbb{N}$*,* $\mathbb{Z}$*, og* $\mathbb{Q}$ *har samme kardinalitet. Man siger, de er* ***tællelige****.*

Sætning 310 $\mathbb{N}$ *og* $\mathbb{R}$ *har ikke samme kardinalitet.*

Sætning 311 $\mathbb{R}$ *har samme kardinalitet som* $\mathbb{C}$ *og som* $\mathbb{R}^n$ *for alle* $n \in \mathbb{N}$*. Man siger, de har* ***kontinuets kardinalitet****.*

Sætning 312 *Lad* M *være en mængde. Da har potensmængden* $P(M)$ *ikke samme kardinalitet som* M*.*

8.5 Opgaver

1. Betragt mængderne

$$A = \{1, 2, 3, 4, 5\} \tag{8.39}$$

og

$$B = \{a, b, c, d\} \tag{8.40}$$

Hvilke af følgende mængder repræsenterer relationer, som er afbildninger fra den ene mængde ind i den anden:

1. $\{(1, a), (2, b), (3, d), (4, d), (5, d)\}$
2. $\{(1, a), (b, 3), (c, 4), (5, d), (2, a)\}$

3. $\{(a,3),(b,4),(c,5),(d,1),(a,2)\}$

4. $\{(a,3),(b,2),(c,2),(d,5)\}$

5. $\{(1,a),(2,b),(3,c),(4,d)\}$

Tegn boller med pile til at illustrere disse funktioner
2. Betragt afbildningen på Figur 8.1:
Bestem følgende elementer og mængder

1. $f(3)$ og $f(6)$

2. $f(\{1,2,3\})$ og $f(\{4,5,6\})$

3. $f^{-1}(\{a,e\})$, $f^{-1}(\{b\})$, $f^{-1}(\{c,d\})$ og $f^{-1}(\{c\})$.

3. Giv eksempler på endelige mængder A og B og afbildninger $f : A \longrightarrow B$ så:

1. f er injektiv, men ikke surjektiv

2. f er surjektiv men ikke injektiv

3. f er både surjektiv og injektiv

4. f er hverken surjektiv eller injektiv.

4. Giv eksempler på afbildninger $f : \mathbb{R} \longrightarrow \mathbb{R}$, som opfylder de fire punkter i opgave 3.

5. Lad U være en given grundmængde. For enhver delmængde M af U definerer vi en afbildning $k_M : U \longrightarrow \{0,1\}$, kaldet den karakteristiske afbildning, ved:

$$k_M(x) = \begin{cases} 1 & \text{når} \quad x \in M \\ 0 & \text{når} \quad x \in \complement M \end{cases} \tag{8.41}$$

Vis at når A og B er vilkårlige delmængder af U, gælder:

$$k_{A\cap B} = k_A \cdot k_B, \quad \text{og} \quad k_{A\cup B} = k_A + k_B - k_A \cdot k_B. \tag{8.42}$$

Angiv afbildningerne

$$k_\emptyset, \quad k_U \quad \text{og} \quad k_{\complement A}. \tag{8.43}$$

Hvad kan man sige om mængderne A og B, når det for alle $x \in U$ gælder

$$k_A(x) \leq k_B(x)? \tag{8.44}$$

6. Bevis hvorvidt hver af de nedenstående funktioner $\mathbb{R} \longrightarrow \mathbb{R}$ er injektive og surjektive:

1. $f(x) = \frac{1}{2}x - 5$

2. $f(x) = 543x^3$

3. $f(x) = x^5 - x$
4. $f(x) = e^x$
5. $f(x) = \cos x$

7. Betragt funktionen $f = \sin : \mathbb{R} \longrightarrow \mathbb{R}$. Bestem

$$f(\mathbb{R}_+) \quad \text{og} \quad f^{-1}(\mathbb{R}_+) \tag{8.45}$$

8. Lad funktionerne f og g $\mathbb{R} \longrightarrow \mathbb{R}$ være givet ved

$$f(x) = x^2 - 1 \quad \text{og} \quad g(x) = x^4 - 1. \tag{8.46}$$

Bevis, at

$$g(\mathbb{Q}) \subset f(\mathbb{Q}) \subset \mathbb{Q}. \tag{8.47}$$

Vis endvidere at hvis $x \in g(\mathbb{Q})$ er $\sqrt{x+1}$ rational.

9. Vis, at når $y \neq 1$, har ligningen

$$\frac{x}{x+1} = y \tag{8.48}$$

netop én løsning; og når $y = 1$ har ligningen ingen løsning.

Vis derved, at funktionen

$$f(x) = \frac{x}{x+1} \tag{8.49}$$

er en bijektion af mængden $\mathbb{R} \setminus \{-1\}$ på $\mathbb{R} \setminus \{1\}$, og angiv den inverse funktion.

10. Givet funktionen $f : \mathbb{R} \longrightarrow \mathbb{R}$ defineret ved

$$f(x) = \frac{x}{|x|+1}. \tag{8.50}$$

Bestem $f(\mathbb{R})$, og vis at f er bijektiv $\mathbb{R} \longrightarrow f(\mathbb{R})$

Find $f^{-1}(y)$ for $y \geq 0$ og for $y < 0$, og vis at f^{-1} kan angives ved

$$f^{-1}(y) = \frac{y}{1-|y|}. \tag{8.51}$$

11. Lad $f : A \longrightarrow B$ være en afbildning. Definer en relation på A ved

$$x \sim y \Leftrightarrow f(x) = f(y). \tag{8.52}$$

Bevis at $\sim$ er en ækvivalensrelation. Beskriv ækvivalensklasserne.

12. Lad $f : A \longrightarrow B$ være en afbildning.

1. Vis at for alle delmængder M af A gælder

$$M \subseteq f^{-1}(f(M)). \tag{8.53}$$

2. Giv et eksempel på at $f^{-1}(f(M))$ ikke behøver være lig med M.

Chapter 9

Kompositionsregler, grupper og isomorfier

9.1 Kompositionsregler

Regneoperationerne som f.eks. addition $+$ og multiplikation $\cdot$ kan opfattes som afbildninger fra $\mathbb{R}^2$ ind i $\mathbb{R}$, defineret ved, at de afbilder $(x, y) \in \mathbb{R}^2$ over i henholdsvist $x + y$ og $x \cdot y$. Hvis F betegner mængden af funktioner af en mængde M ind i sig selv, kan $\circ$ (sammensætning af funktioner) på lignende vis opfattes som en afbildning, der afbilder F^2 ind i F, nemlig den afbildningen der afbilder $(f, g) \in F^2$ over i $f \circ g$. Vi generaliserer denne idé, idet vi definerer begrebet kompositionsregel:

Definition 313 *En kompositionsregel på en mængde M er en afbildning fra $M \times M$ ind i M.*

Notation 314 *Kompositionsregler benævnes ofte med tegn som $\oplus$, $\otimes$, $+$ eller $\cdot$.*

Billedet $\otimes(x, y)$ af $(x, y) \in M \times M$ betegnes med $x \otimes y$.
En mængde M udstyret med en kompositionsregel $\otimes$ betegnes med $(M, \otimes)$.

Eksempel 315 *Her følger en række eksempler på kompositionsregler:*

1. *Addition og multiplikation på $\mathbb{R}$.*
2. *Sammensætning $\circ$ af funktioner i mængden af funktioner af en mængde M ind i sig selv.*
3. *Når M er en mængde, er $\cap$ og $\cup$ kompositionsregler på mængden $P(M)$ af delmængder af M.*
4. *Forskriften $x \oplus y = \frac{1}{2}(x+y)$ (gennemsnitsdannelse) er en kompositionsregel på $\mathbb{R}$.*

5. *Matrixmultiplikation på mængden $Mat_n(\mathbb{R})$ af $n \times n$ matricer.*

6. *Sammensætning af funktioner på mængden af lineære afbildninger af et vektorrum ind i sig selv.*

7. *Addition og multiplikation på $\mathbb{Z}/n$.*

Øvelse 316 *Overvej om de elementære regneoperationer $+, -, \cdot, :$ er kompositionsregler på følgende talmængder: $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{Q}_+, \mathbb{R}, \mathbb{R}_+, \mathbb{R} \setminus \{0\}$. Det der afgør sagen, er om kompositionsreglen (lad os kalde den $\otimes$) er defineret for alle par (x, y) i mængden, og om $x \otimes y$ igen ligger i mængden.*

Definition 317 *Lad $\otimes$ være en kompositionsregel på mængden M.*

1. *$\otimes$ siges at være **associativ,** hvis det for alle $x, y, z \in M$ gælder at*

$$x \otimes (y \otimes z) = (x \otimes y) \otimes z \tag{9.1}$$

2. *$\otimes$ siges at være **kommutativ,** hvis det for alle $x, y \in M$ gælder at*

$$x \otimes y = y \otimes x \tag{9.2}$$

Eksempel 318 *I Eksempel 315 er kompositionsreglerne i 1, 3 og 7. både associative og kommutative.*

Sammensætning af funktioner er associativ (se Sætning 291) men ikke kommutativ (se eksempel 286). Det samme gælder matrixmultiplikation.

Gennemsnitsdannelse nævnt i punkt 4 er kommutativ men ikke associativ (overvej dette).

Subtraktion på $\mathbb{R}$ er hverken kommutativ eller associativ. (overvej).

Bemærkning 319 *Når en kompositionsregel $\otimes$ er associativ, kan man tillade sig at skrive $x \otimes y \otimes z$ i stedet for $x \otimes (y \otimes z)$ eller $(x \otimes y) \otimes z$.*

Når der er tale om associative kompositionsregler, kan man i det hele taget tillade sig at sætte og hæve parenteser uden at ændre i udtryk af formen $x_1 \otimes x_2 \otimes \cdots \otimes x_n$ (overvej dette).

Når en kompositionsregel $\otimes$ er kommutativ og associativ, kan man bytte om på leddenes orden i udtryk af formen $x_1 \otimes x_2 \otimes \cdots \otimes x_n$ (overvej).

Definition 320 *Lad $\otimes$ være en kompositionsregel på mængden M. Et element $e \in M$ kaldes et **neutralt element** for $\otimes$ (eller i $(M, \otimes)$), hvis det for ethvert $x \in M$ gælder, at*

$$x \otimes e = e \otimes x = x \tag{9.3}$$

Eksempel 321

1. *Tallet 0 er neutralt element i $(\mathbb{R}, +)$*

2. *Tallet 1 er neutralt element i $(\mathbb{R}, \cdot)$*

3. *Identitetsafbildningen 1_M er neutralt element for $\circ$ i mængden af funktioner af M ind i sig selv.*

4. *Den tomme mængde $\emptyset$ er neutralt element for $\cup$ på $P(M)$*
5. *M er neutralt element for $\cap$ på $P(M)$.*
6. *Identitetsmatricen er neutralt element for matrixmultiplikation.*
7. *$[0]$ og $[1]$ er neutrale elementer for henholdsvis addition og multiplikation i $\mathbb{Z}/n$.*

Øvelse 322 *Bevis at gennemsnitsdannelse beskrevet i Eksempel 315.4 ikke har noget neutralt element.*

Sætning 323 *Hvis en kompositionsregel $\otimes$ på en mængde M har et neutralt element, da er det entydigt bestemt.*

Bevis. Lad $\otimes$ være en kompositionsregel på mængden M og antag at e_1 og e_2 er neutrale elementer i $(M, \otimes)$. Da e_1 er et neutralt element gælder at $e_1 \otimes e_2 = e_2$. Da e_2 er et neutralt element gælder ligeledes at $e_1 \otimes e_2 = e_1$. Altså er $e_1 = e_2$. ■

Definition 324 *Lad $(M, \otimes)$ være en mængde med en kompositionsregel, og lad e være et neurralt element i $(M, \otimes)$. Ved et **inverst element** til et element $x \in M$ forstås et element $y \in M$, for hvilket*

$$x \otimes y = y \otimes x = e \tag{9.4}$$

Eksempel 325
1. *I $(\mathbb{R}, +)$ er $-x$ det inverse element til x.*
2. *I $(\mathbb{R}_+, \cdot)$ er $1/x$ det inverse element til x.*
3. *I Mængden S_M af bijektive afbildninger af en mængde M på sig selv, udstyret med kompositionsreglen $\circ$, er f^{-1} det inverse element til f. Hvis vi ikke indskrænker os til de bijektive afbildninger, har ikke alle elementer et inverst.*
4. *Hvis A er en invertibel $n \times n$ matrix, da er A^{-1} dens inverse element.*
5. *I $(\mathbb{Z}/n, +)$ er $[-x]$ det inverse element til $[x]$.*

Øvelse 326 *Overvej om ethvert element har et inverst element i følgende mængder med kompositionsregel:* $(\mathbb{N}_0, +)$, $(\mathbb{Z}, +)$, $(\mathbb{Z}, \cdot)$, $(\mathbb{Q}, +)$, $(\mathbb{Q}, \cdot)$, $(\mathbb{Q} \setminus \{0\}, \cdot)$, $(\mathbb{R}, \cdot)$, $(\mathbb{R} \setminus \{0\}, \cdot)$, $(\mathbb{Z}/n, \cdot)$

Sætning 327 *Lad $\otimes$ være en associativ kompositionsregel på mængden M, og lad e være et neutralt element for $\otimes$. Hvis et element $x \in M$ har et inverst element, da er det entydigt bestemt.*

Bevis. Antag, at y og z er inverse elementer til x altså, at $x \otimes y = y \otimes x = e$ og $x \otimes z = z \otimes x = e$. Da gælder:

$$y = y \otimes e = y \otimes (x \otimes z) = (y \otimes x) \otimes z = e \otimes z = z \tag{9.5}$$

■

Sætning 328 *Lad $\otimes$ være en associativ kompositionsregel på mængden M, og lad e være et neutralt element for $\otimes$. Da er e sit eget inverse element.*

Bevis. Vi skal checke at $e \otimes e = e$, men det følger af definitionen af det neutrale element. ■

Bemærkning 329 *Ligesom da vi definerede den inverse afbildning, kunne vi også her definere en højre- og en venstreinvers. Men vi vil nøjes med at indføre den inverse.*

9.2 Grupper

Mængder med en associativ kompositionsregel, hvori der findes et neutralt element, og hvori ethvert element har et inverst element, er så hyppigt forekommende i matematikken, at man giver dem et særligt navn:

Definition 330 *En **gruppe** $(G, \otimes)$ er en mængde G udstyret med en kompositionsregel $\otimes$, der har følgende egenskaber:*

1. *Kompositionsreglen $\otimes$ er associativ; d.v.s. for alle $x, y, z \in G$ gælder*

$$(x \otimes y) \otimes z = x \otimes (y \otimes z). \tag{9.6}$$

2. *Der findes et neutralt element $e \in G$; d.v.s. at for alle $x \in G$ gælder*

$$x \otimes e = e \otimes x = x. \tag{9.7}$$

3. *Ethvert element i G har et inverst element; d.v.s at for alle $x \in G$ findes et element, som vi vil kalde x^{-1}, for hvilket*

$$x \otimes x^{-1} = x^{-1} \otimes x = e \tag{9.8}$$

Antallet af elementer i gruppen kaldes gruppens orden. Hvis gruppen er uendelig er dens orden ∞.

Det fremgår af sætning 323 og 327, at det neutrale element i en gruppe er entydigt bestemt, og ligeledes at det inverse til et givet element også er entydigt. Derfor kan vi tillade os at tale om *det neutrale element, og give det navnet e* og vi kan tale om *det* inverse element til et element $x \in G$ og give det navnet x^{-1}.

Definition 331 *En gruppe $(G, \otimes)$ kaldes kommutativ eller abelsk*[1]*, hvis kompositionsreglen $\otimes$ er kommutativ.*

Eksempel 332 *Følgende er eksempler på abelske grupper:* $(\mathbb{R}, +)$, $(\mathbb{R} \setminus \{0\}, \cdot)$, $(\mathbb{Q}, +)$, $(\mathbb{Q} \setminus \{0\}, \cdot)$, $(\mathbb{Z}, +)$,$(\mathbb{Z}/n, +)$, $(\mathbb{C}, +)$, $(\mathbb{C} \setminus \{0\}, \cdot)$.

Følgende er eksempler på ikke abelske grupper:

[1] Efter den norske matematiker Niels Henrik Abel (1802-1829)

1. *Mængden $(S_M, \circ)$ af bijektive afbildninger af en mængde på sig selv, udstyret med kompositionsreglen $\circ$. Den kaldes den fulde transformationsgruppe af M.*

2. *Mængden af invertible reelle $n \times n$ matricer udstyret med kompositionsreglen matrixmultiplikation. Den kaldes den generelle lineære gruppe af grad n, og betegnes $GL_n(\mathbb{R})$*

3. *Mængden af bijektive lineære afbildninger af et n-dimensionalt vektorrum på sig selv udstyret med kompositionsreglen $\circ$.*

Sætning 333 *Lad $(G, \otimes)$ være en gruppe med neutralt element e. Da gælder*

$$e^{-1} = e \tag{9.9}$$

Bevis. Følger af 328. ■

Sætning 334 *Lad $(G, \otimes)$ være en gruppe og $x \in G$. Da gælder:*

$$\left(x^{-1}\right)^{-1} = x \tag{9.10}$$

Bevis. Vi skal vise at x er det inverse element til x^{-1}, altså at

$$x^{-1} \otimes x = x \otimes x^{-1} = e. \tag{9.11}$$

Men det gælder, fordi x^{-1} er det inverse element til x. ■

Sætning 335 *Lad $(G, \otimes)$ være en gruppe og $x, y \in G$. Da er*

$$(x \otimes y)^{-1} = y^{-1} \otimes x^{-1}. \tag{9.12}$$

Bevis. Ved gentagen brug af associativiteten fås

$$\begin{aligned} & (x \otimes y) \otimes \left(y^{-1} \otimes x^{-1}\right) && (9.13)\\ = \;& x \otimes (y \otimes \left(y^{-1} \otimes x^{-1}\right)) && (9.14)\\ = \;& x \otimes (\left(y \otimes y^{-1}\right) \otimes x^{-1}) && (9.15)\\ = \;& x \otimes (e \otimes x^{-1}) = x \otimes x^{-1} = e. && (9.16) \end{aligned}$$

På samme vis ses at $\left(y^{-1} \otimes x^{-1}\right) \otimes (x \otimes y) = e$. ■

Sætning 336 *Lad $(G, \otimes)$ være en gruppe med neutralt element e, og lad H være en ikke tom delmængde af G. Da er $(H, \otimes)$ en gruppe, hvis*

1. *H er lukket under $\otimes$, d.v.s.*

$$\forall x, y \in H : x \otimes y \in H \tag{9.17}$$

2. *$e \in H$*

3. For alle $x \in H$ gælder det, at $x^{-1} \in H$

Bevis. Lukketheden af $\otimes$ sikrer, at $\otimes$ er en kompositionsregel på H. Vi skal da blot checke, at de tre krav til en gruppe er opfyldt.

Associativiteten af $(H, \otimes)$ følger af associativiteten af $\otimes$ på hele G.

Da $e \in H$, er dette element også neutralt element i $(H, \otimes)$.

Et vilkårligt element $x \in H$ har et inverst x^{-1} i $(G, \otimes)$, og da dette er antaget at ligge i H, er det også et inverst i $(H, \otimes)$. ■

Definition 337 *I en situation som den ovenstående kaldes H en **undergruppe** af G.*

Eksempel 338 $(\mathbb{Q}, +)$ *og* $(\mathbb{Z}, +)$ *er undergrupper af* $(\mathbb{R}, +)$.
$(\mathbb{Q} \setminus \{0\}, \cdot)$ *er en undergruppe af* $(\mathbb{R} \setminus \{0\}, \cdot)$.

Eksempel 339 $\{1, -1, i, -i\}$ *er en undergruppe i* $(\mathbb{C} \setminus \{0\}, \cdot)$ *(overvej dette).*

Eksempel 340 *En afstandsbevarende afbildning af planen på sig selv kaldes en flytning. Mængden af flytninger er en undergruppe af mængden af bijektioner af planen på sig selv udstyret med kompositionsreglen $\circ$.*

Mængden af bijektive lineære afbildninger af et n-dimensionalt vektorrum på sig selv udstyret med kompositionsreglen $\circ$ er en undergruppe af mængden af bijektioner af vektorrummet på sig selv udstyret med kompositionsreglen $\circ$.

Kompositionsregler for endelige mængder, specielt endelige grupper, kan angives ved en **kompositionstavle**, som til ethvert par x, y i mængden angiver $x \otimes y$. Kompositionstavlen for $(\mathbb{Z}/4, +)$ ser således ud:

$+$	$[0]$	$[1]$	$[2]$	$[3]$
$[0]$	$[0]$	$[1]$	$[2]$	$[3]$
$[1]$	$[1]$	$[2]$	$[3]$	$[0]$
$[2]$	$[2]$	$[3]$	$[0]$	$[1]$
$[3]$	$[3]$	$[0]$	$[1]$	$[2]$

(9.18)

Kompositionstavlen for $(\{1, -1, i, -i\}, \cdot)$ ser således ud:

$\cdot$	1	i	-1	$-i$
1	1	i	-1	$-i$
i	i	-1	$-i$	1
-1	-1	$-i$	1	i
$-i$	$-i$	1	i	-1

(9.19)

Eksempel 341 *Betragt de flytninger af planen over i sig selv, som afbilder et rektangel, som ikke er et kvadrat, over i sig selv. Mængden af disse flytninger kalder vi K. K har fire elementer nemlig:*

e: Den identiske afbildning.

a: Spejlingen i midtnormalen for det ene sæt modstående sider.

b: Spejlingen i midtnormalen for det andet sæt modstsående sider.

c: En drejning på 180° *om skæringspunktet for de to omtalte midtnormaler.*

Fra 336 ses let, at $(K, \circ)$ *er en gruppe. Dens kompositionstavle ser således ud:*

$\circ$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

(9.20)

Denne gruppe kaldes ***Klein's firegruppe***[2]*.*

Grupper er et eksempel på en **matematisk struktur**. Andre eksempler er partielt ordnede mængder og mængder med en ækvivalensrelation. En matematisk struktur er en mængde som er udstyret med en kompositionsregel, en relation eller lignende. I en vis forstand er en mængde også en struktur, men en meget fattig struktur. Det eneste, der strukturerer den er "$\in$". Det karakteristiske ved en matematisk struktur er, at man ikke fastlægger hvad elementerne i mængden er for en slags objekter, og man ikke fastlægger hvilken kompositionsregel eller relation eller andet man har med at gøre. Man kræver bare, at visse eksplicit angivne aksiomer er opfyldt. Aksiomerne for en gruppe er de tre krav opremset i definition 330. Aksiomerne for en partielt ordnet mængde er angivet i Definition 209[3]. Senere i dette kursus vil vi indføre andre matematiske strukturer som legemer og ordnede legemer. I lineær algebra indføres vektorrum, som en vigtig slags matematisk struktur. Senere på studiet vil du stifte bekendtskab med andre strukturer som målrum og metriske rum.

Der er flere fordele ved at indføre og arbejde med matematiske strukturer.

1. Når man har vist en sætning om en matematisk struktur, gælder den for alle de konkrete eksempler på denne struktur. Man slår med andre ord mange fluer med et smæk.

2. Man får klargjort den logiske struktur i sin teori ved klart at fastlægge spillereglerne.

3. Det bliver lettere at gennemskue hvilke forudsætninger en sætning bygger på.

4. Ved at fjerne overflødige ting kan man ofte lettere finde beviser for sætninger.

Lad mig her især fremhæve punkt 1. Når man i lineær algebra viser en sætning om vektorrum, har man faktisk vist et væld af sætninger om forskellige vektorrum. Sætningen gælder jo uanset om elementerne i vektorrummet

[2] Efter den tyske matematiker Felix Klein (1849-1925).

[3] Læseren fornemmer nok at der ikke gives en præcis definition af en matematisk struktur. Det skyldes at der ikke findes nogen.

er reelle tal, vektorer i rummet eller funktioner, og uanset betydningen af vektoraddition og multiplikation med skalar. I algebrakurset vil du komme til at vise en lang række sætninger om grupper, og disse vil altså gælde for alle grupper, det være sig $(\mathbb{R},+)$, $(\mathbb{Z}/n,+)$, $(\mathbb{Z},+)$, $(K,\circ)$ eller en anden konkret gruppe. En gruppeteoretisk sætning indeholder altså i sig en lang række sætninger om tal, geometri mm. Lad os give et enkelt eksempel på en sådan generel sætning, som kunne have sparet os nogle overvejelser ovenfor:

Sætning 342 *Lad $(M,\otimes)$ være en mængde, med en associativ kompositionsregel og et neutralt element e. Mængden I af invertible elementer i $(M,\otimes)$ (altså de elementer, som har en invers) udgør en gruppe med kompositionsreglen $\otimes$.*

Bevis. Vi skal først vise at $\otimes$ er en kompositionsregel på I, altså at for $x,y\in I$ gælder, at $x\otimes y\in I$. Antag altså, at $x,y\in I$. Da har x og y inverser x^{-1} og y^{-1}. Som i sætning 335 ses at $y^{-1}\otimes x^{-1}$ er et inverst element til $x\otimes y$. Altså gælder $x\otimes y\in I$.

Vi skal dernæst vise at de tre krav til en gruppe er opfyldt af $(I,\otimes)$.

1. Associativiteten af $\otimes$ arves fra $(M,\otimes)$.

2. Det følger af sætning 328 at det neutrale element $e\in I$. Det er klart at e er det neutrale element i $(M,\otimes)$.

3. Lad $x\in I$. Vi skal vise at $x^{-1}\in I$ altså at x^{-1} har en invers. Men, som i sætning 334 ses at x er invers til x^{-1}, hvorfor $x^{-1}\in I$. ■

Eksempel 343 *Antag nu at vi allerede har vist, at kompositionsreglerne på $(\mathbb{Q},\cdot)$, $(\mathbb{R},\cdot)$ og på mængderne af afbildninger af en mængde ind i sig selv, samt på mængden af lineære afbildninger af et vektorrum ind i sig selv (begge udstyret med $\circ$) alle er associative, og at de indeholder et neutralt element. Da viser sætning 342 i et hug at de følgende mængder er grupper: $(\mathbb{Q}\setminus\{0\},\cdot)$, $(\mathbb{R}\setminus\{0\},\cdot)$, mængden S_M af bijektioner af en mængde på sig selv, og mængden af lineære bijektive afbildninger af et vektorrum på sig selv (begge udstyret med $\circ$). Dette illustrerer styrken i at bevise sætninger om matematiske strukturer.*

9.3 Gruppeisomorfier

Selv om der er mange forskellige grupper, er de altså fælles om alle gruppeteoriens sætninger. På denne måde ligner de alle hinanden. Men der er nogle grupper, der ligner hinanden mere end andre. Det er denne vage idé om lighed vi vil præcisere i dette afsnit. Vi vil ikke sammenligne grupper ud fra hvilken slags elementer de har. Om elementerne er tal, afbildninger, matricer, restklasser eller noget helt femte, er gruppeteorien uvedkommende. Det vi vil sammenligne er gruppernes struktur, altså hvordan kompositionsreglerne virker i grupperne.

Betragt for eksempel de tre grupper $(\mathbb{Z}/4,+)$, $(\{1,-1,i,-i\},\cdot)$, og $(K,\circ)$, hvis kompositionstavler vi skrev op ovenfor. I Kleins firegruppe har alle elementerne den egenskab, at de giver det neutrale element, hvis de sammensættes med sig selv: $x\otimes x=e$. Denne egenskab har de to andre grupper ikke.

De to grupper $(\mathbb{Z}/4, +)$ og $(\{1, -1, i, -i\}, \cdot)$ ligner hinanden meget. Faktisk har de samme kompositionstavle, hvis man ser bort fra elementernes navne. Hvis vi omdøber elementerne i $\{1, -1, i, -i\}$ og kalder dem $1 = \varphi([0])$, $i = \varphi([1])$, $-1 = \varphi([2])$ og $-i = \varphi([3])$, så vil kompositionstavlen for $(\{1, -1, i, -i\}, \cdot)$ få følgende udseende:

$\cdot$	$\varphi([0])$	$\varphi([1])$	$\varphi([2])$	$\varphi([3])$
$\varphi([0])$	$\varphi([0])$	$\varphi([1])$	$\varphi([2])$	$\varphi([3])$
$\varphi([1])$	$\varphi([1])$	$\varphi([2])$	$\varphi([3])$	$\varphi([0])$
$\varphi([2])$	$\varphi([2])$	$\varphi([3])$	$\varphi([0])$	$\varphi([1])$
$\varphi([3])$	$\varphi([3])$	$\varphi([0])$	$\varphi([1])$	$\varphi([2])$

(9.21)

Denne kompositionstavle er identisk med kompositionstavlen for $(\mathbb{Z}/4, +)$ bortset fra φ'erne. Vi kan derfor opfatte de to grupper som ens, bortset fra at elementerne hedder noget forskelligt. Vi vil sige at to grupper er isomorfe (græsk for "samme form"), hvis de ved omdøbning af elementerne, kan få samme kompositionstavle. Denne karakterisering er dog lidt tung, og den duer åbenbart kun for endelige grupper. Derfor vil vi omformulere isomorfibegrebet på en måde, som også kan bruges på uendelige grupper.

Ideen til denne generalisering får vi ved at betragte omdøbningen φ som en funktion $\mathbb{Z}/4 \longrightarrow \{1, -1, i, -i\}$. Denne funktion skal være bijektiv. For at kompositionstavlerne skal have samme form skal φ desuden have den egenskab, at det element i kompositionstavlen for $(\{1, -1, i, -i\}, \cdot)$, som angiver $\varphi(x) \cdot \varphi(y)$, skal være det samme element, som man får, ved at finde $x + y$ i kompositionstavlen for $(\mathbb{Z}/4, +)$ og anvende φ på dette element. Med andre ord, skal der gælde at $\forall x, y \in \mathbb{Z}/4 : \varphi(x + y) = \varphi(x) \cdot \varphi(y)$.

Denne karakterisering af isomorfibegrebet lader sig generalisere til vilkårlige grupper, og derfor vil vi bruge den som definition:

Definition 344 *Lad $(G, \otimes)$ og $(H, *)$ være to grupper. En afbildning $\varphi : G \longrightarrow H$ kaldes en **(gruppe)isomorfi,** hvis*

1. *φ er bijektiv,*
2. *for alle $x, y \in G$ gælder, at*

$$\varphi(x \otimes y) = \varphi(x) * \varphi(y). \tag{9.22}$$

*Hvis der findes en isomorfi $\varphi : (G, \otimes) \longrightarrow (H, *)$, siges $(G, \otimes)$ og $(H, *)$ at være **isomorfe**.*

Eksempel 345
1. *Den ovenfor nævnte omdøbning φ opfattet som en funktion $(\mathbb{Z}/4, +) \longrightarrow (\{1, -1, i, -i\}, \cdot)$ er en isomorfi.*
2. *Den afbildning, som sender en bijektiv lineær afbildning af et n-dimensionalt reelt vektorrum på sig selv over i den tilhørende matrix i en given basis, er en isomorfi af gruppen af bijektive lineære afbildninger på vektorrummet på $GL_n(\mathbb{R})$.*

Øvelse 346 *Vis, at følgende afbildninger er isomorfier:*

1. $x^2 : (\mathbb{R}_+, \cdot) \longrightarrow (\mathbb{R}_+, \cdot)$.
2. $kx : (\mathbb{R}, +) \longrightarrow (\mathbb{R}, +)$.
3. $\log : (\mathbb{R}_+, \cdot) \longrightarrow (\mathbb{R}, +)$.
4. $e^x : (\mathbb{R}, +) \longrightarrow (\mathbb{R}_+, \cdot)$.

Bemærkning 347 *Da en isomorfi specielt er bijektiv, har to isomorfe endelige grupper samme antal elementer. (se Sætning 307)*

Sætning 348 *Lad $\varphi : (G, \otimes) \longrightarrow (H, *)$ være en isomorfi. Da gælder*

$$\varphi(e_G) = e_H, \tag{9.23}$$

og for alle $x \in G$:

$$\varphi(x^{-1}) = (\varphi(x))^{-1}. \tag{9.24}$$

Bemærkning 349 *I formuleringen af sætningen betyder e_G og e_H naturligvis de neutrale elementer i henholdsvis $(G, \otimes)$ og $(H, *)$. Man skal være opmærksom på, at vi har brugt samme notation for dannelsen af den inverse i de to grupper.*

Bevis. af sætning 342. For at vise at $\varphi(e_G) = e_H$, skal vi vise, at $\varphi(e_G)$ har den definerende egenskab (9.3) altså at for alle $y \in H$ gælder, at

$$y * \varphi(e_G) = \varphi(e_G) * y = y. \tag{9.25}$$

Men da φ er bijektiv, findes der et $x \in G$, så $\varphi(x) = y$. Derfor kan vi fra (9.22) slutte, at

$$y * \varphi(e_G) = \varphi(x) * \varphi(e_G) = \varphi(x \otimes e_G) = \varphi(x) = y. \tag{9.26}$$

På samme måde ses, at $\varphi(e_G) * y = y$.

For at vise, at $\varphi(x^{-1}) = (\varphi(x))^{-1}$, skal vi vise, at $\varphi(x^{-1})$ er den inverse af $\varphi(x)$. Der gælder, at

$$\varphi(x^{-1}) * \varphi(x) = \varphi(x^{-1} \otimes x) = \varphi(e_G) = e_H. \tag{9.27}$$

På samme måde ses at $\varphi(x) * \varphi(x^{-1}) = e_H$, hvorfor $\varphi(x^{-1})$ er den inverse af $\varphi(x)$. ■

Sætning 350 *Lad $(G, \otimes)$, $(H, *)$ og $(J, \odot)$ være grupper. Da gælder:*

1. *Hvis $\varphi : (G, \otimes) \longrightarrow (H, *)$ og $\psi : (H, *) \longrightarrow (J, \odot)$ er isomorfier, da er $\psi \circ \varphi : (G, \otimes) \longrightarrow (J, \odot)$ en isomorfi.*
2. *Hvis $(G, \otimes)$ er isomorf med $(H, *)$ og $(H, *)$ er isomorf med $(J, \odot)$, da er $(G, \otimes)$ isomorf med $(J, \odot)$.*

Bevis. Overlades til læseren. ■

Eksempel 351 *Der findes ingen isomorfi af* $(K, \circ)$ *på* $(\mathbb{Z}/4, +)$.

Bevis. Beviset forløber indirekte. Antag derfor, at $\varphi : (K, \circ) \longrightarrow (\mathbb{Z}/4, +)$ er en isomorfi. Vælg et $a \in (K, \circ)$, hvorom det gælder at $a \neq e$ (for eksempel det element vi ovenfor kaldte a). Som vi bemærkede ovenfor er $a \circ a = e$. Deraf fås, at

$$\varphi(a) * \varphi(a) = \varphi(a \circ a) = \varphi(e) = [0]. \tag{9.28}$$

Sidste lighedstegn kommer fra (9.23), idet [0] er det neutrale element i $(\mathbb{Z}/4, +)$. Men fra kompositionstavlen for $(\mathbb{Z}/4, +)$ ses, at [0] og [2] er de eneste elementer i $(\mathbb{Z}/4, +)$, som sammensat med sig selv giver [0]. Derfor kan vi slutte, at $\varphi(a) = [0]$, eller $\varphi(a) = [2]$.

Antag først at $\varphi(a) = [0]$. Fra sætning 348 ved vi, at $\varphi(e) = [0]$. Altså er $\varphi(a) = \varphi(e)$, og da φ var antaget bijektiv, kan vi herfra slutte at $a = e$, i modstrid med, at vi havde antaget, at $a \neq e$. Altså kan vi slutte, at $\varphi(a) = [2]$.

Betragt nu et tredje element $b \in (K, \circ)$ forskelligt fra e og a. Som ovenfor kan vi slutte, at $\varphi(b) = [0] = \varphi(e)$, eller $\varphi(b) = [2] = \varphi(a)$. Da φ er bijektiv medfører det at $b = e$, eller $b = a$, i modstrid med forudsætningen.

Vi har dermed vist, at det fører til modstrid at antage, at $(K, \circ)$ og $(\mathbb{Z}/4, +)$ er isomorfe.

Der findes altså mindst to ikke isomorfe grupper af orden 4, nemlig Kleins firegruppe og $(\mathbb{Z}/4, +)$.

Grupperne, som er isomorfe med $(\mathbb{Z}/4, +)$ kaldes **cykliske** grupper af orden 4. Mere generelt kaldes grupper, der er isomorfe med $(\mathbb{Z}/n, +)$, for cykliske grupper af orden n.

I algebrakurset vil det blive bevist at når p er et primtal, så er alle grupper af orden p isomorfe med $(\mathbb{Z}/n, +)$. Sagt med andre ord: Der findes op til isomorfi kun én gruppe af primtalsorden, nemlig den cykliske gruppe. ■

Øvelse 352 *Betragt* $(\mathbb{Z}/n, \cdot)$. *Argumenter for at delmængden bestående af de invertible elementer er en gruppe. Denne gruppe kaldes* $(\mathbb{Z}/n)^*$.

Bestem mængderne $(\mathbb{Z}/3)^*$ *og* $(\mathbb{Z}/4)^*$, *og opstil kompositionstavlerne for de to grupper. Check om de er cykliske.*

9.4 Ordningsisomorfier

En gruppeisomorfi kan opfattes som en afbildning, som bevarer gruppens struktur. Man definerer generelt en isomorfi, som en bijektiv afbildning, som bevarer en matematisk struktur. For eksempel er bijektive lineære afbildninger vektorrumsisomorfier. På strukturen af partielt ordnede mængder definerer man en ordningsisomorfi som følger:

Definition 353 *Lad* $(A, \leq_A)$ *og* $(B, \leq_B)$ *være partielt ordnede mængder. En bijektion* $\varphi : A \longrightarrow B$ *kaldes en ordningsisomorfi, hvis det for alle* $x, y \in A$ *gælder, at*

$$x \leq_A y \Leftrightarrow \varphi(x) \leq_B \varphi(y) \tag{9.29}$$

To partielt ordnede mængder siges at være ordensisomorfe, hvis der findes en ordningsisomorfi mellem dem.

Sætning 354 *En bijektion $\varphi : (A, \leq_A) \longrightarrow (B, \leq_B)$ er en ordningsisomorfi, hvis og kun hvis der gælder:*

$$x <_A y \Leftrightarrow \varphi(x) <_B \varphi(y) \tag{9.30}$$

Bevis. Overlades til læseren. ■

Sætning 355 *Hvis $\varphi : (A, \leq_A) \longrightarrow (B, \leq_B)$ er en ordningsisomorfi, da er $\varphi^{-1} : (B, \leq_B) \longrightarrow (A, \leq_A)$ også en ordningsisomorfi.*

Bevis. Overlades til læseren. ■

Sætning 356 *Lad $(A, \leq_A)$ og $(B, \leq_B)$ være partielt ordnede mængder og $\varphi : (A, \leq_A) \longrightarrow (B, \leq_B)$ en ordningsisomorfi. Da er $(A, \leq_A)$ totalt ordnet, hvis og kun $(B, \leq_B)$ er totalt ordnet.*

Bevis. Antag at $(A, \leq_A)$ er totalt ordnet. Vi skal vise, at $(B, \leq_B)$ er totalt ordnet. Lad derfor s og t være vilkårlige elementer i B. Da φ er bijektiv, findes der $x, y \in A$, så $s = \varphi(x)$, og $t = \varphi(y)$. Da $(A, \leq_A)$ er totalt ordnet, ved vi pr. definition, at enten er $x \leq_A y$, eller også er $y \leq_A x$. Heraf og af (9.29) slutter vi, at enten er $s = \varphi(x) \leq_B \varphi(y) = t$, eller også er $t = \varphi(y) \leq_B \varphi(x) = s$. Altså er $(B, \leq_B)$ er totalt ordnet.

Beviset for "det omvendte" overlades til læseren. ■

Sætning 357 *Lad $(A, \leq_A)$ og $(B, \leq_B)$ være totalt ordnede mængder og $\varphi : (A, \leq_A) \longrightarrow (B, \leq_B)$ en ordningsisomorfi. Lad $C \subseteq A$ og $a \in A$. Da gælder:*

1. *a er en majorant for C, hvis og kun hvis $\varphi(a)$ er en majorant for $\varphi(C)$.*
2. *a er det største element i C, hvis og kun hvis $\varphi(a)$ er det største element i $\varphi(C)$.*
3. *$a = \sup C \Leftrightarrow \varphi(a) = \sup \varphi(C)$*

Der gælder lignende sætninger om minorant, mindste element og infimum.

Bevis. Vi beviser 1. og 3. og overlader 2. til læseren.

Bevis for 1. Antag, at a er en majorant for C. Vi skal vise, at $\varphi(a)$ er en majorant for $\varphi(C)$. Lad $s \in \varphi(C)$. Da findes der pr. definition et $c \in C$, så $s = \varphi(c)$. Da a er en majorant for C er $c \leq a$, og da φ er en ordningsisomorfi, følger det, at $s = \varphi(c) \leq \varphi(a)$. Altså er $\varphi(a)$ en majorant for $\varphi(C)$.

Antag omvendt, at $\varphi(a)$ er en majorant for $\varphi(C)$. Vi skal da vise, at a er en majorant for C. Lad $c \in C$. Da gælder at $\varphi(c) \in \varphi(C)$, og da $\varphi(a)$ er en majorant for $\varphi(C)$, medfører det, at $\varphi(c) \leq \varphi(a)$. Men da φ er en ordningsisomorfi, følger det heraf, at $c \leq a$. Altså er a en majorant for C.

Bevis for 3. Antag $a = \sup C$. Da er a en majorant for C, så ifølge punkt 1 er $\varphi(a)$ en majorant for $\varphi(C)$. Vi skal vise, at $\varphi(a)$ er den mindste majorant. Det

vises ved et modstridsbevis. Antag derfor at $s < \varphi(a)$, og at s er en majorant for $\varphi(C)$. Da φ er bijektiv, findes et $b \in A$, så $s = \varphi(b)$; og da $s = \varphi(b)$ er en majorant for $\varphi(C)$, gælder for alle $c \in C$, at $\varphi(c) \leq \varphi(b)$. Da nu φ er en ordningsisomorfi gælder det for alle $c \in C$, at $c \leq b$, hvorfor b er en majorant for C. Men da $\varphi(b) = s < \varphi(a)$, følger af sætning 354 at $b < a$. Altså er b en majorant for C, som er mindre end a, i modstrid med at a var antaget at være den mindste majorant.

Den omvendte vej overlades til læseren. ■

Sætning 358 *Sammensætning af to ordningsisomorfier giver igen en ordningsisomorfi. Hvis to ordnede mængder begge er isomorfe med en tredie, er de indbyrdes isomorfe.*

Bevis. Præcisering af sætningen og beviset overlades til læseren. ■

Øvelse 359 *Bestem en ordningsisomorfi mellem* $\mathbb{N}$ *og mængden af lige naturlige tal, med den sædvanlige ordning.*

Øvelse 360 *Lad*

$$A = \left\{ x \in \mathbb{R} \mid x = 1 - \frac{1}{n} \text{ for et } n \in \mathbb{N} \right\}, \tag{9.31}$$

og lad $B = A \cup \{1\}$*. Betragt disse mængder som ordnede mængder under den sædvanlige ordning på* $\mathbb{R}$*. Vis at* A *er ordensisomorf med* $\mathbb{N}$*, men at* B *ikke er ordensisomorf med* $\mathbb{N}$*. (vink: at* B *ikke er ordensisomorf med* $\mathbb{N}$ *vises lettest ved et modstridsbevis). Vis at* $\mathbb{N}$ *har samme kardinalitet som* B*.*

Chapter 10

Aksiomatisk beskrivelse af de reelle tal

De reelle tal spiller en fremtrædende rolle i matematikken. I skolen vænner man sig gradvist til at arbejde med dem, og man kommer til at betragte deres egenskaber som velkendte og naturlige. I dette kursus har vi da også ofte benyttet os af reelle tal som eksempel-materiale. Men hvad er egentligt de reelle tal helt præcist? Det er dette spørgsmål vi vil besvare i dette kapittel.

Vi vil opstille en række aksiomer som entydigt beskriver de reelle tal. Det er klart at aksiomerne er inspireret af de egenskaber ved de reelle tal, som vi har vænnet os til i skolen, men det er vigtigt at understrege, at logisk set er indføringen af de reelle tal i dette kapitel helt uafhængig af vores tidligere usystematiske viden om dem. Det eneste vi bygger på i dette kapitel er logikken og mængdelæren, samt nogle af de resultater vi udledte om grupper. Men intet af dette afhænger af vores tidligere omgang med de reelle tal.

Vi vil definere de reelle tal som en matematisk struktur kaldet et fuldstændigt ordnet legeme. Et legeme er en mængde L med to kompositionsregler, som vi kalder $+$ og $\cdot$, hvorom der gælder en række aksiomer (regneregler). Dette legeme skal være ordnet med en ordning $\leq$. De to kompositionsregler og ordningen spiller sammen på en særlig måde, som er afspejlet i aksiomerne. Et særligt aksiom fastslår, at det ordnede legeme har supremumsegenskaben (vi siger det er fuldstændigt).

I dette kapitel skal vi først indføre legemer, og bevise nogle simple egenskaber om dem. Dernæst skal vi indføre ordningen, og bevise sætninger, der fortæller hvordan den spiller sammen med legemsstrukturen. Til slut ser vi hvilke egenskaber fuldstændigheden tilfører strukturen.

10.1 Legemer

Definition 361 *En mængde L med to kompositionsregler $+$ (kaldet addition) og $\cdot$ (kaldet multiplikation) kaldes et* ***legeme*** *og betegnes $(L, +, \cdot)$, hvis følgende*

aksiomer er opfyldt:

L1: $+$ og $\cdot$ er kommutative, dvs. for alle $x, y \in L$ gælder:

$$x + y = y + x \quad og \quad x \cdot y = y \cdot x. \tag{10.1}$$

L2: $+$ og $\cdot$ er associative, dvs. for alle $x, y, z \in L$ gælder:

$$(x + y) + z = x + (y + z) \quad og \quad (x \cdot y) \cdot z = x \cdot (y \cdot z). \tag{10.2}$$

L.3: Der eksisterer et additivt og et multiplikativt neutralt element kaldet hhv. 0 og 1, d.v.s der eksisterer to elementer $0 \neq 1$ i L så for alle $x \in L$ gælder:

$$x + 0 = x \quad og \quad x \cdot 1 = x. \tag{10.3}$$

L4: Der eksisterer additive og multiplikative inverser, eller mere præcist: For ethvert $x \in L$ findes et element vi vil kalde $-x$ i L, hvorom det gælder, at

$$x + (-x) = 0, \tag{10.4}$$

og for ethvert $x \in L \setminus \{0\}$ findes et element vi vil kalde x^{-1} i L, hvorom det gælder, at

$$x \cdot x^{-1} = 1 \tag{10.5}$$

*L5: Den **distributive lov**: For alle $x, y, z \in L$ gælder, at*

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z). \tag{10.6}$$

Notation 362 *Som vi plejer når vi regner i $\mathbb{R}$, vil vi ofte tillade os at udelade multiplikationstegnet $\cdot$. Ligeså vil vi bruge den sædvanlige konvention om, at multiplikation udføres før addition, så vi kan udelade parenteser om produkter. Derved kan den distributive lov skrives:*

$$x(y + z) = xy + xz. \tag{10.7}$$

Sætning 363 *Hvis $(L, +, \cdot)$ er et legeme, er $(L, +)$ og$(L \setminus \{0\}, \cdot)$ abelske grupper.*

Bevis. Det overlades til læseren at checke at alle aksiomerne for abelske grupper er opfyldt. ■

Bemærkning 364 *Den distributive lov fortæller, hvordan gruppestrukturerne i $(L, +)$ og$(L \setminus \{0\}, \cdot)$ spiller sammen.*

Bemærkning 365 *Når der i det følgende tales om et legeme $(L, +, \cdot)$, er det underforstået, at de neutrale elementer og de inverse elementer betegnes som i definitionen ovenfor.*

Bemærkning 366 *Det følger af Sætning 323 og 327, at de neutrale elementer i et legeme $(L, +, \cdot)$ er entydigt bestemt, og at for et givet $x \in L$ er de additive og multiplikative inverse entydigt bestemt (sidstnævnte dog forudsat at $x \neq 0$). Det er derfor vi kan tillade os at navngive de neutrale og inverse elementer i definitionen af et legeme.*

Da $(L,+)$ og$(L\setminus\{0\},\cdot)$ er abelske grupper, gælder de sætninger vi har vist om disse i Kapitel 9. Lad os specielt fremhæve Sætning 333, 334 og 335. Oversat til notationen i dette kapitel siger disse sætninger det følgende:

Sætning 367 *Lad $(L,+,\cdot)$ være et legeme og lad x,y være vilkårlige elementer i L. Da gælder:*

$$-0=0 \quad og \quad 1^{-1}=1, \tag{10.8}$$

$$-(-x)=x \quad og \quad \left(x^{-1}\right)^{-1}=x, \tag{10.9}$$

$$-(x+y)=(-x)+(-y) \quad og \quad (xy)^{-1}=x^{-1}\cdot y^{-1}. \tag{10.10}$$

De følgende sætninger viser hvordan den additive og den multiplikative struktur på $(L,+,\cdot)$ spiller sammen. Derfor er det klart at den distributive lov skal bruges i beviserne.

Sætning 368 *Det additive neutrale element 0 i et legeme $(L,+,\cdot)$ har følgende multiplikative egenskab:*

For ethvert $x\in L$ gælder:

$$x\cdot 0=0 \tag{10.11}$$

Bevis. Sæt $y=x\cdot 0$. Da gælder:

$$y=x\cdot 0=x(0+0)=x\cdot 0+x\cdot 0=y+y. \tag{10.12}$$

Heraf fås:

$$\begin{aligned} x\cdot 0 &= y=y+0=y+(y+(-y)) && (10.13)\\ &= (y+y)+(-y)=y+(-y)=0 && (10.14)\end{aligned}$$

■

Bemærkning 369 *I ovenstående bevis og i de følgende beviser er det vigtigt, at man kun bruger de regneregler, som er specificeret i aksiomerne for et legeme, og de sætninger, man har udledt derfra. Notationen kunne forføre en til bare at regne, som man plejer i $\mathbb{R}$, men da vi nu er i gang med at opbygge de reelle tal fra grunden, er der ingen "plejer". Først når kapitlet er slut, har vi vist at man i den aksiomatiske struktur, vi indfører, kan regne som vi plejer i $\mathbb{R}$. Indtil da må alle trin i udledningerne checkes nøje, og vi må glemme alle vores tillærte regneregler for reelle tal. Check derfor nøje at vi i ovenstående udregninger kun brugte legemsaksiomerne.*

Hvis $x\in L$ er $-x$ defineret ud fra den additive gruppestruktur. Elementet 1 er derimod defineret ud fra den multiplikative struktur. De spiller dog sammen på følgende måde:

Sætning 370 *Lad $(L,+,\cdot)$ være et legeme. For ethvert $x\in L$ gælder:*

$$-x=(-1)\cdot x. \tag{10.15}$$

Bevis. Vi skal bevise at $(-1)\cdot x$ er det additivt inverse element til x. Det ses af følgende udregning:

$$x+(-1)\cdot x = 1\cdot x+(-1)\cdot x = (1+(-1))\cdot x = 0\cdot x = 0 \tag{10.16}$$

Da det inverse element er entydigt, gælder dermed, at $-x = (-1)\cdot x$. ■

Sætning 371 *Lad $(L,+,\cdot)$ være et legeme. Da gælder:*

$$(-1)(-1) = 1. \tag{10.17}$$

Bevis. Ifølge sætning (370) er $(-1)(-1) = -(-1)$, og ifølge sætning (367) er $-(-1) = 1$. ■

Sætning 372 *Lad $(L,+,\cdot)$ være et legeme. For ethvert $x,y\in L$ gælder:*

$$(-x)(-y) = xy. \tag{10.18}$$

Bevis. Beviset følger af følgende udregning:

$$\begin{aligned}
(-x)(-y) &= ((-1)x)((-1)y) = (-1)(x((-1)y)) && (10.19)\\
&= (-1)((x(-1))y) = (-1)(((-1)x)y) && (10.20)\\
&= (-1)((-1)(xy)) = ((-1)(-1))(xy) && (10.21)\\
&= 1(xy) = xy. && (10.22)
\end{aligned}$$

■

I det ovenstående bevis er hvert enkelt skridt i beviset skrevet ud, således at der i hvert lighedstegn kun er brugt et af aksiomerne. I det følgende vil beviserne ikke altid blive skrevet ud i samme detaljegrad. Da både addition og multiplikation er associative og kommutative, kan vi jo tillade os at hæve og sætte parenteser, og flytte om på leddene. Det vil vi gøre frit i det følgende.

Definition 373 *Lad $(L,+,\cdot)$ være et legeme. Vi definerer de to nye kompositionsregler $-$ (subtraktion) og $:$ (division) på L som følger:*

For ethvert $x,y\in L$ defineres

$$x-y := x+(-y). \tag{10.23}$$

For ethvert $x,y\in L$ med $y\neq 0$ defineres

$$x:y := x\cdot y^{-1} \tag{10.24}$$

Vi skriver også $x:y$ som $\frac{x}{y}$ eller x/y.

Bemærkning 374 *Faktisk er division ikke en kompositionsregel på hele L men kun på $L\setminus\{0\}$.*

Sætning 375 *Lad $(L,+,\cdot)$ være et legeme. For ethvert $x,y \in L$ gælder*

$$x-y=-(y-x) \tag{10.25}$$

og

$$xy^{-1}=\left(yx^{-1}\right)^{-1} \tag{10.26}$$

Bevis. For at vise (10.25), skal vi vise, at $x-y$ den additivt inverse til $y-x$. Det følger af følgende udregning:

$$\begin{aligned}(y-x)+(x-y) &= (y+(-x))+(x+(-y))= &(10.27)\\ y+((-x)+x)+(-y) &= y+0+(-y)=y+(-y)=0. &(10.28)\end{aligned}$$

(10.26) vises på samme måde. ■

Øvelse 376 *Lad $(L,+,\cdot)$ være et legeme. Bevis, at for $x,y \in L$ gælder*

$$x-y=0 \Leftrightarrow x=y \tag{10.29}$$

og hvis $y \neq 0$

$$x/y=1 \Leftrightarrow x=y \tag{10.30}$$

Ligesom for grupper vil vi regne to legemer for essentielt samme legeme, hvis de er isomorfe.

Definition 377 *Lad $(L,+,\cdot)$ og $(M,+,\cdot)$ være legemer. En afbildning φ:$L \longrightarrow M$ kaldes en **legemsisomorfi**, hvis den er bijektiv og for alle $x,y \in L$ opfylder:*

$$\begin{aligned}\varphi(x+y) &= \varphi(x)+\varphi(y), &(10.31)\\ \varphi(x\cdot y) &= \varphi(x)\cdot\varphi(y). &(10.32)\end{aligned}$$

*Hvis der findes en legemsisomorfi mellem to legemer, siges de at være **isomorfe**.*

Bemærkning 378 *Da en legemsisomorfi $(L,+,\cdot) \longrightarrow (M,+,\cdot)$ specielt er en gruppeisomorfi $(L,+) \longrightarrow (M,+)$, er $\varphi(0_l)=0_M$ (sætning 348), hvorfor φ også er en gruppeisomorfi $(L\setminus\{0\},\cdot) \longrightarrow (M\setminus\{0\},\cdot)$. Derfor gælder sætningerne om gruppeisomorfier også om legemsisomorfier.*

Er nu alle legemer isomorfe, eller er der flere essentielt forskellige legemer? Der er faktisk mange forskellige legemer. Ud over de reelle tal, som vi har ladet os inspirere af, er de rationale tal og de komplekse tal, med de sædvanlige regneoperationer, også eksempler på legemer.

Eksempel 379 *Talmængderne $\mathbb{Q}$, $\mathbb{R}$ og $\mathbb{C}$ med de sædvanlige regneoperationer er ikke isomorfe legemer.*

Bevis. Det anses for velkendt at $\mathbb{Q}$, $\mathbb{R}$ og $\mathbb{C}$ er legemer. Fra sætning 309-311 vides at $\mathbb{Q}$, ikke har samme kardinalitet som $\mathbb{R}$ og $\mathbb{C}$ Derfor kan de ikke være isomorfe. Vi skal derfor blot vise at $\mathbb{R}$ og $\mathbb{C}$ ikke er isomorfe.

Beviset føres ved modstrid. Antag altså at $\varphi : \mathbb{C} \longrightarrow \mathbb{R}$ er en legemsisomorfi. Betragt $\varphi(i) \in \mathbb{R}$. Ifølge (10.32, 9.23 og 9.24) gælder

$$\begin{aligned} \varphi(i) \cdot \varphi(i) &= \varphi(i \cdot i) && (10.33)\\ &= \varphi(-1) = -\varphi(1) && (10.34)\\ &= -1. && (10.35) \end{aligned}$$

Men der findes intet reelt tal som ganget med sig selv giver -1. Vi er altså nået til en modstrid, og må konkludere at der ikke findes en isomorfi mellem $\mathbb{C}$ og $\mathbb{R}$.

Præcist samme bevis kan også bruges til at vise, at $\mathbb{Q}$ og $\mathbb{C}$ ikke er isomorfe.

At $\mathbb{Q}$ og $\mathbb{R}$ ikke er isomorfe kan også vises ved en lille ændring i ovenstående bevis. Man skal blot se på billedet af $\sqrt{2}$ og udnytte at der ifølge Sætning 50 ikke findes et element i $\mathbb{Q}$, hvis kvadrat er 2. ■

Der findes også legemer med endeligt mange elementer.

Eksempel 380 *Der findes op til isomorfi præcist ét legeme med to elementer.*

Bevis. Først skal det bemærkes at det i aksiomerne er fastslået at $0 \neq 1$. Altså har et legeme mindst to elementer. For at vise at der faktisk findes et legeme med to elementer, laver vi en analyse:

Vi antager altså at $(L, +, \cdot)$ er et legeme med to elementer. Da L indeholder et additivt neutralt element 0 og et multiplikativt neutralt element 1, og da $0 \neq 1$ må $L = \{0, 1\}$. For at finde ud af om vi kan udstyre $\{0, 1\}$ som et legeme, skal vi prøve at opstille kompositionstavlerne for de to kompositionsregler.

Hvis $(\{0, 1\}, +, \cdot)$ er et legeme, må følgende være opfyldt (se sætn. 368):

$$0 + 0 = 0, \tag{10.36}$$

$$1 + 0 = 0 + 1 = 1, \tag{10.37}$$

$$0 \cdot 0 = 1 \cdot 0 = 0 \cdot 1 = 0, \tag{10.38}$$

$$1 \cdot 1 = 1. \tag{10.39}$$

Det eneste felt i kompositionstavlen, som ikke er fastlagt af disse regler er $1 + 1$. Vi ved at $1 + 1$ enten er 1 eller 0. Antag først at $1 + 1 = 1$. Da gælder:

$$1 = 1 + 0 = 1 + (1 + (-1)) = (1 + 1) + (-1) = 1 + (-1) = 0, \tag{10.40}$$

men det er i modstrid med at $1 \neq 0$. Vi kan altså slutte at hvis $(\{0, 1\}, +, \cdot)$ skal være et legeme er det kun muligt hvis $1 + 1 = 0$. De eneste mulige kompositionstavler ser altså således ud:

$+$	0	1
0	0	1
1	1	0

(10.41)

$$\begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array} \tag{10.42}$$

Vores analyse har nu vist, at hvis der overhovedet findes et legeme med to elementer, må kompositionstavlerne for kompositionsreglerne se ud som ovenfor. Vi har altså vist entydigheden af kompositionstavlen, hvad der netop betyder at to legemer med to elementer må være isomorfe (overvej).

For at vise eksistensen skal vi vise at alle regneregler er opfyldt når addition og multiplikation defineres som angivet i kompositionstavlerne. Det er en endelig opgave, men lidt uoverskuelig, da vi skal checke mange ting for alle $x, y(, z) \in \{0, 1\}$.

Vi kan lette efterprøvningen ved at afsløre at vi allerede har mødt en mængde med to elementer og to kompositionsregler, som har de angivne kompositionstavler, nemlig $\mathbb{Z}/2$. Vi skal bare skrive [0] i stedet for 0 og [1] i stedet for 1 i kompositionstavlerne ovenfor, så får vi kompositionstavlerne for $\mathbb{Z}/2$. (check dette). Vi har allerede overvejet at både addition og multiplikation er kommutative og additive, og vi har bevist at $(\mathbb{Z}/2, +)$ er en gruppe. Vi mangler da bare at vise at den distributive lov er opfyldt, og at [1] har en multiplikativ invers. Den distributive lov nedarves fra $\mathbb{Z}$ (overvej), og det ses af kompositionstavlen, at [1] er sin egen inverse. Dermed har vi bevist at $(\mathbb{Z}/2, +, \cdot)$ (eller hvad der er det samme $(\{0, 1\}, +, \cdot)$ med de ovenstående kompositionstavler) er et legeme. ■

Bemærkning 381 *Da isomorfier er bijektive, ved vi fra Sætning 307, at endelige isomorfe legemer har samme antal elementer.*

Det kan vises at, hvis n er en potens af et primtal, så findes der op til isomorfi præcist ét endeligt legeme med n elementer. Der findes derimod ingen endelige legemer, hvis elementantal ikke er en potens af et primtal.

$(\mathbb{Z}/p, +, \cdot)$ er et legeme, hvis og kun hvis p er et primtal, og det er altså det eneste legeme med p elementer.

Vi skal ikke bevise disse resultater. Den interesserede læser kan læse mere på adressen: http://en.wikipedia.org/wiki/Finite_field.

Grunden til at $(\mathbb{Z}/n, +, \cdot)$ ikke er et legeme når n er et sammensat tal, er at der findes elementer forskellig fra 0, som ikke har en multiplikativ invers. For $n = 4$ beviste du det i øvelse (352).

Øvelse 382 *Bevis at $(\mathbb{Z}/3, +, \cdot)$ er et legeme.*

10.2 Ordnede legemer

Da der er så mange væsensforskellige legemer, er det klart at legemsaksiomerne ikke i sig selv er nok til entydigt at definere de reelle tal. Vi vil derfor indskrænke mulighederne, ved at tilføje en ordensstruktur på legemet. Den mest oplagte fremgangsmåde ville være at kræve, at der eksisterer en ordning $\leq$ på legemet,

som på passende vis spiller sammen med legemsstrukturen, og derfra definere de positive elementer som de elementer, der er større end 0. Det viser sig dog at være mere hensigtsmæssigt at gå omvendt til værks. Vi vil i stedet kræve at der er en delmængde af legemet, som vi vil kalde de positive elementer, og opstille aksiomer for denne mængde. Derefter vil vi så definere en ordning, ved at sige at $x \leq y$ hvis $y - x$ er positiv eller 0.

Definition 383 *Et legeme $(L, +, \cdot)$ kaldes et **ordnet legeme,** hvis der eksisterer en delmængde $L_+ \subseteq L$, hvorom der gælder følgende aksiomer:*

O1: L_+ er lukket under addition og multiplikation, dvs. for alle $x, y \in L_+$ gælder:

$$(x + y) \in L_+ \quad og \quad x \cdot y \in L_+. \tag{10.43}$$

O2: For alle $x \in L$ gælder præcist et af følgende udsagn:

$$x \in L_+, \tag{10.44}$$
$$x = 0, \tag{10.45}$$
$$-x \in L_+. \tag{10.46}$$

Elementerne i L_+ kaldes de positive elementer i L.
De negative elementer er pr. definition elementerne i mængden

$$L_- := \complement \left(L_+ \cup \{0\} \right). \tag{10.47}$$

I det følgende betegner L_+ og L_- de positive, henholdsvist negative elementer i et ordnet legeme $(L, +, \cdot)$.

Øvelse 384 *Bevis ved induktion, at summen af et vilkårligt endeligt antal positive elementer i et ordnet legeme er positiv.*

Sætning 385 *For ethvert element x i et ordnet legeme $(L, +, \cdot)$ gælder præcist et af følgende udsagn:*

$$x \in L_+ \tag{10.48}$$
$$x = 0 \tag{10.49}$$
$$x \in L_-. \tag{10.50}$$

Bevis. Først viser vi at højst et af de tre udsagn kan være sandt, altså at to af dem ikke kan være sande samtidigt:

- $x \in L_+$ og $x = 0$ kan ikke begge være sande i følge aksiom O2.
- $x \in L_-$ og $x = 0$ kan ikke begge være sande, da $0 \notin L_-$ (iflg. definitionen af L_-).
- $x \in L_-$ og $x \in L_+$ kan ikke begge være sande, da $L_- := \complement \left(L_+ \cup \{0\} \right)$.

Dernæst viser vi at mindst et af udsagnene er sandt. Til det formål er det nok at vise, at hvis to af udsagnene er falske, så er det sidste udsagn sandt. Antag derfor, at $x \in L_+$ og $x = 0$ begge er falske udsagn, altså at $x \in \complement L_+$ og $x \in \complement\{0\}$. Ifølge De Morgans love (Sætning 148) gælder altså, at

$$x \in \left(\complement L_+\right) \cap \left(\complement\{0\}\right) = \complement\left(L_+ \cup \{0\}\right) = L_-. \tag{10.51}$$

Altså er det sidste udsagn sandt. ■

Sætning 386 *For ethvert element x i et ordnet legeme $(L, +, \cdot)$ gælder*

$$x \in L_+ \Leftrightarrow -x \in L_-. \tag{10.52}$$

Bevis. "⇒": Antag at $x \in L_+$. Ifølge sætning (385) ved vi, at $-x \in L_+$ eller $-x = 0$ eller $-x \in L_-$. Vi vil ved modstrid vise, at de to første alternativer ikke er mulige:

1. Antag at $-x \in L_+$. Ifølge aksiom O1 vil da $0 = x + (-x) \in L_+$, men det strider mod aksiom O2, som jo bl.a. siger at $0 \notin L_+$.
2. Antag at $-x = 0$. Da er $0 = x + (-x) = x + 0 = x$, som jo atter strider mod aksiom O2, idet vi har antaget at $x \in L_+$.

Da både $-x \in L_+$ og $-x = 0$ er falske udsagn, må $-x \in L_-$ være sand.

"⇐": Antag at $-x \in L_-$. Ifølge sætning (385) gælder da hverken $-x \in L_+$ eller $-x = 0$, så ifølge aksiom O2 må $-(-x) \in L_+$ men ifølge sætning 367) er $-(-x) = x$, hvorfor $x \in L_+$. ■

Den næste sætning viser, at der i et vilkårligt ordnet legeme gælder de sædvanlige fortegns-regneregler:

Sætning 387 *Lad x, y være elementer i et ordnet legeme $(L, +, \cdot)$. Da gælder:*

1. $1 \in L_+$.
2. $x, y \in L_- \Rightarrow x \cdot y \in L_+$.
3. $x, y \in L_- \Rightarrow (x + y) \in L_-$.
4. $(x \in L_+ \wedge y \in L_-) \Rightarrow x \cdot y \in L_-$.

Bevis. *Bevis for 1*: Vi giver et modstridsbevis. Antag altså at $1 \notin L_+$. Da endvidere $1 \neq 0$, følger af aksiom O2 at $-1 \in L_+$. men så gælder ifølge (371) og aksiom O1 at

$$1 = (-1) \cdot (-1) \in L_+. \tag{10.53}$$

Dette strider mod antagelsen om at $1 \notin L_+$, hvorfor vi slutter at $1 \in L_+$.

Bevis for 2: Antag at $x, y \in L_-$. Ifølge sætning (386) gælder da, at $(-x), (-y) \in L_+$. Af sætning (372) og aksiom O1 fås så:

$$x \cdot y = (-x) \cdot (-y) \in L_+ \tag{10.54}$$

Bevis for 3:Lad $x, y \in L_-$. Ifølge sætning (386) gælder da, at $(-x), (-y) \in L_+$, så ifølge aksiom O1 gælder $((-x) + (-y)) \in L_+$. Det følger da af sætning (386), at $-((-x) + (-y)) \in L_-$. Men $-((-x) + (-y)) = x + y$, da de begge er den inverse til $(-x) + (-y)$ (overvej). Altså kan vi konkludere at $(x + y) \in L_-$.

Bevis for 4: Antag at $x \in L_+$ og $y \in L_-$. Fra sætning (370) og associativiteten og kommutativiteten kan vi slutte:

$$\begin{aligned} x \cdot y &= x \cdot (-(-y)) = x \cdot ((-1)(-y)) = (x \cdot (-1))(-y) && (10.55) \\ &= ((-1) \cdot x)(-y) = -(x(-y)). && (10.56) \end{aligned}$$

Iflg. sætning (367) er $y = -(-y)$, og da $-(-y) = y \in L_-$, fås fra Sætning (386), at $-y \in L_+$. Derfor ligger både x og $-y$ i L_+, så ifølge aksiom O1 gælder, at $x(-y) \in L_+$. Fornyet brug af sætning (386) giver da, at $-(x(-y)) \in L_-$, og da $x \cdot y = -(x(-y))$, glæder altså, at $x \cdot y \in L_-$. ■

Sætning 388 *Lad x være et element i et ordnet legeme $(L, +, \cdot)$. Da gælder:*

1. $x \in L_+ \Rightarrow x^{-1} \in L_+$.
2. $x \in L_- \Rightarrow x^{-1} \in L_-$.

Bevis. *Bevis for 1*: Antag at $x \in L_+$. Da $x \cdot x^{-1} = 1 \in L_+$, følger af sætning (387.4) at $x^{-1} \notin L_-$. Ifølge sætning (385) er da enten $x^{-1} = 0$ eller $x^{-1} \in L_+$. Men da $x \cdot 0 = 0$ (sætning (368)) er første alternativ udelukket. Altså gælder $x^{-1} \in L_+$.

Bevis for 2.: Overlades til læseren. ■

Definition 389 *Lad $(L, +, \cdot)$ være et ordnet legeme med positive elementer L_+. Vi definerer da en relation $\leq$ på L, idet vi siger, at for $x, y \in L$ er $x \leq y$, hvis $y - x \in L_+ \cup \{0\}$.*

Sætning 390 *Relationen $\leq$ defineret i definition (389) er en total ordningsrelation på L.*

Bevis. Vi skal vise at $\leq$ er en ordningsrelation på L, og at den er total.

- *Refleksivitet*: Lad $x \in L$. Da $x - x = x + (-x) = 0 \in L_+ \cup \{0\}$, er $x \leq x$. Altså er $\leq$ refleksiv.
- *Antisymmetri*: Lad $x, y \in L$, og antag at $x \leq y$ og $y \leq x$. Det betyder pr. definition, at

  $$(y - x) \in L_+ \cup \{0\} \quad (10.57)$$

  og

  $$(x - y) \in L_+ \cup \{0\}. \quad (10.58)$$

  Ifølge sætning (375) er

  $$x - y = -(y - x). \quad (10.59)$$

Da $(y-x) \in L_+ \cup \{0\}$ følger af aksiom O2, at $x-y = -(y-x) \notin L_+$, men da vi har antaget at $(x-y) \in L_+ \cup \{0\}$, må $x-y=0$. Men så gælder:

$$y = y+0 = y+(x-y) = x+y+(-y) = x+0 = x. \tag{10.60}$$

Altså er $\leq$ antisymmetrisk.

- *Transitivitet*: Lad $x, y, z \in L$, og antag at $x \leq y$ og $y \leq z$, altså at

$$(y-x) \in L_+ \cup \{0\} \tag{10.61}$$

og

$$(z-y) \in L_+ \cup \{0\}. \tag{10.62}$$

Da gælder ifølge aksiom O1 at[1]

$$(z-x) = ((z-y)+(y-x)) \in L_+ \cup \{0\}, \tag{10.63}$$

hvoraf ses, at $x \leq z$. Dermed er påvist, at $\leq$ er transitiv.

- *Totalitet*: Lad x, y være vilkårlige elementer i L. Vi skal vise at $x \leq y$ eller $y \leq x$. Ifølge aksiom O2 og sætning (375) ved vi, at et af følgende udsagn holder:

$$\begin{aligned} y-x &\in L_+, && (10.64)\\ y-x &= 0, && (10.65)\\ x-y &= -(y-x) \in L_+. && (10.66) \end{aligned}$$

Hvis et af de to første udsagn er sande, gælder $x \leq y$ og hvis et af de to sidste udsagn er sande, gælder $y \leq x$. Altså er enten $x \leq y$ eller $y \leq x$.

■

Sætning 391 *Lad $\leq$ være ordningen på et ordnet legeme $(L, +, \cdot)$. Da gælder*

$$x < y \Leftrightarrow y - x \in L_+ \tag{10.67}$$

og

$$0 < x \Leftrightarrow x \in L_+. \tag{10.68}$$

Bevis. Overlades til læseren. (Husk at relationen $<$ er defineret i Definition 218) ■

Vi vil nu se, hvordan ordningen på et ordnet legeme spiller sammen med legemsstrukturen (aritmetikken).

Sætning 392 *Lad $(L, +, \cdot)$ være et ordnet legeme, og $a, b, c, d \in L$. Da gælder:*

[1] Her angives ikke alle mellemregninger. Læseren kan selv fylde dem på.

1. $(a < b) \wedge (c \leq d) \Rightarrow (a + c) < (b + d)$.
2. $(a < b) \wedge 0 < c \Rightarrow ac < bc$.
3. $(0 < a < b) \wedge (0 < c \leq d) \Rightarrow ac < bd$

Bevis. *Bevis for 1*: Antag at $(a < b)$ og $(c \leq d)$, dvs. at

$$(b - a) \in L_+ \quad \text{og} \quad (d - c) \in L_+ \cup \{0\}. \tag{10.69}$$

Ifølge aksiom O1 gælder da (overvej)

$$(b + d) - (a + c) = (b - a) + (d - c) \in L_+, \tag{10.70}$$

hvorfor $(a + c) < (b + d)$.

Bevis for 2: Antag at $(a < b)$ og $0 < c$, dvs. at

$$(b - a) \in L_+ \quad \text{og} \quad c \in L_+ \tag{10.71}$$

Da følger det af aksiom O1 og den distributive lov, at

$$bc - ac = (b - a)\, c \in L_+, \tag{10.72}$$

hvorfor $ac < bc$.

Bevis for 3: Antag at $(0 < a < b) \wedge (0 < c \leq d)$. Ifølge punkt 2, er

$$ac < bc \tag{10.73}$$

og

$$bc \leq bd \tag{10.74}$$

(overvej). Af transitiviteten fås da, at

$$ac < bd \tag{10.75}$$

(overvej). ■

Sætning 393 *Lad $(L, +, \cdot)$ være et ordnet legeme, og $a, b \in L$. Da gælder:*

$$0 < a \leq b \Rightarrow b^{-1} \leq a^{-1}. \tag{10.76}$$

Bevis. Overlades til læseren. ■

Man kan vise mange andre sætninger om uligheder. De ser ud som reglerne i $\mathbb{R}$.

Definition 394 *Lad $(L, +, \cdot)$ være et ordnet legeme, og $a \in L$. Den numeriske værdi (eller absolutværdien) $|a|$ defineres som*

$$|a| = \begin{cases} a & \text{hvis } a \in L_+ \cup \{0\} \\ -a & \text{hvis } a \in L_- \end{cases} \tag{10.77}$$

Man kan vise de fra $\mathbb{R}$ kendte sætninger om numeriske værdier, og deres beviser er helt magen til beviserne i $\mathbb{R}$. Man indfører endvidere de sædvanlige betegnelser for intervaller (se Notation 99). Da gælder følgende sætning:

Sætning 395 *Lad $(L,+,\cdot)$ være et ordnet legeme, og $a,x \in L$ og $\epsilon \in L_+$. Da gælder:*

$$|x-a| < \epsilon \Leftrightarrow x \in]a-\epsilon, a+\epsilon[\tag{10.78}$$

Bevis. Overlades til læseren. ■

Definition 396 *En legemsisomorfi mellem to ordnede legemer kaldes ordensbevarende, hvis den er en ordningsisomorfi. Vi siger at to ordnede legemer er* ***isomorfe****, hvis der eksisterer en ordensbevarende legemsisomorfi mellem dem.*

Sætning 397 *En legemsisomorfi $\varphi : (L,+,\cdot) \longrightarrow (M,+,\cdot)$ er ordensbevarende, hvis og kun hvis $\varphi(L_+) = M_+$.*

Bevis. Lad $\varphi : (L,+,\cdot) \longrightarrow (M,+,\cdot)$ være en legemsisomorfi.

- Antag først at φ er en ordningsisomorfi, altså at for $x,y \in L$ gælder at

$$x \leq y \Leftrightarrow \varphi(x) \leq \varphi(y), \tag{10.79}$$

eller ækvivalent hermed, at

$$x < y \Leftrightarrow \varphi(x) < \varphi(y). \tag{10.80}$$

(sætning 354). Vi skal vise at $\varphi(L_+) = M_+$.

Først viser vi at $\varphi(L_+) \subseteq M_+$. Lad altså $y \in \varphi(L_+)$. Da eksisterer et $x \in L_+$, hvorom det gælder at $\varphi(x) = y$. Da $x \in L_+$ ved vi fra (10.68) at $0 < x$, så af (10.80) slutter vi at

$$0 = \varphi(0) < \varphi(x) = y. \tag{10.81}$$

Fra (10.68) fås da, at $y \in M_+$.

Dernæst viser vi, at $\varphi(L_+) \supseteq M_+$. Lad altså $y \in M_+$. Da φ er bijektiv, eksisterer et $x \in L$, hvorom det gælder, at $\varphi(x) = y$. Da $y \in M_+$, følger af (10.68), at

$$\varphi(0) = 0 < y = \varphi(x), \tag{10.82}$$

hvorfra vi ved brug af (10.80) kan slutte, at $0 < x$, altså at $x \in L_+$. Men det betyder at $y = \varphi(x) \in \varphi(L_+)$.

- Antag dernæst at $\varphi(L_+) = M_+$, og lad $x,y \in L$. Da gælder

$$\begin{aligned} x \quad &< \quad y \Leftrightarrow (y-x) \in L_+ \Leftrightarrow \varphi(y-x) \in M_+ && (10.83)\\ &\Leftrightarrow \quad \varphi(y+(-x)) \in M_+ \Leftrightarrow (\varphi(y)+\varphi(-x)) \in M_+ && (10.84)\\ &\Leftrightarrow \quad (\varphi(y)-\varphi(x)) \in M_+ \Leftrightarrow \varphi(x) < \varphi(y). && (10.85) \end{aligned}$$

Afbildningen φ er altså en ordningsisomorfi.

■

Eksempel 398 $\mathbb{R}$ *og* $\mathbb{Q}$ *er begge eksempler på ordnede legemer, hvis* $\mathbb{R}_+$ *og* $\mathbb{Q}_+$ *tillægges den sædvanlige mening. Som vi så i (379) er disse legemer ikke isomorfe som legemer. De er derfor heller ikke isomorfe som ordnede legemer.*

Derimod er $\mathbb{C}$ *og de endelige legemer (f.eks.* $\mathbb{Z}/p$*) ikke ordnede legemer.*

Bevis. Vi vil vise at $\mathbb{C}$ ikke er et ordnet legeme. Det gøres ved modstrid: Hvis $\mathbb{C}$ var et ordnet legeme ville i ifølge 385 enten tilhøre $\mathbb{C}_+$ eller $\mathbb{C}_-$ eller være lig med 0. Sidstnævnte mulighed er forkert, og da $i \cdot i = -1 \in \mathbb{C}_-$ strider de to første muligheder mod aksiom O1 og 387(2). Altså er $\mathbb{C}$ ikke et ordnet legeme.

Nedenfor vil vi se at ethvert ordnet legeme er uendeligt. Dermed er det klart at endelige legemer ikke er ordnede legemer. ■

Vi vil nu vise at ethvert ordnet legeme $(L, +, \cdot)$ indeholder kopier af $\mathbb{N}$, $\mathbb{Z}$, og $\mathbb{Q}$ i den forstand at der findes delmængder af L, som udstyret med kompositionsreglerne og ordningen på L er isomorfe[2] med $\mathbb{N}$, $\mathbb{Z}$, og $\mathbb{Q}$. Dette medfører specielt at ethvert ordnet legeme er uendeligt.

Lidt løst sagt er ideen den at vi først konstruerer en mængde i L som er isomorf med $\mathbb{N}$. Det gør vi ved at se på mængden N bestående af elementerne:1_L, $1_L + 1_L$, $1_L + 1_L + 1_L, ...$ Her har vi brugt betegnelsen 1_L for det multiplikative neutrale element i L for at skelne det fra det naturlige tal 1. Vi viser at alle disse elementer er forskellige. Den afbildning ϕ, som afbilder $n \in \mathbb{N}$ over i elementet $(1_L + 1_L + 1_L, ... + 1_L)$ (hvor der er n addender) kan nu vises at være en isomorfi mellem $\mathbb{N}$ og N.

Mængden $Z \subseteq L$ af differenser mellem elementer i N er da isomorf med $\mathbb{Z}$, og mængden $Q \subseteq L$ af kvotienter mellem elementer i Z er isomorf med $\mathbb{Q}$.

Ved konstruktionen af mængden N vil vi gå lidt mere formelt til værks, idet vi vil bruge vores viden om rekursion og induktion.

Definition 399 *Lad* $(L, +, \cdot)$ *være et ordnet legeme. Afbildningen* $\phi : \mathbb{N} \longrightarrow L$ *defineres som den afbildning, som opfylder*

$$\phi(1) = 1_L \tag{10.86}$$

$$\phi(n+1) = \phi(n) + 1_L \quad \textit{for alle } n \in \mathbb{N}. \tag{10.87}$$

Bemærkning 400 *At der her er defineret en entydigt bestemt afbildning, følger af rekursionssætningen (Sætning 90) idet vi går ud fra afbildningen* $f : L \longrightarrow L$ *givet ved* $f(x) = x + 1_L$*.*

Sætning 401 *Afbildningen* $\phi : \mathbb{N} \longrightarrow L$ *opfylder, at*

$$\phi(m+n) = \phi(m) + \phi(n) \quad \textit{for alle } m, n \in \mathbb{N}. \tag{10.88}$$

[2] $\mathbb{Q}$ er isomorf med et dellegeme af L. $\mathbb{N}$ er ikke et legemer, men det er isomorft med en delmængde N af L i den forstand at der findes en ordensisomorfi $\phi : \mathbb{N} \longrightarrow N$, som også opfylder (10.31 og 10.32). Det samme gælder $\mathbb{Z}$.

Bevis. Vi lader $m \in \mathbb{N}$ være et vilkårligt fastholdt naturligt tal, og fører så beviset ved induktion efter n.

Induktionsstarten: Ifølge (10.86 og 10.87) gælder at

$$\phi(m+1) = \phi(m) + 1_L = \phi(m) + \phi(1). \tag{10.89}$$

Dermed er (10.88) sand for $n = 1$.

Induktionsskridtet: Antag nu at (10.88) er sand for en bestemt værdi n_0 af n altså at

$$\phi(m+n_0) = \phi(m) + \phi(n_0). \tag{10.90}$$

Vi skal da vise at (10.88) er sand for $n = n_0+1$.Det følger af følgende udregning, hvor det midterste lighedstegn er et resultat af induktionsantagelsen (10.90), og de to andre lighedstegn er et resultat af (10.87):

$$\phi(m+n_0+1) = \phi(m+n_0)+1_L = \phi(m)+\phi(n_0)+1_L = \phi(m)+\phi(n_0+1). \tag{10.91}$$

■

Sætning 402 *Lad $(L,+,\cdot)$ være et ordnet legeme, og lad $\phi : \mathbb{N} \longrightarrow L$ være afbildningen defineret i definition 399. Da er $\phi(n) \in L_+$ for alle $n \in \mathbb{N}$.*

Bevis. Induktion efter n. ■

Sætning 403 *Afbildningen $\phi : \mathbb{N} \longrightarrow L$ opfylder, at*

$$m < n \Rightarrow \phi(m) < \phi(n). \tag{10.92}$$

Bevis. Antag at $m < n$. Da er $n - m \in \mathbb{N}$, og $n = m + (n - m)$. Af sætning 401 følger at

$$\phi(n) = \phi(m + (n - m)) = \phi(m) + \phi(n - m), \tag{10.93}$$

så

$$\phi(n) - \phi(m) = \phi(n - m). \tag{10.94}$$

Da nu $\phi(n - m) \in L_+$ ifølge sætning 402, følger at $\phi(m) < \phi(n)$. ■

Korollar 404 *Afbildningen $\phi : \mathbb{N} \longrightarrow \phi(\mathbb{N})$ er bijektiv og en ordningsisomorfi.*

Bevis. Overlades til læseren. ■

Sætning 405 *Afbildningen $\phi : \mathbb{N} \longrightarrow L$ opfylder, at*

$$\phi(m \cdot n) = \phi(m) \cdot \phi(n) \quad \textit{for alle } m, n \in \mathbb{N}. \tag{10.95}$$

Bevis. Beviset kan føres ved at fastholde m og lave induktion efter n. Det overlades til læseren. ■

Nu har vi altså konstrueret en uendelig delmængde $N := \phi(\mathbb{N})$ af L, som udstyret med strukturen i L er isomorf med $\mathbb{N}$.

Definition 406 *Delmængden Z af L defineres ved:*

$$Z := \{\phi(n) - \phi(m) \mid m, n \in \mathbb{N}\}. \tag{10.96}$$

Definer så afbildningen $\phi_1 : \mathbb{Z} \longrightarrow Z$ ved

$$\phi(n - m) = \phi(n) - \phi(m) \text{ for alle } m, n \in \mathbb{N}. \tag{10.97}$$

Bemærkning 407 *Det er ikke på forhånd klart at ϕ_1 er veldefineret ved (10.97). Ethvert tal i $\mathbb{Z}$ kan jo skrives på formen $(n - m)$ på uendeligt mange måder. Spørgsmålet er da, om $\phi(n) - \phi(m)$ kan antage flere forskellige værdier. Du kan selv bevise at det kan den ikke.*

Sætning 408 *$\phi_1 : \mathbb{Z} \longrightarrow Z$ er en ordningsisomorfi, som for alle $x, y \in \mathbb{Z}$ opfylder:*

$$\phi_1(x + y) = \phi_1(x) + \phi_1(y), \tag{10.98}$$
$$\phi_1(x \cdot y) = \phi_1(x) \cdot \phi_1(y). \tag{10.99}$$

Beviset udelades.

Definition 409 *Delmængden Q af L defineres ved:*

$$Q := \left\{ \frac{x}{y} \mid x, y \in Z \wedge y \neq 0 \right\}. \tag{10.100}$$

Definer så afbildningen $\phi_2 : \mathbb{Q} \longrightarrow Q$ ved

$$\phi_2(\frac{x}{y}) = \frac{\phi_1(x)}{\phi_1(y)}. \tag{10.101}$$

Igen skal det checkes at ϕ_2 er veldefineret.

Sætning 410 *Afbildningen $\phi_2 : \mathbb{Q} \longrightarrow Q$ er en ordensbevarende legemsisomorfi.*

Bevis. Beviset udelades. ■

Vi har hermed vist at der sidder en kopi af de rationale tal inde i ethvert ordnet legeme. Man kan altså sige at $\mathbb{Q}$ er det mindste ordnede legeme.

Sætning 411 *Ethvert ordnet legeme er uendeligt.*

Bevis. Lad L være et ordnet legeme. Da $\mathbb{N}$ er uendelig, og N er isomorf med $\mathbb{N}$ er N uendelig (Sætning 307). Da $N \subseteq L$ er L også uendelig (Sætning 308). ■

10.3 Fuldstændigt ordnede legemer. De reelle tal

Da der altså er flere ordnede legemer end de reelle tal, er aksiomerne for ordnede legemer ikke tilstrækkelige til at fastlægge disse entydigt. Sagt med andre ord, aksiomerne for et ordnet legeme er ikke tilstrækkelige til at kunne udlede nogle af de sætninger, vi ønsker skal gælde om de reelle tal. for eksempel kan vi ikke vise, at ethvert positivt element i et ordnet legeme har en kvadratrod. Denne sætning gælder jo ikke i det ordnede legeme $\mathbb{Q}$ (se Sætning 50).

Vi skal derfor tilføje endnu et aksiom, nemlig supremumsegenskaben:

Definition 412 *Et ordnet legeme kaldes et* ***fuldstændigt ordnet legeme****, hvis det har supremumsegenskaben, altså hvis enhver ikke tom opadtil begrænset delmængde har et supremum.*

Sætning 413 *Alle fuldstændige ordnede legemer er isomorfe.*

Vi skal ikke bevise denne sætning. Den tillader os at definere de reelle tal:

Definition 414 ***De reelle tal*** $\mathbb{R}$ *er det op til isomorfi entydigt bestemte fuldstændigt ordnede legeme.*

Ud fra aksiomerne for et fuldstændigt ordnet legeme, kan vi nu udlede alle de kendte sætninger om reelle tal. Vi skal ikke gennemføre denne øvelse i detaljer, men indskrænke os til at bevise nogle få centrale sætninger.

Korollar 415 *Enhver ikke tom nedadtil begrænset delmængde af* $\mathbb{R}$ *har et infimum.*

Bevis. Følger af Sætning 251. ■

Sætning 416 *Lad A være en ikke-tom opadtil begrænset delmængde af $\mathbb{R}$. Lad a være en majorant for A. Da er følgende betingelser ækvivalente:*

1. $a = \sup A$.
2. $\forall \epsilon \in \mathbb{R}_+ \exists x \in A : |x - a| < \epsilon$.
3. $\forall \epsilon \in \mathbb{R}_+ \exists x \in A : x \in]a - \epsilon, a + \epsilon[$

Bevis. At betingelserne er ækvivalente betyder, at hver af dem medfører enhver af de andre. Man kan vise dette ved (eksempelvis) at vise implikationerne: (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (1).

Bevis for (1) $\Rightarrow$ (2): Antag at $a = \sup A$, og lad $\epsilon \in \mathbb{R}_+$ være vilkårlig. Da er $a - \epsilon < a$, thi $a - (a - \epsilon) = \epsilon \in \mathbb{R}_+$. Da $a = \sup A$, er $a - \epsilon$ altså ikke en majorant for A. Men det betyder at der findes et $x \in A$, hvorom det gælder, at $a - \epsilon < x$. Ved brug af sætning 392 fås heraf, at $a - x < \epsilon$. Da endvidere a er en majorant for A, er $x \leq a$, så $a - x \geq 0$. Altså er $|x - a| = a - x < \epsilon$.

Bevis for (2) $\Rightarrow$ (3): Det følger af sætning 395.

Bevis for (3) $\Rightarrow$ (1). På grund af sætningens forudsætninger ved vi at a er en majorant for A. Vi skal vise at det er den mindste majorant. Vi fører beviset ved modstrid. Antag derfor, at A har en majorant b, som er mindre end a: $b < a$. Da er $\epsilon = a - b \in \mathbb{R}_+$. Ifølge (3) findes der da et $x \in A$, så $x \in]a-(a-b)\,, a+(a-b)\,[=]b, a+(a-b)\,[$. Altså er $x > b$, i modstrid med at b er en majorant. ■

Vi antager at vi har indlejret $\mathbb{N}$, $\mathbb{Z}$ og $\mathbb{Q}$ i de relle tal, som beskrevet ovenfor.

Sætning 417 ***(Den arkimediske egenskab)***[3]. *For ethvert reelt tal x findes der et naturligt tal n, så $x < n$.*

Bevis. Beviset føres ved modstrid. Antag altså at der findes et $x \in \mathbb{R}$, så $x < n$ er falsk for alle $n \in \mathbb{N}$. I så fald er $n \leq x$ for alle $n \in \mathbb{N}$, så $\mathbb{N}$ er opadtil begrænset af x. Ifølge supremumsegenskaben findes da et supremum y for $\mathbb{N}$. Da y er det mindste overtal for $\mathbb{N}$, og $y - 1 < y$ er $y - 1$ ikke et overtal. Der findes altså et naturligt tal m så $y - 1 < m$. Ifølge sætning (392) er da $y < m + 1$, men da $m + 1 \in \mathbb{N}$ strider det mod at y er et overtal. ■

Sætning 418 *For ethvert positivt reelt tal x findes et naturligt tal n, så $1/n < x$.*

Bevis. Antag $x \in \mathbb{R}_+$. Ifølge sætning (388) er da $x^{-1} \in \mathbb{R}_+$. Altså findes et $n \in \mathbb{N}$ så $x^{-1} < n$. Ifølge sætning (393) er da $x = \left(x^{-1}\right)^{-1} > 1/n$ (overvej). ■

Sætning 419 *Ethvert positivt reelt tal har en positiv kvadratrod. Sagt med andre ord:*

$$\forall a \in \mathbb{R}_+ \exists b \in \mathbb{R}_+ : b^2 = a. \tag{10.102}$$

Bevis. Betragt mængden

$$A = \left\{ x \in \mathbb{R} \mid x^2 < a \right\}. \tag{10.103}$$

Da $0 \in A$ er A ikke tom. Vi vil vise at A også er opadtil begrænset, mere præcist at $a + 1$ er et overtal. Vi skal altså vise at

$$x^2 < a \Rightarrow x \leq a + 1. \tag{10.104}$$

Det gør vi ved kontraposition. Antag altså at $x > a + 1$. Da følger af regnereglerne i sætning (392) at

$$x^2 \geq (a+1)x \geq (a+1) \cdot 1 = a + 1 \geq a. \tag{10.105}$$

hvilket netop er negationen af venstresiden i (10.104).

Ifølge supremumsegenskaben har A altså et supremum b. Det kan vises, at $b > 0$ og at

$$b^2 = a, \tag{10.106}$$

[3] Efter den græske matematiker Arkimedes (287-212 f.Kr.).

med andre ord, at b er en positiv kvadratrod af a. Beviset for at b er positiv overlades til læseren. Beviset for at $b^2 = a$ forløber helt parallelt med beviset i eksempel (245) idet vi viser at både.$b^2 < a$ og $b^2 > a$ fører til modstrid, hvoraf trikotomiloven giver, at $b^2 = a$.

Da vi i Sætning 50 har bevist at 2 ikke har en kvadratrod i $\mathbb{Q}$, ligger der altså irrationale tal i $\mathbb{R}$. ■

Beviserne for de følgende sætninger overlader vi til læseren. Se opgave 6 og 7.

Sætning 420 *For alle $x \in \mathbb{R}$, findes et $n \in \mathbb{Z}$, hvorom det gælder, at*

$$n \leq x < n + 1 \tag{10.107}$$

Sætning 421 *For ethvert $x \in \mathbb{R}$ og ethvert $\epsilon \in \mathbb{R}_+$ eksisterer et rationalt tal q så*

$$|x - q| < \epsilon. \tag{10.108}$$

*Man siger at $\mathbb{Q}$ ligger **tæt** i $\mathbb{R}$.*

Sætning 422 *Mellem to vilkårlige forskellige reelle tal ligger der et rationalt tal. Med andre ord:*

$$\forall x, y \in \mathbb{R} : (x < y) \Rightarrow (\exists q \in \mathbb{Q} : x < q < y) \tag{10.109}$$

Sætning 423 *Mellem to vilkårlige forskellige reelle tal ligger der et irrationalt tal.*

10.4 Opgaver

1. Lad $(L, +, \cdot)$ være et legeme, og $a, b, c, d \in L$ med $c, d \neq 0$. Vis at

$$\frac{ab}{cb} = \frac{a}{c}, \tag{10.110}$$

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, \tag{10.111}$$

$$\frac{1}{c/d} = \frac{d}{c}. \tag{10.112}$$

2. Lad $\varphi : (L, +, \cdot) \longrightarrow (M, +, \cdot)$ være en isomorfi mellem to legemer og $x, y, z \in L$ med $z \neq 0$. Vis at der gælder:

$$\varphi(x - y) = \varphi(x) - \varphi(y), \tag{10.113}$$

$$\varphi(x/y) = \varphi(x) / \varphi(y). \tag{10.114}$$

3. Overvej, hvorfor man ikke kan definere 0^{-1}.

4. Formuler flere regneregler for uligheder i stil med reglerne i sætning (392) og bevis dem.

5. Bevis sætning 395

6. Bevis sætning 420. Vink: Brug sætning 417 til at vise at $A = \{n \in \mathbb{Z} \mid n \leq x\}$ er ikke tom og opadtil begrænset. Betragt sup A.

7. Bevis sætning 421. Vink: Vis først at for ethvert naturligt tal N findes der et $n \in \mathbb{Z}$ så

$$\frac{n}{N} \leq x < \frac{n+1}{N}. \tag{10.115}$$

brug dernæst sætning 418 til at vælge N passende.

8. Bevis sætning 422. Vink: Vælg N så $1/N < (y - x)$, og brug ideen fra foregående opgave.

Chapter 11

Beviser i reel analyse, i tilknytning til Lindstrøms "Kalkulus"

De reelle tal spiller en central rolle i den matematiske analyse, det vil sige den fortsættelse af differential- og integralregning, som benytter grænseværdier. Vi skal i dette kapitel indøve de metoder, som anvendes i denne del af matematikken. Det drejer sig især om de såkaldte epsilon-delta-definitioner og -beviser, som er karakteriseret ved intensiv brug af kvantorer og uligheder. Disse metoder blev udvikler i 1800-tallet, for at gøre argumenter om grænseværdier, differentiation og integration stringente. En tilvænning til den type argumenter er en forudsætning for alle senere kurser i matematisk analyse.

Argumentationsformen er hovedsagen i MatM, men vi skal naturligvis argumentere på noget. Vi vælger at behandle nogle emner, som allerede er behandlet på Mat-Intro. I har nok set beviserne for nogle af sætningerne før. Formålet med at tage dem op igen er, at I selv skal kunne mestre metoderne. I skal kunne gengive beviser for centrale sætninger selv, og selv finde på beviser for simple sætninger.

De to emnekredse, vi skal behandle, er reelle talfølger og kontinuerte reelle funktioner. Specielt vil vi vise, at enhver Cauchyfølge er konvergent, og at en kontinuert funktion på et lukket begrænset interval afbilder dette på et lukket begrænset interval. Dette er to hovedsætninger i den reelle analyse. De er behandlet i Tom Lindstrøm: "Kalkulus" Afsnit 4.3-4.4 og 5.1-5.3.

Dette kapitel er kun et supplement til Lindstrøms bog. Det vil uddybe enkelte punkter og give et alternativt bevis for en af sætningerne. Desuden vil vi lave en nøje analyse af et af beviserne, og diskutere, hvordan man bør læse matematik.

11.1 Reelle talfølger

Lindstrøm indfører en følge, som en uendelig følge af tal

$$a_1, a_2, a_3, \ldots, a_n, \ldots \tag{11.1}$$

Mere formelt kan vi definere følger ud fra funktionsbegrebet, som vi jo har funderet i mængdelæren:

Definition 424 *En **reel talfølge** (a_n) er en afbildning $\mathbb{N} \longrightarrow \mathbb{R}$. I stedet for at betegne billedet af $n \in \mathbb{N}$ med $a(n)$ betegner man det ved a_n.*

Bemærkning 425 *Der er forskel på mængden $\{a_n \mid n \in \mathbb{N}\}$ og følgen (a_n). I følgen er rækkefølgen af betydning, hvorimod den ikke er af betydning i mængden. Derfor er det også mere almindeligt at bruge almindelige parenteser om en følge, som vi gør det her, end at bruge tuborg-parenteser, som Lindstrøm gør.*

Gennemarbejd nu følgende dele af afsnit 4.3 og 4.4 i Lindstrøm (inklusive beviserne):

4.3.1, 4.3.3, 4.3.9 (inklusiv den foregående definition på voksende og aftagende følger), 4.4.1 - 4.4.4 og 4.4.6 - 4.4.10.

I slutbemærkningen formulerer Lindstrøm følgende sætning:

Sætning 426 *Lad L være et ordnet legeme. Da har L supremumsegenskaben, hvis og kun hvis L har følgende to egenskaber:*

- *I L er enhver Cauchy-følge konvergent.*
- *L har den arkimediske egenskab.*

11.2 Kontinuerte funktioner

Gennemarbejd følgende dele af afsnit 5.1 i Lindstrøm:

5.1.1, 5.1.5, 5.1.7, 5.1.9 - 5.1.11.

11.2.1 Kontinuitet af et produkt. Et bevis, og en bevisanalyse.

I sætning 5.1.5 beviser Lindstrøm kun kontinuiteten af summen af to kontinuerte funktioner. Jeg vil her give et bevis for den mere komplicerede sætning om produktet. Først giver jeg et poleret bevis, hvori δ'er trækkes op af hatten uden forklaring, og derefter vil jeg afsløre den bagvedliggende analyse:

Sætning 427 *(Del af sætning 5.1.5 i Lindstrøm). Lad $A \subseteq \mathbb{R}$, og lad $f, g : A \longrightarrow \mathbb{R}$ være kontinuerte i $a \in A$. Da er $f \cdot g$ kontinuert i a.*

Bevis. Antag at f og g er kontinuerte i a, altså at

$$\forall \epsilon > 0 \exists \delta > 0 \forall x \in A : |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon \tag{11.2}$$

$$\forall \epsilon > 0 \exists \delta > 0 \forall x \in A : |x - a| < \delta \Rightarrow |g(x) - g(a)| < \epsilon. \tag{11.3}$$

Vi skal vise at $f \cdot g$ er kontinuert i a, altså at

$$\forall \epsilon > 0 \exists \delta > 0 \forall x \in A : |x - a| < \delta \Rightarrow |f(x)g(x) - f(a)g(a)| < \epsilon \tag{11.4}$$

Lad $\epsilon > 0$.
Ifølge (11.2) kan vi bestemme δ_1, så det for alle $x \in A$ gælder, at

$$|x - a| < \delta_1 \Rightarrow |f(x) - f(a)| < 1. \tag{11.5}$$

Ifølge (11.2) kan vi bestemme δ_2, så det for alle $x \in A$ gælder, at

$$|x - a| < \delta_2 \Rightarrow |f(x) - f(a)| < \frac{\epsilon}{2\left(|g(a)| + 1\right)}. \tag{11.6}$$

Endelig kan vi ifølge (11.3) bestemme δ_3, så det for alle $x \in A$ gælder, at

$$|x - a| < \delta_3 \Rightarrow |g(x) - g(a)| < \frac{\epsilon}{2\left(|f(a)| + 1\right)}. \tag{11.7}$$

Sæt nu $\delta = \min(\delta_1, \delta_2, \delta_3)$, og antag, at $x \in A$, og at $|x - a| < \delta$. Vi skal da vise, at

$$|f(x)g(x) - f(a)g(a)| < \epsilon. \tag{11.8}$$

Da $|x - a| < \delta_1$, har vi at $|f(x) - f(a)| < 1$. Trekantsuligheden (Lindstrøm 2.1.3) fortæller os nu at

$$|f(x)| - |f(a)| \leq ||f(x)| - |f(a)|| \leq |f(x) - f(a)| < 1, \tag{11.9}$$

hvoraf vi slutter at

$$|f(x)| < |f(a)| + 1 \tag{11.10}$$

Da endvidere $|x - a| < \delta_2$ og $|x - a| < \delta_3$, kan vi ved brug af trekantsuligheden slutte, at

$$\begin{aligned}
& |f(x)g(x) - f(a)g(a)| && (11.11)\\
= \; & |f(x)g(x) - f(x)g(a) + f(x)g(a) - f(a)g(a)| && (11.12)\\
\leq \; & |f(x)g(x) - f(x)g(a)| + |f(x)g(a) - f(a)g(a)| && (11.13)\\
= \; & |f(x)|\,|g(x) - g(a)| + |f(x) - f(a)|\,|g(a)| && (11.14)\\
< \; & \left(|f(a)| + 1\right) \frac{\epsilon}{2\left(|f(a)| + 1\right)} + \frac{\epsilon}{2\left(|g(a)| + 1\right)} |g(a)| < \epsilon && (11.15)
\end{aligned}$$

■

Bevisanalyse: Lad os se på den bevisanalyse, som førte frem til dette ret komplicerede bevis. Specielt skal vi se hvordan vi kom på vurderingerne i (11.5) - (11.7)

Det er klart at fg er en funktion $A \longrightarrow \mathbb{R}$. Det vi skal vise er at den er kontinuert i a, altså at

$$\forall \epsilon > 0 \exists \delta > 0 \forall x \in A : |x - a| < \delta \Rightarrow |f(x)g(x) - f(a)g(a)| < \epsilon. \tag{11.16}$$

Det er altså dette udsagn, som er endemålet for vores kæde af deduktioner. Fra formuleringen af problemet er det også ret klart, at vi undervejs i beviset skal bruge at f og g er kontinuerte, altså at

$$\forall \epsilon > 0 \exists \delta > 0 \forall x \in A : |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon \tag{11.17}$$

$$\forall \epsilon > 0 \exists \delta > 0 \forall x \in A : |x - a| < \delta \Rightarrow |g(x) - g(a)| < \epsilon. \tag{11.18}$$

Disse udsagn, som vi altså kan behandle som sande, skal nok være blandt de udsagn som deduktionen begynder med. Det er mindre klart, hvilke andre ting vi skal bruge som start på eller undervejs i deduktionerne. Skal vi bruge nogle særlige egenskaber ved de reelle tal som for eksempel fuldstændigheden? Det må vise sig ved analysen af beviset.

Lad os nu se på det udsagn vi skal ende med nemlig (11.16). Det starter med en alkvantor. Vi "lader" derfor ϵ være et tal større end 0. For denne givne værdi ϵ, skal vi vise, at der eksisterer et $\delta > 0$, så

$$\forall x \in A : |x - a| < \delta \Rightarrow |fg(x) - fg(a)| < \epsilon. \tag{11.19}$$

For at bevise eksistensen af et sådant δ, skal vi finde en kandidat og vise, at den duer. Hvor får vi kandidaten fra? Jo, vi skal naturligvis bruge (11.17) og (11.18). Mere uformelt kan vi sige, at vi skal gøre $|f(x)g(x) - f(a)g(a)|$ lille, når $|x - a|$ er lille, og det vi ved er, at både $|f(x) - f(a)|$ og $|g(x) - g(a)|$ kan gøres små, når $|x - a|$ er lille. Det ville altså være rart, hvis vi kunne omforme $|f(x)g(x) - f(a)g(a)|$ til et udtryk, der involverer $|f(x) - f(a)|$ og $|g(x) - g(a)|$. Desværre gælder det ikke at $|f(x)g(x) - f(a)g(a)| = |f(x) - f(a)|\,|g(x) - g(a)|$. Vi må være mere snedige. I stedet bruger vi følgende omskrivning:

$$\begin{aligned}
& |f(x)g(x) - f(a)g(a)| && (11.20)\\
= \; & |f(x)g(x) - f(x)g(a) + f(x)g(a) - f(a)g(a)| && (11.21)\\
\leq \; & |f(x)g(x) - f(x)g(a)| + |f(x)g(a) - f(a)g(a)| && (11.22)\\
= \; & |f(x)|\,|g(x) - g(a)| + |f(x) - f(a)|\,|g(a)| && (11.23)
\end{aligned}$$

Denne omskrivning er et inspireret gæt for en nybegynder, men bliver en standard metode, når man har set nogle beviser af denne type. Ved omformningen har vi opnået, at $|f(x) - f(a)|$ og $|g(x) - g(a)|$ er kommet ind i billedet. Dem kan vi gøre små. Hvis vi kan gøre dem så små, at $|f(x)|\,|g(x) - g(a)|$ og $|f(x) - f()|\,|g(a)|$ begge bliver mindre end $\epsilon/2$, er vi hjemme. Problemet er bare, at $|f(x) - f(a)|$ og $|g(x) - g(a)|$ er ganget sammen med henholdsvist $|g(a)|$ og $|f(x)|$.

Sidste led: Her giver $|g(a)|$ ikke noget problem, da a og dermed $|g(a)|$ er en fast konstant. For at gøre $|f(x) - f(a)|\,|g(a)|$ mindre end $\epsilon/2$, skal vi bare

gøre $|f(x)-f(a)|$ mindre end $\frac{\epsilon}{2|g(a)|}$. Dette udtryk giver dog kun menig hvis $|g(a)| \neq 0$. Det kan vi komme uden om på to måder: Enten kan vi bemærke at hvis $|g(a)| = 0$, er leddet $|f(x)-f(a)|\,|g(a)|$ selv lig 0, og så behøver vi ikke vurdere yderligere på det. Eller også kan vi bare gøre $|f(x)-f(a)|$ mindre end $\frac{\epsilon}{2(|g(a)|+1)}$[1], som jo altid giver mening og er mindre end $\frac{\epsilon}{2|g(a)|}$ når $|g(a)| \neq 0$. Vi vælger sidste udvej, fordi vi derved undgår at behandle de to tilfælde $|g(a)| \neq 0$ og $|g(a)| = 0$ forskelligt. For at gøre $|f(x)-f(a)|$ mindre end $\frac{\epsilon}{2(|g(a)|+1)}$ skal vi bare vælge et δ_1 i (11.17) svarende til et $\epsilon_1 = \frac{\epsilon}{2(|g(a)|+1)}$.

Første led: Det er ikke så let at få vurderet $|f(x)|\,|g(x)-g(a)|$ mindre end $\epsilon/2$. For her varierer $|f(x)|$ med x, og i definitionen (11.18) skal δ vælges uafhængigt af x (x optræder i udsagnet efter δ). Vi ville dog kunne lave en vurdering magen til vurderingen ovenfor, hvis vi kunne finde en øvre grænse på $|f(x)|$ (bemærk her den typiske analysestrategi: vi ved hvor vi vil hen og finder så et udsagn, som ville medføre denne konklusion). Vi kunne måske prøve at bevise, at funktionen er begrænset på hele mængden A. Analysen har ført til et åbent problem: Er en kontinuert funktion begrænset? Svaret er nej. Man kunne så spørge: Kan man gøre en antagelse om A, så en kontinuert funktion defineret på A er begrænset. Vi kommer til at se at svaret er ja: Hvis A er et lukket og begrænset interval, er en kontinuert funktion begrænset på A. Men vi har ikke forudsat at A er et lukket og begrænset interval, og selv hvis det var, ville brug af denne sætning være at skyde spurve med kanoner.

Vi skal jo huske på, at vi ikke behøver at finde en øvre grænse for $|f(x)|$ i hele A, men kun i en lille δ-omegn om a, hvor vi selv kan vælge δ. Nu bemærker vi så, at vi kan gøre $|f(x)-f(a)| < \epsilon_2$

hvis bare $|x-a|$ er mindre end et passende δ, og når afstanden fra $f(a)$ til $f(x)$ er mindre end ϵ_2, så må $|f(x)|$ jo være mindre end $|f(a)|+\epsilon_2$ (trekantsuligheden). Altså kan $|f(a)|+\epsilon_2$ bruges som øvre grænse for $|f(x)|$ i en δ_2-omegn om a. Vi kan vælge ϵ_2 vilkårlig, så længe den bare er positiv; vi vælger $\epsilon_2 = 1$.[2]

Nu, hvor vi har vurderet $|f(x)| < |f(a)|+1$, kan vi gøre $|f(x)|\,|g(x)-g(a)|$ mindre end $\epsilon/2$, hvis vi bare gør $|g(x)-g(a)|$ mindre end $\frac{\epsilon}{2(|f(a)|+1)}$, og det gør vi ved at vælge et δ_3 i (11.18) svarende til $\epsilon_3 = \frac{\epsilon}{2(|f(a)|+1)}$.

I vurderingen af (11.20) skal vi bruge alle de tre ovenstående vurderinger, så vi skal gøre $|x-a|$ mindre end de tre δ'er vi bestemte ovenfor, altså mindre end $\delta = \min(\delta_1, \delta_2, \delta_3)$.

Dermed er vi klar til det syntetiske bevis, som blev givet ovenfor.

11.3 Hvordan man skal læse matematik

Ovenfor og i kapitlet om analyse og syntese har vi set på hvordan man kan få ideer til beviser, og vi har især i sidste eksempel set på, hvordan man så til

[1] Her har vi valgt at lægge 2 til i nævneren. Det er et helt arbitrært valg. ethvert tal større end 0 ville lige så vel kunne bruges.

[2] Vi kunne naturligvis også vælge ϵ_1 lig det givne ϵ, men det vil give det forkerte indtryk, at valget afhænger af det givne ϵ. Valget ϵ_1 giver i øvrigt en æstetisk symmetri mellem de to led i vurderingen.

slut omformulerer sine ideer til det formelle bevis, som kan "publiceres" (eller afleveres til læreren). I dette afsnit skal vi se på det omvendte problem: Hvordan skal man læse de publicerede formelle beviser og forstå dem? Som eksempel ser vi på det bevis, som vi formulerede i forrige afsnit.

Lad altså som om, du ser sætningen og beviset for første gang. Sandsynligvis vil du allerede ved læsningen af sætningen danne dig nogle billeder af hvad den siger. Måske noget i stil med: Når hverken f eller g ('s graf) springer, så springer deres produkt heller ikke. Det lyder intuitivt plausibelt. Naturligvis ved vi godt at kontinuitet betyder noget mere præcist end "springer ikke", og vi kender måske også eksempler på at denne intuition kan være vildledende. Men det er vigtigt, at vi har disse intuitive billeder af sætningens indhold for øje. Uden intuition, analogier mm. bliver matematik en indholdsløs formalisme og en samling indholdsløse manipulationer, som ikke er til at huske eller anvende, og som ikke kan fænge læseren. Når du læser matematik, så prøv at trænge ind bag den formelle facade, og dan dig intuitive billeder af objekter og sætninger, men vær samtidig klar over at disse billeder kan være vildledende og måske senere bør udskiftes med andre billeder.

Ihukommende definitionen af kontinuitet, kunne næste trin i forståelsen af sætningen måske være at du tænkte: Sætningen udsiger at hvis ændringen i f kan gøres lille og ændringen i g kan gøres lille, så vil ændringen i fg kunne gøres lille. Når du således har prøvet at danne dig en intuition om hvad sætningen siger, kan du gå i gang med at læse beviset.

Første problem i læsningen af beviset er at forstå hvert enkelt skridt i kæden af slutninger. Ved første gennemlæsning kan det være det eneste man kærer sig om. Når man er kommet igennem denne proces, har man overbevist sig om, at beviset holder, og at sætningen derfor er korrekt.

Men man bør ikke stille sig tilfreds med denne overfladiske forståelse. Man skal prøve at forstå, hvorfor de enkelte skridt i beviset ser ud som de gør. Ved første gennemlæsning af beviset er der sikkert flere steder, hvor du undrer dig over hvorfor man nu lige skal gøre, som der står i teksten (hvis ikke er du et matematisk geni eller alt for autoritetstro). Hvorfor skal vi gøre $|f(x) - f(a)|$ mindre end 1, og hvorfor ser vi lige på de underlige størrelser $\frac{\epsilon}{2(|g(a)|+1)}$ og $\frac{\epsilon}{2(|f(a)|+1)}$? Disse spørgsmål kan du nok besvare ved at gennemgå beviset endnu en gang, efter at du har set omformningen (11.13). Du bør blive ved med at gennemgå bevisets dele, indtil alle trin forekommer dig naturlige.

Ja, du bør gå endnu videre. Du bør undersøge beviset som en helhed, indtil dets struktur forekommer dig naturlig. En god test af om man har forstået et bevis, er om man kan gengive det uden at huske de enkelte trin udenad. I det ovenstående bevis bør man først være tilfreds med sin forståelse, når man kan gengive beviset uden at huske de enkelte vurderinger udenad. Det eneste trin, man måske skal lære sig udenad, er omformningen af $|f(x)g(x) - f(a)g(a)|$ ved at trække $f(x)g(a)$ fra og lægge det til igen[3]. Dette trin er et trick, som først bliver naturligt, når man har set det i spil i flere beviser.

[3]Du bør naturligvis også indse, at du lige så godt kunne trække $f(a)g(x)$ fra og lægge det til igen.

For en dybere forståelse af sætningen, bør man gå videre og spørge sig selv, hvor de enkelte forudsætninger blev brugt i beviset. Hvis ikke alle forudsætningerne blev brugt, kan du generalisere sætningen til en situation, hvor du kun forudsætter det du bruger i beviset. Du bør også være sikker på, at du forstår, hvorfor man ikke skal forudsætte mere end man gør. For eksempel bør du i det her analyserede tilfælde kunne forstå, hvorfor det ikke er nødvendigt at forudsætte, at funktionerne er definerede i en hel omegn om a. Hvad siger sætningen for eksempel, hvis funktionerne er defineret på $[a, b]$, eller hvis a er et isoleret punkt i definitionsmængden?

Man kan også spørge sig selv om ikke lignende sætninger gælder for summen, differensen og kvotienten af to kontinuerte funktioner (det gør de, se Lindstrøm 5.1.5), og man kan prøve at lave beviser for disse sætninger. Man kan også spørge, om der mon ikke gælder lignende sætninger for grænseværdier af følger, eller lignende.

Når man således har forstået beviset for sætningen og har afsøgt mulighederne for at generalisere den, bør man også prøve at forstå, hvilken plads den har i den deduktive struktur af den teori, den indgår i. En test af om denne struktur er forstået er, at man synes den er naturlig i den forstand, at hver sætning indtager en naturlig plads i bygningsværket. Man bør kunne gøre rede for strukturen, uden at kunne de enkelte sætninger udenad. Alle disse øvelser vil føre til en bedre forståelse af det behandlede stykke matematik. **Når I som studerende læser matematik bør I stræbe efter denne dybe forståelse.**

Forsøgene på at generalisere sætninger er et første skridt på vejen mod egen matematisk kreativitet. Næste skridt er at prøve at formulere helt nye beviser, helt nye strukturer af det behandlede stykke matematik, helt nye problemer og sætninger inden for teorien, helt nye begreber og til slut helt nye matematiske områder. Hvis I når så langt er I matematiske forskere.

11.4 Mellemværdisætningen

Mellemværdisætningen siger, at en kontinuert funktion defineret på et interval $[a, b]$ antager alle værdier mellem $f(a)$ og $f(b)$. Først viser vi den sætning, som i Lindstrøm kaldes Skæringssætningen (5.2.1). Vi skal give et andet bevis end i Lindstrøm.

Vi starter med et lemma, som kan være nyttigt i mange andre sammenhænge:

Lemma 428 *Lad $f : [a, b] \longrightarrow \mathbb{R}$ være en funktion, som er kontinuert i et punkt $c \in [a, b]$, og antag at $f(c) > 0$.*

Da findes et $\delta > 0$ så

$$x \in [a, b] \cap [c - \delta, c + \delta] \Rightarrow f(x) > 0. \tag{11.24}$$

Løst sagt: Hvis en kontinuert funktion er positiv i et punkt, så er den positiv i en hel omegn af punktet.

Bevis. Sæt $\epsilon = \frac{1}{2}f(c)$. Da er $\epsilon > 0$, og da f er kontinuert i c, eksisterer et $\delta_1 > 0$, så

$$x \in [a,b] \wedge |x-c| < \delta_1 \Rightarrow |f(x) - f(c)| < \frac{1}{2}f(c). \tag{11.25}$$

Sæt nu $\delta = \frac{1}{2}\delta_1$. Da gælder:

$$x \in [a,b] \cap [c-\delta, c+\delta] \Rightarrow x \in [a,b] \wedge |x-c| \leq \frac{1}{2}\delta_1 \tag{11.26}$$

$$\Rightarrow \quad x \in [a,b] \wedge |x-c| < \delta_1 \Rightarrow |f(x) - f(c)| < \frac{1}{2}f(c) \tag{11.27}$$

$$\Rightarrow \quad -f(x) + f(c) < \frac{1}{2}f(c) \Rightarrow \frac{1}{2}f(c) < f(x) \Rightarrow f(x) > 0. \tag{11.28}$$

■

Lemma 429 *Lad $f : [a,b] \longrightarrow \mathbb{R}$ være en funktion, som er kontinuert i et punkt $c \in [a,b]$, og antag at $f(c) < 0$.*

Da findes et $\delta > 0$, så

$$x \in [a,b] \cap [c-\delta, c+\delta] \Rightarrow f(x) < 0. \tag{11.29}$$

Bevis. Brug lemma (428) på funktionen $g := -f$. ■

Sætning 430 ***Skæringssætningen.*** *Lad $f : [a,b] \longrightarrow \mathbb{R}$ være en kontinuert funktion, hvor $f(a)$ og $f(b)$ har modsatte fortegn. Da findes et $c \in]a,b[$, så $f(c) = 0$.*

Bevis. At $f(a)$ og $f(b)$ har modsatte fortegn betyder, at vi enten har

$$f(a) < 0 \quad \text{og} \quad f(b) > 0 \tag{11.30}$$

eller

$$f(a) > 0 \quad \text{og} \quad f(b) < 0. \tag{11.31}$$

Vi kan nu bemærke, at dersom vi kan bevise sætningen for alle situationer af type (11.30), så følger den også for alle situationer af type (11.31): For har vi for en given kontinuert funktion f tilfældet (11.31) foreliggende, kan vi betragte den ligeledes kontinuerte funktion $g := -f$, for hvilken situationen (11.30) da foreligger; har vi vist sætningen for tilfælde (11.30), følger da eksistensen af et $c \in]a,b[$, så $g(c) = 0$. Men da er jo også $f(c) = 0$.

Vi kan altså - som man siger - "gerne antage" tilfældet (11.30), hvilket vi nu gør.

Betragt nu mængden:

$$C := \{x \in [a,b] \mid f(x) \leq 0\}. \tag{11.32}$$

Da $f(a) < 0$, er $a \in C$, og C er således ikke tom. På den anden side er $C \subseteq [a,b]$, og specielt er b en majorant for C, så C er opadtil begrænset. Idet $\mathbb{R}$ har supremumsegenskaben, følger det, at C har et supremum i $\mathbb{R}$. Sæt:

$$c := \sup C. \tag{11.33}$$

Hvis vi kan vise, at $f(c) = 0$, er det klart at $c \in]a, b[$, og sætningen vil dermed være bevist. Trikotomiloven siger, at enten er $f(c) = 0$, eller $f(c) > 0$, eller $f(c) < 0$. Vi kan altså vise at $f(c) = 0$, og dermed sætningen, ved at vise at $f(c) > 0$ og $f(c) < 0$ begge fører til modstrid. Det gør vi nu:

Antag først at $f(c) > 0$. Ifølge lemma 428 findes et $\delta > 0$, så

$$x \in [a, b] \cap [c - \delta, c + \delta] \Rightarrow f(x) > 0. \tag{11.34}$$

Da $c = \sup C$, eksisterer der et $x_0 \in C$, så $|x_0 - c| < \delta$ (overvej). Da $x_0 \in C$, vil $f(x_0) \leq 0$; men da $|x_0 - c| < \delta$, fås af (11.34), at $f(x_0) > 0$.Det er en modstrid.

Antag derefter $f(c) < 0$. Ifølge lemma 429 findes et $\delta > 0$, så

$$x \in [a, b] \cap [c - \delta, c + \delta] \Rightarrow f(x) < 0. \tag{11.35}$$

Da $f(b) > 0$, fås ved kontraposition af (11.35), at $b \notin [a, b] \cap [c - \delta, c + \delta]$ og dermed, at $b \notin [c - \delta, c + \delta]$. Da endvidere $b > c$, må det gælde, at $b > c + \delta$. Altså gælder at $(c + \delta) \in [a, b] \cap [c - \delta, c + \delta]$, så af (11.35) ses, at $f(c + \delta) < 0$. Det betyder, at $(c + \delta) \in C$, i modstrid med at $c = \sup C$ (og dermed specielt at c er en majorant for C). ■

Læs også Lindstrøms bevis for skæringssætningen. De to forskellige beviser illustrerer, at matematiske sætninger ofte kan bevises på ret forskellige måder. Hvilket bevis man foretrækker, er et spørgsmål om smag. Lindstrøms bevis er måske mere intuitivt apellerende, og derfor måske mere pædagogisk. På den anden side involverer det talfølger, der sådan set er sagen uvedkommende. Beviset ovenfor er mere direkte.

Desuden indeholder Lindstrøms bevis forskellige løse ender. For det første er det ikke helt klart, at der findes et $n \in \mathbb{N}$ så $c + 1/n < b$. Dette problem kan klares ved at bytte om på de to dele af beviset. For hvis vi først har bevist, at $f(c) \leq 0$, og vi ved at $f(b) > 0$, følger det at $b > c$. De reelle tals arkimediske egenskab siger da, at der findes et $n \in \mathbb{N}$, så $1/n < b - c$, hvorfra vi slutter at $c + 1/n < b$.

Desuden bruger Lindstrøms bevis, at når alle leddene i en følge a_n er positive, og følgen konvergerer mod c, da er $c \geq 0$. Det er en konsekvens af følgende lemma:

Lemma 431 *Lad (a_n) være en reel talfølge, og antag at $a_n \to a$ for $n \to \infty$. Da gælder:*

1. *Hvis $a_n \geq b$ for alle $n \in \mathbb{N}$, da er $a \geq b$.*
2. *Hvis $a_n \leq b$ for alle $n \in \mathbb{N}$, da er $a \leq b$.*

Bevis. Vi beviser punkt 1. Punkt 2 overlades til læseren.

Antag altså at $a_n \to a$ for $n \to \infty$ og at $a_n \geq b$. Beviset føres ved modstrid. Vi antager altså, at $a < b$. Da er $\epsilon := \frac{b-a}{2}$ positiv, og da $a_n \to a$ for $n \to \infty$, eksisterer der et $N \in \mathbb{N}$, så

$$|a_n - a| < \frac{b - a}{2} \text{ for alle } n \geq N. \tag{11.36}$$

Heraf følger, at

$$a_n - a < \frac{b-a}{2} \text{ for alle } n \geq N, \tag{11.37}$$

hvoraf

$$a_n < a + \frac{b-a}{2} = b - \left(\frac{b-a}{2}\right) < b \text{ for alle } n \geq N. \tag{11.38}$$

Men det er i modstrid med antagelsen om at $a_n \geq b$ for alle $n \in \mathbb{N}$. ■

Sætning 432 ***Mellemværdisætningen.*** *Lad $f : [a,b] \longrightarrow \mathbb{R}$ være en kontinuert funktion. Lad y_0 være et tal mellem $f(a)$ og $f(b)$. Da eksisterer et $c \in [a,b]$, så $f(c) = y_0$.*

Bevis. Hvis $f(a) = f(b)$, kan c vælges lig a.

Hvis $f(a) \neq f(b)$, opfylder funktionen $f - y_0$ antagelserne i skæringssætningen. Der eksisterer derfor et $c \in]a,b[\subseteq [a,b]$, så $f(c) - y_0 = 0$. Heraf fås som ønsket, at $f(c) = y_0$. ■

Gennemgå nu Lindstrøms Korollar 5.2.2 og de efterfølgende eksempler og bemærkninger.

11.5 Ekstremalværdisætningen

Læs Lindstrøms afsnit 5.3. I beviset for sætning 5.3.2 og 5.3.5 får du brug for følgende to lemmaer:

Lemma 433 *Lad (a_n) og (b_n) være reelle talfølger. Hvis*

$$\lim_{n\to\infty} a_n = c \tag{11.39}$$

og

$$\lim_{n\to\infty} (a_n - b_n) = 0, \tag{11.40}$$

da er (b_n) konvergent og

$$\lim_{n\to\infty} b_n = c. \tag{11.41}$$

Bevis. Vi kan skrive

$$b_n = a_n - (a_n - b_n), \tag{11.42}$$

så da (a_n) og $(a_n - b_n)$ begge er konvergente (med grænseværdier henholdsvis c og 0), er $b_n = a_n - (a_n - b_n)$ også konvergent ifølge Lindstrøm sætning 4.3.3, og dens grænseværdi er $c - 0 = c$. ■

Lemma 434 ***Klemme-lemma.*** *Lad (a_n), (b_n), og (c_n) være reelle talfølger med*

$$a_n \leq c_n \leq b_n \text{ for alle } n \in \mathbb{N}. \tag{11.43}$$

Hvis (a_n) og (b_n) er konvergente med samme grænseværdi

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = c, \tag{11.44}$$

da er (c_n) konvergent og har samme grænseværdi:

$$\lim_{n\to\infty} c_n = c \tag{11.45}$$

Bevis. Hvis vi kan vise, at følgen $(a_n - c_n)$ er konvergent med grænseværdi 0, så følger sætningen af Lemma 433.

Ifølge Lindstrøm sætning 4.3.3 gælder, at

$$\lim_{n\to\infty} (a_n - b_n) = 0. \tag{11.46}$$

Til et givet $\epsilon > 0$ kan vi altså bestemme et $N \in \mathbb{N}$, så $|a_n - b_n| < \epsilon$ for alle $n > N$. Men fra (11.43) ses, at

$$|a_n - c_n| \leq |a_n - b_n| \text{ for alle } n \in \mathbb{N}., \tag{11.47}$$

hvoraf $|a_n - c_n| < \epsilon$ for alle $n > N$. Heraf ses, at

$$\lim_{n\to\infty} (a_n - c_n) = 0. \tag{11.48}$$

■

I Lindstrøms beviser for sætning 5.3.2 og ekstremalværdisætningen, henviser han intetsteds til, at funktionens definitionsinterval er lukket. Man skulle så tro, at sætningen ville gælde også for åbne intervaller, men som han viser med eksempler (side 226) er sætningerne ikke sande, hvis definitionsintervallet ikke er lukket.

Grunden er, at beviset faktisk benytter at intervallet er lukket, uden at dette nævnes eksplicit. Beviserne bygger jo på, at følgen (c_n) af elementer i intervallet, konvergerer mod et punkt *i intervallet*. Hvis intervallet er lukket, følger det af Lemma 431. Men hvis intervallet ikke er lukket, kan det ske at grænseværdien ligger uden for intervallet. For eksempel gælder at $1/n \in]0,1]$ for alle $n \in \mathbb{N}$, men $\lim_{n\to\infty} 1/n = 0 \notin]0,1]$. Hvis man prøver at gennemføre Lindstrøms bevis for sætning 5.3.2 i det tilfælde, at funktionen er $1/x$ defineret på $]0,1]$, vil man netop få en følge (c_n), hvor $\lim_{n\to\infty} c_n = 0$. Og så virker beviset ikke længere.

Vi vil nu kombinere mellemværdisætningen og ekstremalværdisætningen til følgende hovedsætning om kontinuerte funktioner:

Sætning 435 *En kontinuert funktion afbilder et lukket begrænset interval på et lukket begrænset interval.*

Bevis. Lad Lad $f : [a,b] \longrightarrow \mathbb{R}$ være en kontinuert funktion. Vi skal vise at $f([a,b])$ er et lukket begrænset interval.

Fra ekstremalværdisætningen ved vi at der findes $c, d \in [a,b]$ så

$$\max f([a,b]) = f(c), \tag{11.49}$$
$$\min f([a,b]) = f(d). \tag{11.50}$$

Fra mellemværdisætningen ved vi, at f antager alle værdier mellem $f(c)$ og $f(d)$. Derfor er $f([a,b]) = [f(d), f(c)]$, altså et lukket begrænset interval. ■

11.6 Opgaver

1. Bevis sætning 4.3.3 (ii) i Lindstrøm.
2. Bevis sætning 4.3.3 (iii) i Lindstrøm.
3. Bevis sætning 4.3.3 (iv) i Lindstrøm.
4. Bevis den del af sætning 5.1.5 i Lindstrøm som handler om $f - g$.
5. Bevis den del af sætning 5.1.5 i Lindstrøm som handler om f/g.

Chapter 12

Construction of the real numbers.

12.1 Motivation.

It will not come as a big surprise to anyone when I say that we need the real numbers in mathematics. More to the point, we need to be able to talk *precisely* about real numbers and their properties, and to rigorously prove theorems whose statements and/or proofs involve the system of real numbers.

In mathematics, to be able to 'talk precisely' about real numbers can of course only mean 1 thing: Namely, to be able to talk about them *on an axiomatic basis*.

Fine. Why don't we just introduce some axioms for the real numbers that seem to work all right? For instance, why don't we just accept the axioms introduced in Chapter 10 and then work on that basis? That seems to work well enough, right?

There are 2 reasons why this is not acceptable in the long run. The first reason is that it is a question of aesthetics: If we can introduce the real numbers without accepting any further axioms than the axioms of set theory then that seems like a more satisfactory and much nicer situation. This reason also ties up with the principle of 'Occam's razor' which states that in science one should not introduce any more hypotheses than absolutely necessary.

The second reason is far more important though. On a deeper level it is in fact connected with the first reason but goes in the direction of answering the following question: 'How do we know that the axiomatic basis of mathematics does not allow us to prove contradictory statements?' I.e., how can we know that it is not possible to prove both a theorem and the negation of that theorem on the basis of our accepted axiomatic system?

The point is that it is usually extremely hard to answer such questions. For instance, one can prove that it is *not possible* for the system of axioms

of set theory − i.e., the modern basis of all mathematics − to prove its own 'consistency', that is, that there are no hidden contradictions, − unless, of course, the system does in fact contain such contradictions (for in that case one would be able to prove *any* statement on the basis of the axioms). (Do not worry too much about this though; no contradictions have turned up after a century of scrutiny, and in the unlikely event that a contradiction should turn up you can be sure that bridges will not suddenly start to collapse, or that space ships will begin to miss their destinations because of that. If a contradiction turned up we would simply have to reconsider the situation and construct a new axiomatic system that does for us what we want of it).

However, we can take a few precautions: We can be as conservative as we can when it comes to introducing new axioms in mathematics. Thus, by constructing the real numbers on the basis of the axioms that we already have we can be certain that there will not result any *new* contradictions for some deep or hidden reason.

So, how do we construct the real numbers?

The fundamental idea about the real numbers is that − whatever they are − they should be objects that can be approximated by rational numbers. For instance, if we think of real numbers as having − possibly infinite − decimal expansions then this is just one way of thinking about real numbers as 'limits' of rational numbers: For a(n) (infinite) decimal expansion is nothing but a sequence of *finite* decimal expansions where we add more and more digits in the expansion. And a finite decimal expansion is a rational number (Exercise 464).

But there is nothing special about decimal expansions: We can consider any sequences of rational numbers. The important question then is: Which sequences of rational numbers should be thought of as 'approximating real numbers'? Notice that, from a formal point of view, the question is meaningless: For since we have not yet constructed the real numbers we can not attach any precise meaning to the phrase 'approximating real numbers' ...

But that the question is formally meaningless does *not* mean that it's mathematically pointless or trivial. Rather, the point is that we are searching for a *good definition*, i.e., a definition that will ultimately lead to a satisfactory theory.

To make a long story short, it turns out that the notion of a 'Cauchy sequence of rational numbers' − see the precise definition in the next section − not only captures much of the intuition we may have about 'approximating sequences of rational numbers', but also leads to a satisfactory theory. How people came up with this definition in the first place is a question in the history of mathematics and we will not go into that in this course.

When we have defined what Cauchy sequences of rational numbers are the great idea is then this: We want to think about the Cauchy sequences as sequences of approximations to real numbers but we still don't have the real

numbers. How do we get them? Answer: We simply *identify* them with these sequences ...

This causes a small problem to consider: Our intuition tells us that there are many different sequences approximating a real number. For instance, it seems reasonable to think of the sequences

$$(1+1, 1+\frac{1}{2}, 1+\frac{1}{4}, \ldots, 1+\frac{1}{2^n}, \ldots) \quad \text{and} \quad (1+1, 1+\frac{1}{2}, 1+\frac{1}{3}, \ldots, 1+\frac{1}{n}, \ldots)$$

as both being sequences approximating 1. But when should we generally consider 2 sequences $(a_1, a_2, \ldots)$ and $(b_1, b_2, \ldots)$ as approximations of the 'same real number'? Intuition tells us that this should be so precisely if the sequence $(a_1 - b_1, a_2 - b_2, \ldots)$ approximates the rational number 0; and if this is the case, we should consider the sequences $(a_1, a_2, \ldots)$ and $(b_1, b_2, \ldots)$ as being − or rather *representing* − the same real number.

The mathematically exact way of doing this is to introduce a certain equivalence relation between Cauchy sequences of rational numbers. The real numbers are then *defined* as the corresponding equivalence classes.

12.2 Definition of the real numbers.

On the basis of the motivational remarks above we now proceed to actually construct the system of real numbers.

12.2.1 Fundamental definitions and basic properties.

Definition 436 *Let $\alpha = (a_1, a_2, \ldots, a_n, \ldots)$ be a sequence of rational numbers, i.e., $a_n \in \mathbb{Q}$, for all $n \in \mathbb{N}$.*

We say that α is a ***Cauchy sequence*** *of rational numbers if for every positive rational number ϵ there is (depending on ϵ) an $N \in \mathbb{N}$ such that:*

$$|a_m - a_n| < \epsilon \quad \textit{whenever } m, n \geq N \ .$$

I.e., α is called a Cauchy sequence if

$$\forall \epsilon \in \mathbb{Q}_+ \exists N \in \mathbb{N} \forall m, n \in \mathbb{N} \ : \quad m, n \geq N \Rightarrow |a_m - a_n| < \epsilon \ .$$

We denote the set of Cauchy sequences of rational numbers by $\mathcal{C}$:

$$\mathcal{C} := \{\alpha \ \textit{ sequence of rational numbers} \mid \alpha \textit{ is a Cauchy sequence}\} \, .$$

If $\alpha = (a_1, a_2, \ldots, a_n, \ldots)$ is a sequence of rational numbers and $a \in \mathbb{Q}$ we say that α converges to a if for every positive rational number ϵ there is (depending on ϵ) an $N \in \mathbb{N}$ such that:

$$|a - a_n| < \epsilon \quad \textit{whenever } n \geq N \ .$$

I.e., we say that α converges to $a \in \mathbb{Q}$ if

$$\forall \epsilon \in \mathbb{Q}_+ \exists N \in \mathbb{N} \forall n \in \mathbb{N} \ : \quad n \geq N \Rightarrow |a - a_n| < \epsilon \ .$$

If α converges to a we also write $\alpha \to a$, or $a_n \to a$ for $n \to \infty$, or we say that a ***is the limit of*** *α.*

We call the sequence α ***convergent in*** *$\mathbb{Q}$ if it converges to some $a \in \mathbb{Q}$.*

The sequence α is called a ***null-sequence*** *if it converges to the rational number 0. We denote the set of null-sequences by $\mathcal{N}$:*

$$\mathcal{N} := \{\alpha \ \ \textit{sequence of rational numbers} \mid \alpha \ \ \textit{is a null-sequence}\} \ .$$

Proposition 437 *Let $\alpha = (a_1, a_2, \ldots, a_n, \ldots)$ be a sequence of rational numbers.*

If α is convergent in $\mathbb{Q}$ then α is a Cauchy sequence.

In particular, every null-sequence is a Cauchy sequence:

$$\mathcal{N} \subseteq \mathcal{C}.$$

Bevis. Suppose that α is convergent in $\mathbb{Q}$. To show that α is a Cauchy sequence, let the positive rational number ϵ be given.

Since α converges in $\mathbb{Q}$ there is some $a \in \mathbb{Q}$ such that α converges to a. Since $\epsilon/2$ is a positive rational number there is an $N \in \mathbb{N}$ such that:

$$|a - a_n| < \epsilon/2 \quad \text{whenever } n \geq N \ .$$

Let $m, n \in \mathbb{N}$ be such that $m, n \geq N$. We then find:

$$|a_m - a_n| = |(a_m - a) + (a - a_n)| \leq |(a_m - a)| + |(a - a_n)| < \epsilon/2 + \epsilon/2 = \epsilon \ .$$

Since ϵ was arbitrary we have shown that α is a Cauchy sequence. ■

Definition 438 *Let $a \in \mathbb{Q}$. The sequence $(a, a, \ldots, a, \ldots)$ is called the* ***constant sequence*** *with term a. It is clear from the definition that any such constant sequence is a Cauchy sequence which in fact converges in $\mathbb{Q}$ to $a \in \mathbb{Q}$.*

It is a little harder to show the existence of Cauchy sequences that are *not* convergent in $\mathbb{Q}$. We will construct an example in exercise 470.

Next we proceed to show that Cauchy sequences can be added and multiplied in a natural way. We first need a little lemma.

Lemma 439 *Suppose that $\alpha = (a_1, a_2, \ldots, a_n, \ldots)$ is a Cauchy sequence of rational numbers.*

Then there is (depending on α) a positive rational number c such that:

$$|a_n| \leq c \quad \textit{for all } n \in \mathbb{N} \ .$$

(One says that the sequence α is ***bounded****).*

If α is not a null-sequence there is (depending on α) a positive rational number d and an $M \in \mathbb{N}$ such that:

$$|a_n| \geq d \quad \textit{whenever } n \geq M \ .$$

Bevis. Since α is a Cauchy sequence there is an $N \in \mathbb{N}$ such that $|a_m - a_n| < 1$ whenever $m, n \geq N$. Let c_0 be the maximum of the finitely many rational numbers $|a_1|, \ldots, |a_N|$, and put $c := c_0 + 1$. Then c is a positive rational number.

Let $n \in \mathbb{N}$. If $n \leq N$ we have $|a_n| \leq c_0 < c$. And if $n \geq N$ we find:

$$|a_n| = |(a_n - a_N) + a_N| \leq |a_n - a_N| + |a_N| \leq 1 + c_0 = c \ .$$

Suppose then that α is a Cauchy sequence which is *not* a null-sequence. This means that α does not converge to 0. By definition, this means that:

$$\exists \epsilon \in \mathbb{Q}_+ \forall N \in \mathbb{N} \exists n \in \mathbb{N} \ : \quad n \geq N \ \text{ and } \ |0 - a_n| \geq \epsilon \ ,$$

i.e., there is a positive rational number ϵ so that whenever $N \in \mathbb{N}$ we can always find an $n \geq N$ such that $|a_n| \geq \epsilon$. Fix one such positive rational number ϵ.

Now, since α is a Cauchy sequence, and since $\epsilon/2$ is a positive rational number there is an $N \in \mathbb{N}$ such that

$$|a_m - a_n| < \epsilon/2 \quad \text{whenever } m, n \geq N \ .$$

By the above there is an $n_0 \geq N$ such that $|a_{n_0}| \geq \epsilon$. But then we have for any $n \geq N$ that:

$$|a_n| = |a_{n_0} - (a_{n_0} - a_n)| \geq |a_{n_0}| - |a_{n_0} - a_n| \geq \epsilon - \epsilon/2 = \epsilon/2 \ .$$

So we can take $d = \epsilon/2$ and $M = N$. ■

Proposition 440 *Suppose that $\alpha = (a_1, a_2, \ldots, a_n, \ldots)$ and $\beta = (b_1, b_2, \ldots, b_n, \ldots)$ are Cauchy sequences of rational numbers. Define the sequences $\alpha + \beta$ and $\alpha \cdot \beta$ (also often simply written as $\alpha\beta$) of rational numbers as follows:*

$$\alpha + \beta := (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n, \ldots) \ ,$$

and

$$\alpha \cdot \beta := (a_1 b_1, a_2 b_2, \ldots, a_n b_n, \ldots) \ .$$

Then $\alpha + \beta$ and $\alpha \cdot \beta$ are both Cauchy sequences of rational numbers.

Bevis. Let a positive rational number ϵ be given.

Since α and β are Cauchy sequences there exist according to Lemma 439 positive rational numbers c_1 and c_2 such that $|a_n| \leq c_1$ and $|b_n| \leq c_2$ for all $n \in \mathbb{N}$. Let c be the largest of the 3 numbers 1, c_1, and c_2. Then we certainly have $|a_n| \leq c$ and $|b_n| \leq c$ for all $n \in \mathbb{N}$.

The number $\frac{\epsilon}{2c}$ is a positive rational number. Again since α and β are Cauchy sequences there exist by definition $N_1, N_2 \in \mathbb{N}$ such that $|a_m - a_n| < \frac{\epsilon}{2c}$ if $m, n \geq N_1$, and such that $|b_m - b_n| < \frac{\epsilon}{2c}$ if $m, n \geq N_2$.

Now let N be the largest of the 2 numbers N_1 and N_2. Then, whenever $m, n \geq N$ we deduce:

$$\begin{aligned} |(a_m + b_m) - (a_n + b_n)| &= |(a_m - a_n) + (b_m - b_n)| \\ &\leq |a_m - a_n| + |b_m - b_n| < \frac{\epsilon}{2c} + \frac{\epsilon}{2c} = \frac{\epsilon}{c} \leq \epsilon \ , \end{aligned}$$

since $c \geq 1$, and furthermore:

$$\begin{aligned} |a_m b_m - a_n b_n| &= |a_m(b_m - b_n) + b_n(a_m - a_n)| \leq |a_m||b_m - b_n| + |b_n||a_m - a_n| \\ &< c \cdot \frac{\epsilon}{2c} + c \cdot \frac{\epsilon}{2c} = \epsilon \ . \end{aligned}$$

■

If $\alpha = (a_1, a_2, \ldots, a_n, \ldots) \in \mathcal{C}$ and $q \in \mathbb{Q}$, we can define:

$$q \cdot \alpha := (qa_1, qa_2, \ldots, qa_n, \ldots) \tag{12.1}$$

which we may also simply denote by $q\alpha$. But one notices that $q\alpha$ is in fact nothing but the product of the constant sequenc ewith term q with the sequence α.Thus, by Proposition 440 we have $q\alpha \in \mathcal{C}$.

Also, if $\alpha = (a_1, a_2, \ldots, a_n, \ldots)$ and $\beta = (b_1, b_2, \ldots, b_n, \ldots)$ are elements of $\mathcal{C}$ we can define their difference $\alpha - \beta$:

$$\alpha - \beta := (a_1 - b_1, a_2 - b_2, \ldots, a_n - b_n, \ldots), \tag{12.2}$$

but we notice that this is just the sequence $\alpha + (-1) \cdot \beta$. Again by Proposition 440 we deduce that $\alpha - \beta \in \mathcal{C}$.

The following proposition is completely straightforward to prove.

Proposition 441 *Let $\alpha, \beta, \gamma \in \mathcal{C}$.Then:*

$$\alpha + \beta = \beta + \alpha \ , \quad \alpha\beta = \beta\alpha \ , \tag{12.3}$$

$$\alpha \pm (\beta \pm \gamma) = (\alpha \pm \beta) \pm \gamma, \quad \alpha\,(\beta\gamma) = (\alpha\beta)\,\gamma, \tag{12.4}$$

and

$$\alpha\,(\beta \pm \gamma) = \alpha\beta \pm \alpha\gamma \tag{12.5}$$

In particular, if $q \in \mathbb{Q}$ we have:

$$q\,(\alpha\beta) = (q\alpha)\,\beta = \alpha\,(q\beta)\,, \quad q\,(\alpha \pm \beta) = q\alpha \pm q\beta \tag{12.6}$$

The proof of the following proposition will be left to the exercises. You can model the proof on the proof of proposition 440.

Proposition 442 *Suppose that $\alpha = (a_1, a_2, \ldots, a_n, \ldots)$ and $\beta = (b_1, b_2, \ldots, b_n, \ldots)$ are sequences of rational numbers that converge in $\mathbb{Q}$ to a and b, respectively. Let q be an arbitrary rational number.*

Then the sequences $\alpha + \beta$, $q\alpha$, and $\alpha \cdot \beta$ are also convergent in $\mathbb{Q}$, with limits $a + b$, qa, and $a \cdot b$, respectively.

Korollar 443 *Suppose that $\alpha = (a_1, a_2, \ldots, a_n, \ldots)$ and $\beta = (b_1, b_2, \ldots, b_n, \ldots)$ are null-sequences of rational numbers. Suppose also that q is a rational number.*

Then the sequences $\alpha + \beta$, $q\alpha$, and $\alpha \cdot \beta$ are also null-sequences.

Bevis. This follows immediatly from the definition of "null-sequence" and Proposition 442 ■

The fact that $q\alpha$ is a null-sequence if α is can be generalized a bit:

Proposition 444 *Suppose that $\alpha = (a_1, a_2, \ldots, a_n, \ldots)$ and $\beta = (b_1, b_2, \ldots, b_n, \ldots)$ are Cauchy sequences of rational numbers and that β is a null-sequence.*

Then $\alpha\beta$ is also a null-sequence.

Bevis. Let ϵ be any positive rational number.

Since α is a Cauchy sequence there is according to Lemma 439 a positive rational number c such that $|a_n| \leq c$ for all $n \in \mathbb{N}$.

Then ϵ/c is also a positive rational number. Since β is a null-sequence there is $N \in \mathbb{N}$ such that $|b_n| < \epsilon/c$ whenever $n \geq N$. But then:

$$|a_n b_n| = |a_n||b_n| < c \cdot \epsilon/c = \epsilon$$

for all $n \geq N$.

We conclude that $\alpha\beta$ is a null-sequence. ■

12.2.2 Definition of the set of real numbers.

Definition 445 *Define a relation $\sim$ between elements of $\mathcal{C}$, i.e., between Cauchy sequences of rational numbers, as follows: If $\alpha, \beta \in \mathcal{C}$ we write $\alpha \sim \beta$ if the Cauchy sequence $\alpha - \beta$ is a null-sequence. In other words:*

$$\alpha \sim \beta \overset{def}{\Leftrightarrow} \alpha - \beta \in \mathcal{N} \ .$$

Proposition 446 *The relation $\sim$ is an equivalence relation.*

Bevis. *Proof that $\sim$ is reflexive*: Let $\alpha = (a_1, a_2, \ldots, a_n, \ldots) \in \mathcal{C}$. Then:

$$\alpha - \alpha = (0, 0, \ldots) \tag{12.7}$$

is the constant sequence with term 0. This is obviously a null-sequence, so we have $\alpha - \alpha \in \mathcal{N}$, i.e. $\alpha \sim \alpha$ by definition.

Proof that $\sim$ is symmetric: Suppose that $\alpha, \beta \in \mathcal{C}$, and suppose that $\alpha \sim \beta$. Then $\alpha - \beta$ is a null-sequence. Then by Proposition 441 we have

$$\beta - \alpha = (-1) \cdot (\alpha - \beta) \tag{12.8}$$

and so by Corollary 443 $\beta - \alpha$ is again a null-sequence. That is $\beta \sim \alpha$.

Proof that $\sim$ is transitive: Let $\alpha, \beta, \gamma \in \mathcal{C}$, and suppose that $\alpha \sim \beta$ and $\beta \sim \gamma$. That is, the sequences $\alpha - \beta$ and $\beta - \gamma$ are both null-sequences. Now, by Proposition 441 we have

$$\alpha - \gamma = (\alpha - \beta) + (\beta - \gamma) \tag{12.9}$$

and so by Corollary 443 the sequence $\alpha - \gamma$ is again a null-sequence. That is, we have $\alpha \sim \gamma$. ■

Definition 447 *The set of real numbers $\mathbb{R}$ is defined as the set of equivalence classes of elements of $\mathcal{C}$ with respect to the equivalence relation $\sim$:*

$$\mathbb{R} := \mathcal{C}/\sim . \tag{12.10}$$

If α is a Cauchy sequence of rational numbers we will write $\overline{\alpha}$ for the equivalence class containing α in $\mathcal{C}/\sim$. Thus, for $\alpha \in \mathcal{C}$ we have an element $\overline{\alpha} \in \mathbb{R}$.

We would now like to define sums and products of elements in $\mathbb{R}$. Since we have already defined sums and products of elements in $\mathcal{C}$ it seems rather clear how to proceed: If $x, y \in \mathbb{R}$ there are elements $\alpha, \beta \in \mathcal{C}$ such that $x = \overline{\alpha}$ and $y = \overline{\beta}$; it seems reasonable to define:

$$x + y := \overline{\alpha + \beta} \tag{12.11}$$

that is, we are proposing to define $+$ on equivalence classes in $\mathcal{C}/\sim$ thus:

$$\overline{\alpha} + \overline{\beta} = \overline{\alpha + \beta}. \tag{12.12}$$

Be sure that you understand fully that this is *not* a statement but a proposed definition.

Before we can formally make this definition we must be certain that it makes sense. What is the problem? The problem is that the above 'definition' relied on a choice: We chose the representatives α and β of the classes x and y in $\mathbb{R} = \mathcal{C}/\sim$. We must convince ourselves that the definition of $x + y$ does not depend on these choices; that is, we must show that the class $\overline{\alpha + \beta}$ does not depend on the choices of representatives α and β for the classes x and y. In mathematical jargon we must show that our proposed $+$ on real numbers is 'well-defined' (i.e., does not depend on any choices made). This is the purpose of the next proposition.

Proposition 448 *Let $x, y \in \mathbb{R}$. Suppose that α_1, $\alpha_2 \in \mathcal{C}$ are both representatives of the class x, and that β_1, $\beta_2 \in \mathcal{C}$ are both representatives of the class y.*

Then:

$$\overline{\alpha_1 + \beta_1} = \overline{\alpha_2 + \beta_2}$$

and

$$\overline{\alpha_1 \beta_1} = \overline{\alpha_2 \beta_2} .$$

Bevis. The statement $\overline{\alpha_1 + \beta_1} = \overline{\alpha_2 + \beta_2}$ means that the 2 Cauchy sequences $\alpha_1 + \beta_1$ and $\alpha_2 + \beta_2$ are in the same class with respect to $\sim$, that is, that

$$\alpha_1 + \beta_1 \sim \alpha_2 + \beta_2 .$$

By definition, this is the statement that the Cauchy sequence $(\alpha_1 + \beta_1) - (\alpha_2 + \beta_2)$ is a null-sequence. Now:

$$(\alpha_1 + \beta_1) - (\alpha_2 + \beta_2) = (\alpha_1 - \alpha_2) + (\beta_1 - \beta_2)$$

by proposition 441.

But $\alpha_1 - \alpha_2$ is a null-sequence since α_1 and α_2 are both representatives for the class x. Similarly, the sequence is $\beta_1 - \beta_2$ is a null-sequence. Now Corollary 443 implies that $(\alpha_1 + \beta_1) - (\alpha_2 + \beta_2)$ is a null-sequence, as required.

We also have:

$$\alpha_1\beta_1 - \alpha_2\beta_2 = \alpha_1(\beta_1 - \beta_2) + \beta_2(\alpha_1 - \alpha_2) \ .$$

Combining Proposition 444 with Corollary 443 we conclude that $\alpha_1\beta_1 - \alpha_2\beta_2$ is a null-sequence. That is:

$$\overline{\alpha_1\beta_1} = \overline{\alpha_2\beta_2} \ .$$

■

We can now formally introduce sums and products of elements in $\mathbb{R}$:

Definition 449 *Let x and y be real numbers. We define the sum $x + y$ and the product xy as follows:*

Let α and β be Cauchy sequences of rational numbers such that:

$$x = \bar{\alpha} \ , \quad y = \bar{\beta} \ .$$

Then we define:

$$x + y := \overline{\alpha + \beta} \quad \textit{and} \quad xy := \overline{\alpha \cdot \beta} \ .$$

By Proposition 448 this is well-defined, i.e., these definitions do not depend on the choices of representatives α and β of x and y, respectively.

Proposition 450 *Consider the map $\phi\colon\ \mathbb{Q} \to \mathbb{R}$ given by:*

$$\phi(a) := \overline{(a, a, \ldots, a \ldots)} \ ,$$

i.e., by mapping a rational number a to the class in $\mathbb{R} = \mathcal{C}/\sim$ containing the constant sequence with term a.

The map ϕ is injective and has the properties:

$$\phi(a + b) = \phi(a) + \phi(b) \quad \textit{and} \quad \phi(ab) = \phi(a)\phi(b) \ .$$

Bevis. Suppose that a and b are rational numbers such that $\phi(a) = \phi(b)$. By definition of ϕ this means that:

$$\overline{(a, a, \ldots, a \ldots)} = \overline{(b, b, \ldots, b \ldots)}$$

i.e., that $(a, a, \ldots, a \ldots) \sim (b, b, \ldots, b \ldots)$. By definition, this means that the sequence

$$(a, a, \ldots, a \ldots) - (b, b, \ldots, b \ldots) = (a - b, a - b, \ldots, a - b \ldots)$$

is a null-sequence. So, the rational number $|a - b|$ is smaller than any positive rational number (*why?*). Since $|a - b| \geq 0$ this can only happen if $|a - b| = 0$. So we must have $a - b = 0$, i.e., $a = b$. We have proved that ϕ is injective.

The proof that ϕ has the other stated properties is left to the exercises. ■

Notice in the Proposition that we have 2 different 'plusses' in the game: In the equality $\phi(a+b) = \phi(a)+\phi(b)$ the $+$ in $a+b$ means addition of the rational numbers a and b whereas the $+$ in $\phi(a) + \phi(b)$ is the addition of real numbers that was defined in Definition 449. Similarly, in the equality $\phi(ab) = \phi(a)\phi(b)$ there are 2 different (and implicit) multiplication signs $\cdot$.

Proposition 450 means that we can view $\mathbb{R}$ as containing a copy of the rational numbers, − namely the set $\phi(\mathbb{Q})$. Abusing notation and denoting this subset $\phi(\mathbb{Q})$ again by the symbol $\mathbb{Q}$ we can then say that $\mathbb{Q}$ is a subset of $\mathbb{R}$. So, if a is a rational number we will view it as an element $a \in \mathbb{R}$. That is, we write a but actually mean $\phi(a)$.

Does this not introduce a potential ambiguity? Namely, if a and b are rational numbers the expression $a + b$ can be interpreted in 2 different ways: Either we can say that $a+b$ is the 'original' sum of a and b as rational numbers. Or, we can say that $a + b$ is the sum of a and b viewed as elements of $\mathbb{R}$. But the proposition states that if we view the sum as an element of $\mathbb{R}$ then it is the same whether we view it in the one or the other way. Thus, the ambiguity is harmless. Similarly, there is in principle an ambiguity in the expression ab, but again by the proposition it does not matter much.

Using notation in a slightly ambiguous way like in the above occurs frequently in mathematics. The reason is not that mathematicians love ambiguity but rather that they like notation as convenient and transparent as possible. Thus, in the above we would like to be able to write $a \in \mathbb{R}$ though we really should be writing $\phi(a) \in \mathbb{R}$.

Ambiguities such as the ones we have just discussed are perfectly all right to use *as long* as one is aware of them, and is able − at any time − to explain the situation without any ambiguities at all.

Now we can prove the first theorem on the properties of the set of real numbers $\mathbb{R}$.

Sætning 451 *The set $\mathbb{R}$ of real numbers equipped with the 2 binary operations $+$ and $\cdot$ defined in Definition 449 is a field.*

That is, the field axioms as given in Definition 361 are satisfied for $\mathbb{R}$ with the binary operations $+$ and $\cdot$.

Bevis. We have a lot of small things to verify as well as one larger one. We will only give the proof of the larger thing and leave the rest to the exercises.

But let us first notice the following: In the field axioms there occur certain elements 0 and 1 of $\mathbb{R}$. We must clarify what we mean by these symbols. But this is easy: We let them mean exactly what they seem to mean, namely the rational numbers 0 and 1, – interpreted, however, as elements of $\mathbb{R}$ as in the discussion following Proposition 450. In other words, the element $0 \in \mathbb{R}$ is actually the class:

$$0 := \overline{(0, 0, \ldots, 0, \ldots)}$$

containing the constant sequence with term 0. Similarly, 1 as an element of $\mathbb{R}$ is the element:

$$1 := \overline{(1, 1, \ldots, 1, \ldots)} \ .$$

It is these elements $0, 1 \in \mathbb{R}$ that we will use in the field axioms. I.e., the statement is now that with these elements $0, 1 \in \mathbb{R}$ we can prove the field axioms as stated in definition 361.

We prove now the least trivial of the properties which is the existence of multiplicative inverses to non-zero elements of $\mathbb{R}$. That is, we must prove that if $x \in \mathbb{R}$ with $x \neq 0$ then there exists $y \in \mathbb{R}$ such that:

$$xy = 1 \ .$$

So let $x \in \mathbb{R}$ be arbitrary but with $x \neq 0$. Let $\alpha = (a_1, a_2, \ldots, a_n, \ldots)$ be a Cauchy sequence representing x, i.e., such that $x = \bar{\alpha}$. Since $x \neq 0$ we have:

$$\overline{(a_1, a_2, \ldots, a_n, \ldots)} = \bar{\alpha} \neq 0 = \overline{(0, 0, \ldots, 0, \ldots)} \ ;$$

by definition, this means that the sequence

$$(a_1, a_2, \ldots, a_n, \ldots) - (0, 0, \ldots, 0, \ldots) = (a_1, a_2, \ldots, a_n, \ldots)$$

is not a null-sequence. By Lemma 439 there is then a positive rational number d and an $M \in \mathbb{N}$ such that:

$$|a_n| \geq d \quad \text{whenever } n \geq M \ .$$

Since d is positive we must then have $a_n \neq 0$ for $n \geq M$. Since a_n is a rational number we can meaningfully speak of the rational number $\frac{1}{a_n}$ whenever $a_n \neq 0$, – and hence in particular for $n \geq M$. We have:

$$\frac{1}{|a_n|} \leq \frac{1}{d} \quad \text{whenever } n \geq M \ . \tag{12.13}$$

Now let us consider the following sequence β of rational numbers:

$$\beta := (\underbrace{0, \ldots, 0}_{M}, \frac{1}{a_{M+1}}, \frac{1}{a_{M+2}}, \ldots) \ .$$

We first claim that β is a Cauchy sequence. To see this let ϵ be an arbitrary positive rational number. Then $d^2\epsilon$ is also a positive rational number. Since $(a_1, a_2, \ldots, a_n, \ldots)$ is a Cauchy sequence there is $N \in \mathbb{N}$ such that:

$$|a_m - a_n| < d^2\epsilon \quad \text{whenever } m, n \geq N \ . \tag{12.14}$$

Let now L be the largest of the 2 numbers M and N. If $m, n \geq L$ we then find:

$$\left|\frac{1}{a_m} - \frac{1}{a_n}\right| = \left|\frac{a_m - a_n}{a_m a_n}\right| = \frac{|a_m - a_n|}{|a_m||a_n|} < d^2\epsilon \cdot \frac{1}{d^2} = \epsilon \ ,$$

because of (12.13) and (12.14). Hence β is a Cauchy sequence.

So we can consider the class $y := \bar{\beta}$ in $\mathbb{R} = \mathcal{C}/\sim$. We claim that $xy = 1$. To prove this, we have to show that $\bar{\alpha}\bar{\beta} = 1 := \overline{(1, 1, \ldots, 1, \ldots)}$ in $\mathbb{R}$, i.e., that

$$\alpha\beta \sim (1, 1, \ldots, 1, \ldots) \ .$$

We compute:

$$\begin{aligned} \alpha\beta &= (a_1, a_2, \ldots, a_n, \ldots) \cdot (\underbrace{0, \ldots, 0}_{M}, \frac{1}{a_{M+1}}, \frac{1}{a_{M+2}}, \ldots) \\ &= (\underbrace{0, \ldots, 0}_{M}, 1, 1, \ldots) \ , \end{aligned}$$

so that

$$\begin{aligned} \alpha\beta - (1, 1, \ldots, 1, \ldots) &= (\underbrace{0, \ldots, 0}_{M}, 1, 1, \ldots) - (1, 1, \ldots, 1, \ldots) \\ &= (\underbrace{-1, \ldots, -1}_{M}, 0, 0, \ldots) \ ; \end{aligned}$$

since this is clearly a null-sequence, the claim follows. ■

12.3 The order in $\mathbb{R}$ and further properties.

This section should be read together with sections 10.2 and 10.3.

12.3.1 Order and absolute value.

We wish to be able to measure sizes of real numbers. The key to doing that is to introduce an order relation. As in section 10.2 the key to define the order relation is to specify a subset $\mathbb{R}_+$ which will play the role of the *positive* real numbers. So let me emphasize that we now have to *define* a certain subset $\mathbb{R}_+$ of $\mathbb{R}$ such that the order axioms O1 and O2 in Definition 383 are satisfied. Once we have $\mathbb{R}_+$ then as described in Definition 389 we get a total order $\leq$ on the set of real numbers.

Definition 452 *Let $x \in \mathbb{R}$ and let $\alpha = (a_1, a_2, \ldots, a_n, \ldots)$ be a Cauchy sequence of rational numbers that represents x in $\mathbb{R}$.*

*We say that x is **positive** if there exists a positive rational number d and an $N \in \mathbb{N}$ such that:*

$$a_n \geq d \quad \text{whenever } n \geq N \ .$$

The set of positive real numbers is denoted by $\mathbb{R}_+$.

Once again, we have a potential problem with the definition: We must show that the definition does not depend on the choice of the representative α. That is, we must show that a given x could not simultaneously be shown to be positive by choosing one representative, and be shown to be not positive by choosing another representative. The following lemma rules out exactly this hypothetical situation and thus makes the definition that we have just made possible.

Lemma 453 *Let $\alpha = (a_1, a_2, \ldots, a_n, \ldots)$ and $\beta = (b_1, b_2, \ldots, b_n, \ldots)$ be Cauchy sequences of rational numbers that both represent a given $x \in \mathbb{R}$. Suppose that there exists a positive rational number d and an $N \in \mathbb{N}$ such that:*

$$a_n \geq d \quad \text{whenever } n \geq N \ .$$

Then there exists a positive rational number d', and a natural number N' such that:

$$b_n \geq d' \quad \text{whenever } n \geq N' \ .$$

Bevis. Since α and β represent the same real number x, we have that $\alpha - \beta$ is a null-sequence. Now, the number $d/2$ is a positive rational number. Since $\alpha - \beta$ is a null-sequence there exists an $M \in \mathbb{N}$ such that:

$$|a_n - b_n| < d/2 \quad \text{whenever } n \geq M \ .$$

In particular, we have $a_n - b_n < d/2$ and so $b_n > a_n - d/2$ for $n \geq M$. If also $n \geq N$ we deduce $b_n > d - d/2 = d/2$.

So we can take $d' := d/2$ and let N' be the largest of the 2 numbers M and N. ■

Now we can prove that the order axioms stated in 383 hold true with the above definition of the subset $\mathbb{R}_+$ of $\mathbb{R}$.

Sætning 454 *Suppose that $x, y \in \mathbb{R}_+$. Then $x + y, xy \in \mathbb{R}_+$.*

If x is any real number then exactly one of the following holds: (i) $x \in \mathbb{R}_+$, (ii) $x = 0$, (iii) $-x \in \mathbb{R}_+$.

Bevis. The first statement will be left to the exercises.

Let us then prove the second statement. So let $x \in \mathbb{R}$ be arbitrary. Let $\alpha = (a_1, a_2, \ldots, a_n, \ldots)$ be a Cauchy sequence of rational numbers that represents x in $\mathbb{R}$.

Suppose that $x \neq 0$. This means that α is not a null-sequence. By Lemma 439 there is then a positive rational number d and an $M \in \mathbb{N}$ such that:

$$|a_n| \geq d \quad \text{whenever } n \geq M \ ,$$

i.e., for any $n \geq M$ we have either $a_n \geq d$ or $a_n \leq -d$.

Now, since the number d is a positive rational number and since α is a Cauchy sequence there is $N \in \mathbb{N}$ such that

$$|a_m - a_n| < d \quad \text{whenever } m, n \geq N \ . \tag{12.15}$$

Let K be the largest of the 2 numbers M and N. Then we have either $a_K \geq d$ or $a_K \leq -d$.

Suppose that $a_K \geq d$ and let $m \geq K$. Then either $a_m \geq d$ or $a_m \leq -d$; but if $a_m \leq -d$ we would obtain $a_m - a_K \leq -2d$ and thus $|a_m - a_K| \geq 2d$ which contradicts (12.15). So, we conclude that $a_m \geq d$ for all $m \geq K$. By definition this means that $x \in \mathbb{R}_+$.

If on the other hand we have $a_K \leq -d$ then we have $-a_K \geq d$. Now, the sequence $-\alpha = (-a_1, -a_2, \ldots, -a_n, \ldots)$ is a representative for $-x$ in $\mathbb{R}$; repeating the previous argument with $-x$ instead of x we can then conclude that $-x \in \mathbb{R}_+$.

So we have now shown that at least 1 of the 3 possibilities *(i), (ii), (iii)* materializes for any $x \in \mathbb{R}$. We must still show that the 3 possibilities are mutually exclusive. So for instance we must show that the real number 0 is not positive.

The verifications are left to the exercises. ■

As in section 10.2 we can now introduce the order $\leq$ in $\mathbb{R}$: If $x, y \in \mathbb{R}$ we write $x \leq y$ if $y - x \in \mathbb{R}_+ \cup \{0\}$. We also write $x < y$ if $y - x \in \mathbb{R}_+$. As usual we write $x \geq y$ if $y \leq x$, and correspondingly with $>$.

As in section 10.2 we can prove that $\leq$ is a total order on the set $\mathbb{R}$ of real numbers.

We can also introduce the absolute value $|x|$ of a real number x: $|x|$ is defined to be x if $x \in \mathbb{R}_+ \cup \{0\}$, that is, if $x \geq 0$. Otherwise, $|x|$ is defined to be $-x$. As remarked in section 10.2 we can easily prove that this absolute value has the properties that we are used to.

Now again we have a certain ambiguity: Suppose that a and b are *rational* numbers. Suppose that $a \leq b$ *as rational numbers*. Is it then still true that $a \leq b$ if we now view a and b as real numbers via the map ϕ of Proposition 450? The answer is yes and this is stated formally in the next proposition. The proofs are easy and are left to the exercises.

Proposition 455 *Consider the injective map $\phi\colon\ \mathbb{Q} \to \mathbb{R}$ of Proposition 450.*

For $a, b \in \mathbb{Q}$ we have:

$$a \leq b \Rightarrow \phi(a) \leq \phi(b) \ ,$$

and

$$|\phi(a)| = \phi(|a|) \ .$$

We will now prove a theorem which says intuitively that the set $\mathbb{Q}$ of rational numbers is 'dense' in $\mathbb{R}$, or, alternatively, that any real number can be approximated arbitrarily close by rational numbers.

Sætning 456 *(1). (The Archimedean property of $\mathbb{R}$). Let ϵ be a positive real number and let x be any real number. Then there is a natural number k such that:*

$$k \cdot \epsilon > x \ .$$

(2). Let ϵ be any positive real number. Then there is a positive rational number q such that:

$$q < \epsilon \ .$$

(3). Let x be a real number, and let ϵ be any positive real number.
Then there is a rational number q such that:

$$|x - q| < \epsilon \ .$$

Bevis. *Proof of (1):* Let $(e_1, e_2, \ldots, e_n, \ldots)$ and $(a_1, a_2, \ldots, a_n, \ldots)$ be Cauchy sequences of rational numbers representing ϵ and x in $\mathbb{R}$, respectively. By Lemma 439 there is a rational number c such that $|a_n| \leq c$ for all $n \in \mathbb{N}$. Thus in particular, $a_n \leq c$ for all $n \in \mathbb{N}$.

Since ϵ is positive we know by definition that there exists a positive rational number d and an $N \in \mathbb{N}$ such that:

$$e_n \geq d \quad \text{whenever } n \geq N \ .$$

Consider the rational number $\frac{c+1}{d}$. There is a natural number k such that $k > \frac{c+1}{d}$. With such a k we have:

$$k \cdot e_n > \frac{c+1}{d} \cdot d = c + 1 \geq a_n + 1$$

for all $n \geq N$. This shows that the terms of the Cauchy sequence

$$(k \cdot e_1 - a_1, k \cdot e_2 - a_2, \ldots, k \cdot e_n - a_n, \ldots)$$

are ≥ 1 for $n \geq N$. By definition this means that the sequence represents a positive real number. But it clearly represents the real number $k \cdot \epsilon - x$. So, $k \cdot \epsilon - x$ is a positive real number, i.e., we have $x < k \cdot \epsilon$.

Proof of (2): The number $1 \in \mathbb{R}$ is positive. The number ϵ is positive and hence in particular not 0. So we can consider its inverse ϵ^{-1}.

By *(1)* we know that there is $k \in \mathbb{N}$ such that $k \cdot 1 > \epsilon^{-1}$. But then $1/k < \epsilon$ and $1/k$ is a positive rational number.

Proof of (3): If $x = 0$ we can choose $q = 0$. (Why?).

Also, suppose that we have proved the statement for x positive. If then x is such that $-x$ is positive, and if q is a rational number such that $|-x-q| < \epsilon$ then we have $|x-(-q)| < \epsilon$. (Why?).

So, by Theorem 454 it is sufficient to prove the theorem for positive x.

Assume then that x is positive. If $x < \epsilon$ we can take $q = 0$, so we may, and will, further assume that $x \geq \epsilon$. According to *(2)* there is a positive rational number d such that $d < \epsilon$. According to *(1)* there is $k \in \mathbb{N}$ such that $k \cdot d > x$. Let $k \in \mathbb{N}$ be smallest with this property. Since $d < \epsilon \leq x$ we must have $k \geq 2$. Since k was chosen smallest possible we must then have $(k-1) \cdot d \leq x$. But then:

$$0 \leq x - (k-1) \cdot d < k \cdot d - (k-1) \cdot d = d < \epsilon \ .$$

So, we can take $q = (k-1) \cdot d$. ■

12.3.2 Sequences in $\mathbb{R}$ and completeness.

Now we are ready to study sequences of real numbers. We introduce the notions of Cauchy sequences of real numbers, of convergent sequences, and of limits of sequences in complete analogy with Definition 436:

Definition 457 *Let $(x_1, x_2, \ldots, x_n, \ldots)$ be a sequence of real numbers, i.e., $x_n \in \mathbb{R}$, for all $n \in \mathbb{N}$.*

*We say that the sequence is a **Cauchy sequence** if for every positive real number ϵ there is (depending on ϵ) an $N \in \mathbb{N}$ such that:*

$$|x_m - x_n| < \epsilon \quad \textit{whenever } m, n \geq N \ .$$

*The sequence $(x_1, x_2, \ldots, x_n, \ldots)$ is said to **converge** to an $a \in \mathbb{R}$ if for every positive real number ϵ there is (depending on ϵ) an $N \in \mathbb{N}$ such that:*

$$|a - x_n| < \epsilon \quad \textit{whenever } n \geq N \ .$$

*If the sequence converges to a we also write $x_n \to a$ for $n \to \infty$, or we say that a **is the limit of the sequence** $(x_1, x_2, \ldots, x_n, \ldots)$*

*We call the sequence **convergent** if it converges to some $a \in \mathbb{R}$.*

The following small lemma is useful. The proof is very easy and is left to the exercises.

Lemma 458 *Let $(x_1, x_2, \ldots, x_n, \ldots)$ be a sequence of real numbers.*

The sequence is a Cauchy sequence if and only if for every positive rational number ϵ there is (depending on ϵ) an $N \in \mathbb{N}$ such that:

$$|x_m - x_n| < \epsilon \quad \textit{whenever } m, n \geq N \ .$$

The sequence converges to the real number a if and only if for every positive rational number ϵ there is (depending on ϵ) an $N \in \mathbb{N}$ such that:

$$|a - x_n| < \epsilon \quad \textit{whenever } n \geq N \ .$$

Now we can say that the real numbers is the set of limits of Cauchy sequences of rational numbers:

Sætning 459 *Every real number is the limit of a Cauchy sequence of rational numbers.*

More precisely, if the real number x is represented by the Cauchy sequence $\alpha = (a_1, a_2, \ldots, a_n, \ldots)$ of rational numbers then α converges to x in $\mathbb{R}$.

Bevis. Let ϵ be any positive rational number. Then $\epsilon/2$ is also a positive rational number. Since α is a Cauchy sequence there is an $N \in \mathbb{N}$ such that

$$|a_m - a_n| < \epsilon/2 \quad \text{whenever } m, n \geq N \ .$$

That is,

$$-\epsilon/2 \leq a_m - a_n \leq \epsilon/2 \quad \text{whenever } m, n \geq N \ ,$$

that we can also write as:

$$\epsilon/2 \leq a_m - a_n + \epsilon \quad \text{and} \quad \epsilon/2 \leq -a_m + a_n + \epsilon \ , \quad \text{whenever } m, n \geq N \ . \tag{12.16}$$

Now let n be any natural number with $n \geq N$ and consider the sequence

$$(a_1 - a_n + \epsilon, a_2 - a_n + \epsilon, \ldots, \epsilon, a_{n+1} - a_n + \epsilon, \ldots) = \alpha - (a_n - \epsilon, a_n - \epsilon, \ldots, a_n - \epsilon, \ldots)$$

which is a Cauchy sequence of rational numbers. Hence it represents a real number. By (12.16) and Definition 452 it represents a positive real number. On the other hand, we see that it represents the real number $x - (a_n - \epsilon)$. So this number is positive, that is, $x - a_n + \epsilon > 0$. We proved this solely under the assumption that $n \geq N$. So we can conclude that

$$-\epsilon < x - a_n \quad \text{for all} \ \ n \geq N \ .$$

Arguing similarly with the inequality $\epsilon/2 \leq -a_m + a_n + \epsilon$ which holds for all $m, n \geq N$ we deduce that

$$x - a_n < \epsilon \quad \text{for all} \ \ n \geq N \ .$$

Combining the 2 inequalities we have that

$$|x - a_n| < \epsilon \quad \text{for all} \ \ n \geq N \ . \tag{12.17}$$

We have shown: Given any positive rational number ϵ there is $N \in \mathbb{N}$ such that (12.17) holds. By Lemma 458 we conclude that the sequence $(a_1, a_2, \ldots, a_n, \ldots)$ converges in $\mathbb{R}$ to the real number x. ■

The relation between Cauchy sequences and convergent sequences is much simpler in $\mathbb{R}$ than in $\mathbb{Q}$:

Sætning 460 *A sequence $(x_1, x_2, \ldots, x_n, \ldots)$ of real numbers is convergent if and only if it is a Cauchy sequence.*

Bevis. That the sequence is Cauchy if it is convergent is proved in exactly the same manner as in the case of sequences of rational numbers that converge in $\mathbb{Q}$.

Suppose that $(x_1, x_2, \ldots, x_n, \ldots)$ is a Cauchy sequence. We must show that it converges to some number $y \in \mathbb{R}$.

Let n be any natural number. Then $1/n$ is a positive real number. So, according to Theorem 456, *(3)*, there exists a rational number q such that $|x_n - q| < 1/n$. We choose for each $n \in \mathbb{N}$ such a rational number q_n. I.e., we choose q_n such that:

$$|x_n - q_n| < 1/n \quad \text{for every} \quad n \in \mathbb{N} \; . \tag{12.18}$$

Consider now the sequence $(q_1, q_2, \ldots, q_n, \ldots)$ of rational numbers.

Suppose that ϵ is any positive real number. Then $\epsilon/2$ is also a positive real number. Since $(x_1, x_2, \ldots, x_n, \ldots)$ is a Cauchy sequence there exists an $N_1 \in \mathbb{N}$ such that:

$$|x_m - x_n| < \epsilon/2 \quad \text{whenever } m, n \geq N_1 \; .$$

Since $1 \in \mathbb{R}$ is positive there is according to Theorem 456, *(1)*, an $N_2 \in \mathbb{N}$ such that $N_2 = N_2 \cdot 1 > 4 \cdot \epsilon^{-1}$. Then, if we denote by N the largest of the 2 numbers N_1 and N_2 we obtain for $m, n \geq N$:

$$\begin{aligned} |q_m - q_n| &= |q_m - x_m + x_m - x_n + x_n - q_n| \\ &\leq |x_m - q_m| + |x_m - x_n| + |x_n - q_n| \\ &\leq \frac{1}{m} + \frac{\epsilon}{2} + \frac{1}{n} \leq \frac{2}{N} + \frac{\epsilon}{2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \; . \end{aligned}$$

In particular, we now know that $(q_1, q_2, \ldots, q_n, \ldots)$ is a Cauchy sequence of rational numbers. Let y denote the class of this sequence in $\mathbb{R}$, i.e.:

$$y := \overline{(q_1, q_2, \ldots, q_n, \ldots)} \; .$$

We claim that the sequence $(x_1, x_2, \ldots, x_n, \ldots)$ converges to y in $\mathbb{R}$.

For let ϵ be an arbitrary positive real number.

We know by Theorem 459 that the sequence $(q_1, q_2, \ldots, q_n, \ldots)$ converges in $\mathbb{R}$ to the number y. Consequently, there is $M \in \mathbb{N}$ such that:

$$|y - q_n| \leq \epsilon/2 \quad \text{whenever } n \geq M \; . \tag{12.19}$$

Again by Theorem 456, *(1)*, there is $K \in \mathbb{N}$ such that $K = K \cdot 1 > 2\epsilon^{-1}$, i.e.,

$$\frac{1}{K} < \epsilon/2 \; . \tag{12.20}$$

Now, if $n \in \mathbb{N}$ is larger than each of the 2 numbers M and K, we obtain by combining (12.18), (12.19), and (12.20) that:

$$|x_n - y| = |x_n - q_n + q_n - y| \leq |x_n - q_n| + |q_n - y| \leq \frac{1}{n} + \epsilon/2 \leq \frac{1}{K} + \epsilon/2 < \epsilon \ ,$$

and the claim follows. ■

12.3.3 The least upper bound property.

We will now prove that the set of real numbers has the 'least upper bound property'(Definition 412). For us, now that we have constructed the system of real numbers, this is not an axiom but actually a theorem.

But let us first consider the precise definitions.

Definition 461 *Let $A \subseteq \mathbb{R}$ be a non-empty subset of the real numbers.*

*We say that A is **bounded from above** if there exist $y \in \mathbb{R}$ such that $x \leq y$ for all $x \in A$.*

*Any y with this property is called an **upper bound of** A.*

*If $A \subseteq \mathbb{R}$ is non-empty and bounded from above we say that a number $y \in \mathbb{R}$ is **a least upper bound of** A if the following conditions hold:*

- *y is an upper bound of A ,*
- *If z is any upper bound of A then $z \geq y$.*

If A is a non-empty subset of $\mathbb{R}$ which has a least upper bound y then y is the only least upper bound of A. In other words, the least upper bound is uniquely determined if it exists. This is easy to see and is left as an exercise.

In this case, we also call y the **supremum** of A and write $y = \mathrm{Sup}A$.

The 'least upper bound property' of the system of real numbers is the content of the following theorem.

Sætning 462 *Every non-empty subset of $\mathbb{R}$ which is bounded from above has a least upper bound.*

Before the proof we prove a useful lemma, the so-called 'squeeze lemma':

Lemma 463 *(The squeeze lemma). Suppose that we have sequences of real numbers $\alpha = (x_1, x_2, \ldots, x_n, \ldots)$ and $\beta = (y_1, y_2, \ldots, y_n, \ldots)$ such that:*

$$x_1 \leq x_2 \leq \ldots \leq x_n \leq \ldots \ ,$$

and

$$y_1 \geq y_2 \geq \ldots \geq y_n \geq \ldots \ ,$$

and such that the sequence $(y_1 - x_1, y_2 - x_2, \ldots, y_n - x_n, \ldots)$ converges to 0.

Then the sequences α and β both converge and have the same limit.

Bevis. We first claim that

$$x_m \leq y_n \quad \text{for all} \quad m, n \in \mathbb{N} \ .$$

For suppose not. Then there would exist a pair of natural numbers k, l such that $y_l < x_k$. But then we would find $y_n \leq y_l < x_k \leq x_n$ whenever n was larger then both of k and l. We would thus have

$$x_n - y_n \geq x_k - y_n \geq x_k - y_l$$

for all such n; since $x_k - y_l$ was a positive number this contradicts the fact that the sequence $(y_1 - x_1, y_2 - x_2, \ldots, y_n - x_n, \ldots)$ converges to 0.

Let now ϵ be an arbitrary positive real number. Since the sequence:

$$(y_1 - x_1, y_2 - x_2, \ldots, y_n - x_n, \ldots)$$

converges to 0 there is an $N \in \mathbb{N}$ such that $|y_n - x_n| < \epsilon$ for all $n \geq N$. In particular, we have $y_n - x_n < \epsilon$ for $n \geq N$.

Suppose now that $m, n \in \mathbb{N}$ with $m, n \geq N$. Using what we showed above we then get that $0 \leq x_m - x_n \leq y_n - x_n < \epsilon$ if $m \geq n$; and if $n \geq m$ we get $0 \leq x_n - x_m \leq y_m - x_m < \epsilon$. I.e., we have:

$$|x_m - x_n| < \epsilon \quad \text{whenever } m, n \geq N \ .$$

We have proved that α is a Cauchy sequence of real numbers. By Theorem 460 we have then that this sequence converges in $\mathbb{R}$. Call its limit a.

In an analogous manner we deduce that the sequence β is convergent in $\mathbb{R}$. Call its limit b.

Since $(y_1 - x_1, y_2 - x_2, \ldots, y_n - x_n, \ldots)$ converges to 0 and since

$$\alpha + (y_1 - x_1, y_2 - x_2, \ldots, y_n - x_n, \ldots) = \beta$$

in the sense of Exercise 476, the result of that exercise shows that we have:

$$a = a + 0 = b \ .$$

■

Proof of Theorem 462. Let $A \subseteq \mathbb{R}$ be a non-empty subset that is bounded from above.

Since A is not empty there is some number $x \in A$. Since A is bounded from above there is some real number y such that $z \leq y$ for all $z \in A$.

In particular, we have $x \leq y$. Put:

$$\eta := y - x \ .$$

We will define by recursion 2 sequences of real numbers

$$\alpha = (x_1, x_2, \ldots, x_n, \ldots) \quad \text{and} \quad \beta = (y_1, y_2, \ldots, y_n, \ldots) \ .$$

We start by defining $x_0 = x_1 := x$ and $y_0 = y_1 := y$. If we have defined x_n, y_n we proceed by defining x_{n+1}, y_{n+1} as follows:

$$(x_{n+1}, y_{n+1}) := \begin{cases} (z, y_n) & \text{if there is some } z \in A \text{ with } \frac{x_n+y_n}{2} < z \leq y_n \\ (x_n, \frac{x_n+y_n}{2}) & \text{if there is no } z \in A \text{ with } \frac{x_n+y_n}{2} < z \leq y_n \ . \end{cases}$$

We claim that the sequences have the following properties:

- $x_n \in A$ and $x_{n-1} \leq x_n$,
- y_n is an upper bound of A and $y_n \leq y_{n-1}$,
- $0 \leq y_n - x_n \leq \frac{1}{2^{n-1}} \cdot \eta$.

This is proved by induction on $n \in \mathbb{N}$. For $n = 1$ the claim is readily checked. We do then the induction step: assume the properties for a certain $n \in \mathbb{N}$. We must show that there are valid for the value $n + 1 \in \mathbb{N}$.

Suppose first that $(x_{n+1}, y_{n+1}) = (z, y_n)$ where $z \in A$ is such that

$$\frac{x_n + y_n}{2} < z \leq y_n \ .$$

Then clearly $x_{n+1} = z \in A$. Also, $(x_n + y_n)/2 \geq x_n$ since $y_n \geq x_n$ by the induction hypothesis. Hence, $x_{n+1} = z \geq x_n$.

We also have $y_{n+1} = y_n$, so certainly $y_{n+1} \leq y_n$ and y_{n+1} is an upper bound of A since y_n is. We have $y_{n+1} - x_{n+1} = y_n - z \geq 0$, and

$$y_{n+1} - x_{n+1} = y_n - z \leq y_n - \frac{x_n + y_n}{2} = \frac{y_n - x_n}{2} \ ;$$

since $y_n - x_n \leq \frac{1}{2^{n-1}} \cdot \eta$ by the induction hypothesis we can also conclude that

$$y_{n+1} - x_{n+1} \leq \frac{1}{2^n} \cdot \eta \ .$$

The other possibility is that $(x_{n+1}, y_{n+1}) = (x_n, \frac{x_n+y_n}{2})$ and there is no $z \in A$ with $\frac{x_n+y_n}{2} < z \leq y_n$. Then clearly $x_{n+1} \geq x_n$ and $x_{n+1} \in A$.

Suppose that y_{n+1} were not an upper bound of A; then there would exist $\xi \in A$ such that $\xi > y_{n+1}$; on the other hand, by the induction hypothesis, y_n is an upper bound of A; so we would certainly have $y_n \geq \xi$; thus, there would be an element ξ of A with $\frac{x_n+y_n}{2} = y_{n+1} < \xi \leq y_n$; we would then have a contradiction, and can conclude that y_{n+1} is in fact an upper bound of A.

We have $y_{n+1} = \frac{x_n+y_n}{2} \leq y_n$ since $x_n \leq y_n$ by the induction hypothesis and furthermore,

$$y_{n+1} - x_{n+1} = \frac{y_n - x_n}{2} \leq \frac{1}{2^n} \cdot \eta \ ,$$

again by using the induction hypothesis.

We have now sequences $\alpha = (x_1, x_2, \ldots, x_n, \ldots)$ and $\beta = (y_1, y_2, \ldots, y_n, \ldots)$ with the above properties. In particular, they satisfy:

$$x_1 \leq x_2 \leq \ldots \leq x_n \leq \ldots \ ,$$

and

$$y_1 \geq y_2 \geq \ldots \geq y_n \geq \cdots \ ,$$

and we have $|y_n - x_n| \leq \frac{1}{2^{n-1}} \cdot \eta$ for all $n \in \mathbb{N}$. This last property implies that the sequence:

$$(y_1 - x_1, y_2 - x_2, \ldots, y_n - x_n, \ldots)$$

converges to 0: For let $\epsilon > 0$ be arbitrary; by Theorem 456, *(1)*, there is an $N \in \mathbb{N}$ such that $N \cdot \epsilon > \eta$. Then $n \cdot \epsilon > \eta$ for all $n \geq N$. Since one easily proves by induction on k that $2^k > k$ for all $k \in \mathbb{N}$ we can conclude that $2^n \cdot \epsilon > \eta$ for all $n \geq N$. But then $|y_n - x_n| < \epsilon$ for all $n \geq N + 1$.

Now we can apply Lemma 463 to conclude that the sequences α and β are both convergent with the same limit. Call this common limit a.

We claim that a is the least upper bound of A. So we must show that a is an upper bound of A, and that it is smallest among all upper bounds of A.

Let $z \in A$ be any element. Suppose that we had $z > a$. Then $\epsilon := z - a$ would be a positive real number. As the sequence β converges to a there would then certainly be some $n \in \mathbb{N}$ such that $|a - y_n| < \epsilon$. But then we would have $y_n - a < \epsilon = z - a$ and hence $y_n < z$; since this contradicts the fact that y_n is an upper bound of A we can conclude that $z \leq a$

Hence, a is an upper bound of A.

Let now b be any upper bound of A. We must prove that $b \geq a$. Suppose to the contrary that we had $b < a$. Then $\epsilon := a - b$ would be a positive real number. As the sequence α converges to a there would then certainly be some $n \in \mathbb{N}$ such that $|a - x_n| < \epsilon$. But then we would have $a - x_n < \epsilon = a - b$ and hence $b < x_n$; since $x_n \in A$ this would contradict the fact that b is an upper bound of A. We can conclude that $a \leq b$.

Hence, a is the least upper bound of A. ■

12.4 Final remarks.

In the previous sections we constructed a set $\mathbb{R}$ with certain properties: $\mathbb{R}$ is a field with a total order $\leq$ for which the following holds: If $x, y, z \in \mathbb{R}$ then:

$$x \leq y \Rightarrow x + z \leq y + z$$

and

$$(0 \leq x \wedge 0 \leq y) \Rightarrow 0 \leq xy \ .$$

A field with a total order satisfying these conditions is called an ordered field. Notice that the statement of Theorem 462 makes sense for any ordered field; if the Theorem is true for a given ordered field one then naturally says that that ordered field has the *least upper bound property*.

So we can say that the set $\mathbb{R}$ that we constructed is an ordered field with the least upper bound property.

One can prove the following *uniqueness theorem*: If L is an ordered field with the least upper bound property then L is isomorphic to $\mathbb{R}$. This means that there is a bijective map $\mathbb{R} \to L$ which respects in the natural way the field operations as well as the order relations.

The uniqueness theorem means that we can speak of the set of real numbers without ambiguity: We may have other constructions of this set (such as the construction via 'Dedekind cuts'), but the theorems we can prove about the system of real numbers are independent of which construction we choose.

So why did we choose the construction of the previous sections via Cauchy sequences of rational numbers?

The answer is that this construction is one that generalizes to other interesting situations. Thus, in analysis you will later learn about so–called 'metric spaces' which are sets equipped with a measure of distance between points. For metric spaces one has a construction called 'completion' and this construction is a direct generalization of the above process going from $\mathbb{Q}$ to $\mathbb{R}$.

The construction is also important in number theory: Looking at the field $\mathbb{Q}$ of rational numbers we could define the distance between 2 numbers a and b as the rational number $|a-b|$. This notion of distance is the so–called archimedean metric (it is thus called because the distance between 0 and a natural number n goes to ∞ with n, – cf. Theorem 456). You can see that this notion of distance is occurring in the definitions of Cauchy sequences, convergence, null–sequences, etc..

But there are other measures of distance between rational numbers than the archimedean metric: In fact, there is a metric called the p-adic metric attached to any prime number p. It is defined as follows: Given $a, b \in Q$ we can write:

$$a - b = p^s \cdot \frac{m}{n}$$

where $s \in \mathbb{Z}$ and $m, n \in \mathbb{Z}$ are both not divisible by p. One then defines:

$$|a - b|_p := e^{-s} \ .$$

Thus, in the p-adic metric the numbers close to 0 are those that are divisible by high powers of p.

Now, if one goes through with the construction via Cauchy sequences, null–sequences, etc., replacing everywhere $|\cdot|$ by $|\cdot|_p$ one obtains from $\mathbb{Q}$ not the real numbers but another field $\mathbb{Q}_p$ called the field of p-adic numbers. It is not an ordered field and so one can not speak about the least upper bound property and so on. But apart from that the field $\mathbb{Q}_p$ has some of properties of $\mathbb{R}$; for instance, sequences in $\mathbb{Q}_p$ converge if and only if they are Cauchy sequences. The fields $\mathbb{Q}_p$ also have some other, more complicated structures.

The point of these new fields $\mathbb{Q}_p$ is that they can be used to build important theories in number theory and arithmetic geometry that can be used to study such diverse topics as solutions to polynomial equations, for instance Fermat's last theorem, algebraic numbers, and cryptographic systems.

So if you want to study any of those topics you have made a good investment of your time learning the material of the previous sections.

12.5 Exercises

Øvelse 464 *Prove that any finite decimal number (i.e., for instance* 2.1415*) is a rational number. (Hint: Multiply the decimal number by a suitable power of* 10*).*

Øvelse 465 *Give several examples of sequences of rational numbers that are not Cauchy sequences. You should prove rigorously that your examples are not Cauchy sequences.*

Øvelse 466 *Convince yourself of the truth of Proposition 441.*

Øvelse 467 *Prove Proposition 442. You can model the proof on the proof of Proposition 440.*

Øvelse 468 *Finish the proof of Proposition 450.*

Øvelse 469 *Finish the proof of Theorem 451.*

Øvelse 470 *Construct via the following steps a Cauchy sequence of rational numbers that is not convergent in* $\mathbb{Q}$*:*

Øvelse 471 *(1). For any* $n \in \mathbb{N}$ *the integer* $2^{2n+5} - 2^{n+4}$ *is positive. Let* k_n *be the smallest integer* ≥ 0 *such that*

$$k_n^2 \geq 2^{2n+5} - 2^{n+4} \ .$$

Thus, we have in fact $k_n \geq 1$*. Show that we must have* $k_n \leq 2^{n+3}$*. (Hint: Otherwise, we would be able to prove* $(k_n - 1)^2 \geq 2^{2n+6} \geq 2^{2n+5} - 2^{n+4}$ *contradicting the minimality of* k_n*).*

Use this to prove that $k_n^2 \leq 2^{2n+5}$. *(Hint: Otherwise, we could again show* $(k_n - 1)^2 \geq 2^{2n+5} - 2^{n+4}$ *and obtain a contradiction).*

(2). For $n \in \mathbb{N}$ *let* a_n *denote the rational number* $k_n/2^{n+2}$*:*

$$a_n := \frac{k_n}{2^{n+2}} \ .$$

Show that:

$$2 - \frac{1}{2^n} \leq a_n^2 \leq 2 \quad \textit{for all } n \in \mathbb{N} \ . \tag{12.21}$$

(3). Show that $\alpha := (a_1, a_2, \ldots, a_n, \ldots)$ *is a Cauchy sequence of rational numbers: First use (12.21) to conclude that* $a_n \geq 1$ *for all* $n \in \mathbb{N}$. *Then use (12.21) to show:*

$$-\frac{1}{2^m} \leq a_m^2 - a_n^2 \leq \frac{1}{2^n}$$

for all $m, n \in \mathbb{N}$. *Write* $a_m^2 - a_n^2 = (a_m + a_n)(a_m - a_n)$ *and conclude that:*

$$-\frac{1}{2^{m+1}} \leq a_m - a_n \leq \frac{1}{2^{n+1}}$$

for all $m, n \in \mathbb{N}$. *Use this to prove the desired.*

(4). Show that α *is not convergent in* $\mathbb{Q}$. *(Hint: Show that if* α *converged to a rational number* q *then we would be able to prove* $q^2 = 2$. *However, it is known that there exists no rational number with this property).*

(5). Consider the real number $x := \bar{\alpha}$. *Show that* $x^2 = 2$ *in* $\mathbb{R}$. *Thus, we have proved the existence of a square root of* 2 *in* $\mathbb{R}$.

Øvelse 472 *Prove the first statement of Theorem 454 and finish the proof of the theorem.*

Øvelse 473 *Prove Proposition 455.*

Øvelse 474 *Show that (2) of Theorem 456 is a special case of (3) of the theorem.*

Øvelse 475 *Prove Lemma 458. (Hint: Use Theorem 456).*

Øvelse 476 *Let* $\alpha = (x_1, x_2, \ldots, x_n, \ldots)$ *and* $\beta = (y_1, y_2, \ldots, y_n, \ldots)$ *be sequences of real numbers and define in a natural way the sequences* $\alpha + \beta$ *and* $\alpha\beta$*:*

$$\alpha + \beta := (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n, \ldots) \ ,$$

$$\alpha\beta := (x_1 y_1, x_2 y_2, \ldots, x_n y_n, \ldots) \ .$$

Suppose that the sequences α *and* β *are both convergent with limits* a *and* b*, respectively.*

Show that the sequences $\alpha + \beta$ *and* $\alpha\beta$ *are then also convergent with limits* $a + b$ *and* ab*, respectively.*

Appendix A

Talsystemets opbygning

I forrige kapitel har vi set hvordan man ud fra de rationale tal kan konstruere de reelle tal. Det er et af de skridt man skal igennem, hvis man vil opbygge talsystemet på mængdelæren. I dette appendix skal vi skitsere trinene i denne opbygning. Hvert trin kræver en del overvejelser, som vi ikke vil gå ind på. Dog er det behandlede trin fra de rationale til de reelle tal det mest komplicerede trin.

Grunden til at man kan være interesseret i at opbygge talsystemet fra mængdelæren er, som det blev beskrevet i de motiverende bemærkninger i forrige kapitel , at man derved har sikkerhed for at man ikke indfører nye inkonsistenser.

1. Fra mængdelæren til $\mathbb{N}_0$.Definer rekursivt:

   (a) $0 = \emptyset$

   (b) $1 = \{\emptyset, \{\emptyset\}\}$

   (c) $2 = \{\emptyset, \{\emptyset, \{\emptyset\}\}\}$ ...

   Indfør da efterfølgerfunktionen på oplagt måde på mængden af alle disse mængder. Så er Peanos aksiomer opfyldt, og man kan indføre regneoperationerne (kompositionsreglerne) + og $\cdot$, samt en ordning.

2. Fra $\mathbb{N}$ til $\mathbb{Z}$. Vi definerer et helt tal x som en ækvivalensklasse af par (m, n) af naturlige tal (tænk på $m - n$), idet vi siger at (m, n) er ækvivalent med (m_1, n_1), hvis $m + n_1 = m_1 + n$. På naturlig vis defineres regneoperationerne, og ordningen, og det vises at man kan indlejre de naturlige tal i $\mathbb{Z}$.

3. Fra $\mathbb{Z}$ til $\mathbb{Q}$. Vi definerer et rationalt tal x som en ækvivalensklasse af par af hele tal (m, n) hvor $n \neq 0$ (tænk på $\frac{m}{n}$), idet vi siger at (m, n) er ækvivalent med (m_1, n_1), hvis $mn_1 = m_1 n$. På naturlig vis defineres regneoperationerne, og ordningen, og det vises at man kan indlejre de hele tal i $\mathbb{Q}$.

4. Fra $\mathbb{Q}$ til $\mathbb{R}$. Som i forrige kapitel.

5. Fra $\mathbb{R}$ til $\mathbb{C}$. Vi definerer et komplekst tal som et par (x, y) af reelle tal (tænk på $x + iy$)

6. Vi kan nu lave en model af den Euklidiske plangeometri inden i $\mathbb{R}^2$, og af rumgeometrien inden i $\mathbb{R}^3$.

Index

ægte delmængde, 58
ækvivalensklasse, 78
ækvivalensrelation, 77, 163
æselbroen, 29

abelsk gruppe, 114
absolut værdi, 136
addition, 165
afbildning, 96, 97
aksiomatisk-deduktive metode, ix
aksiomer, 26
al-kvantor, 6
almindelige begreber, 26
analyse, 31, 34, 41, 147
antisymmetrisk, 76
arkimedisk egenskab, 141
associativ, 112, 126

begrænset, 86, 160
bevis, 13
bevis delt op i tilfælde, 20
bevis ved kontraposition, 16
biimplikation, 2
bijektion, 99
bijektiv, 99
billede, 100
billedmængde, 97

Cauchy følge af reelle tal, 172
Cauchy-følge af rationale tal, 159
cyklisk gruppe, 121

De Morgans love, 68
deduktion, 13
definitioner, 10
definitionsmængde, 97
delmængde, 57
den arkimediske egenskab, 171
differens mængde, 65
dilemma, 13
direkte bevis, 14
disjunkte mængder, 61
disjunktion, 2
disjunktiv syllogisme, 13
distributive lov, 126
distributive love (mængder), 66, 67
dominans af konnektiver, 3

eksistenskvantor, 6
eksistensproblemer, 32
eksistenssætninger, 21
en-entydig, 98
endelig mængde, 108
endeligt legeme, 131
entydighed, 33
entydighedssætninger, 23
Euklids Elementer, 25

fællesmængde, 60
følge, 146
familier af mængder, 62
Fibonaccital, 40
foreningsmængde, 61
formodning, 15
fri variabel, 1
fuldstændig induktion, 47
fuldstændigt ordnet legeme, 140
funktion, 96, 97

geometrisk sted, 59
Goldbachs formodning, 16
graf af funktion, 97
grundmængde, 65
gruppe, 114
gruppeisomorfi, 118, 119

gyldig slutning, 13

højreinvers, 105
Hilbert, 25
hypotese, 2
hypotetisk syllogisme, 13

identiske afbildning, 105
ikke-konstruktive beviser, 22
implikation, 2
indbyrdes primiske tal, 18
indirekte bevis, 20
induktion, 50
Induktionsaksiomet, 48
induktionsantagelse, 44
induktionsbevis, 43, 44
induktionsskridt, 44
induktionsstart, 44
infimum, 90
infimumsegenskaben, 90
injektiv, 98
invers, 105
inverst element, 113, 126
irrationalt tal, 18
irreducibelt polynomium, 53
isomorfi, 118, 119
isomorfi mellem ordnede legemer, 137
isoperimetriske problem, 39

kaniner, 46
kardinalitet, 108
klassedeling, 80
Klein's firegruppe, 117
klemme-lemma, 154
kommutativ, 112, 126
komplementærmængde, 66
kompositionsregel, 111
kompositionstavle, 116
kongruens modulo n, 75
kongruente trekanter, 28
konjunktion, 2
konklusion, 2
konnektiver, 2
konstruktionsproblemer, 37
kontinuerte funktioner, 146
kontinuets kardinalitet, 108
kontraposition, 4, 16
konvergens, 160, 172
kreativitet, 31
kvadratrod 2, 18
kvantorer, 6
kvantorers rækkefølge, 7

læsevejledning, 149
legeme, 125, 166
legemsisomorfi, 129
lige tal, 10
ligningsløsning, 34
logisk ækvivalens, 4
logisk huskeseddel, 4, 9

mægtighed, 108
mængde, 55
mængdealgebra, 66
mængdedifferens, 65
majorant, 86
matematisk struktur, 117
maximalt element, 84
maximum, 84
mellemværdisætningen, 151, 154
Mersenne-primtal, 17
mindste element, 85
minimalt element, 85
minimum, 85
minorant, 86
modeksempel, 15
modstrid, 3
modstridsbevis, 17
modulær aritmetik, 82
modus ponens, 13
modus tollens, 13

nedadtil begrænset, 86
negation, 2
neutralt element, 112, 126
nul-følge, 160
nummerisk værdi, 136

omvendt afbildning, 107
omvendt implikation, 5
opadtil begrænset, 86
ordnet legeme, 132

ordnet par, 69
ordning, 168
ordningsisomorfi, 121
ordningsrelation, 83
originalmængde, 101

på, 98
Peanos aksiomer, 48
polynomium, 53
positiv, 169
postulater, 26
potensmængde, 71
prædikat, 1
primærmængde, 76
primtal, 10, 19
produkt, 165
produktmængde, 69
Pythagoras' læresætning, 30

rationalt tal, 18
reducibelt polynomium, 53
reelle tal, 125, 141, 157, 164
refleksiv, 76
rekursion, 50
rekursionssætningen, 51
relation, 76
restklasser modulo n, 78, 79, 81
Russels paradoks, 71

sætningsanalyse, 41
sammensat afbildning, 103
sammensat tal, 10
sammensatte udsagn, 2
sandhedsmængde, 56
sandhedstabeller, 2
sekundærmængde, 76, 97
simpel induktion, 44
skæringssætningen, 152
største element, 84
studievejledning, 149
supremum, 87
supremumsegenskaben, 88, 175
surjektiv, 98
symmetrisk, 76
syntese, 33, 34

tællelig mængde, 108
tæt, 143
talfølge, 146
tautologi, 3
tomme mængde, 56
total ordning, 83
transitiv, 76
trikotomi, 84

udsagn, 1
uendelig mængde, 108
uforkortelig brøk, 18
ulige tal, 10, 19
undergruppe, 116
uniform kontinuitet, 8
urbillede, 101

værdimængde, 97
venn-diagram, 59
venstreinvers, 105