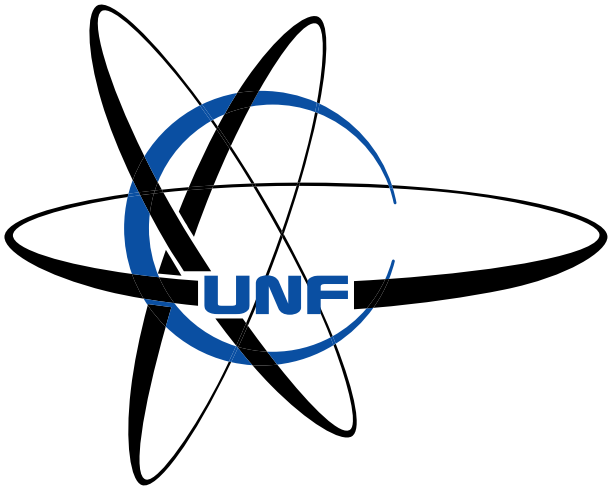


UNF Matematik Camp 2024

Faglige:

Marie Stuhr Kaltoft (ansvarlig)	mark@unf.dk
Rasmus Hauge Hansen (ansvarlig aspirant)	rhh@unf.dk
Anna Mai Østergård	moe@unf.dk
Benjamin Muntz	bmu@unf.dk
Emma Weiss Nielsen (hjælper)	ewn@unf.dk
Erik Søndergård Gimsing	esg@unf.dk
Jonas Nipgaard Dissing (hjælper)	jond@unf.dk
Mathias Weirsøe Klitgaard	mwk@unf.dk
Nanna Wiberg Nielsen	nwn@unf.dk
Rasmus Frigaard Lemvig	rle@unf.dk
Victor Sylvest Blente Heeks	syhe@unf.dk

Ungdommens Naturvidenskabelige Forening

Kompendium til UNF Matematik Camp 2024

Kompendiet er skrevet af Marie Stuhr Kaltoft (ansvarlig), Rasmus Hauge Hansen (ansvarlig aspirant), Anna Mai Østergård, Benjamin Muntz, Emma Weiss Nielsen (hjælper), Erik Søndergård Gimsing, Jonas Nipgaard Dissing (hjælper), Mathias Weirsøe Klitgaard, Nanna Wiberg Nielsen, Rasmus Frigaard Lemvig og Victor Sylvest Blente Heeks.

Layout: Marie Stuhr Kaltoft på forarbejde af Esben Skovhus Ditlefsen, Niels Jakob Søe Loft og Mick Althoff Kristensen.

Opsætning/TEXnisk ansvarlig: Marie Stuhr Kaltoft.

Indhold

Symbolliste

Faglig Introduktion

$\{\cdots\}$	Mængde.
$\emptyset$	Den tomme mængde.
$\subseteq$	Delmængde af.
$\subsetneq$	Ægte delmængde af.
$[a,b]$	Lukket interval fra a til b.
$]a,b[$	Åbent interval fra a til b.
$\in$	Element i.
$\notin$	Ikke element i.
$\mathbb{N}$	Naturlige tal.
$\mathbb{Z}$	Heltal.
$\mathbb{Q}$	Rationale tal.
$\mathbb{R}$	Reelle tal.
$\cup$	Foreningsmængde.
$\cap$	Fællesmængde.
$\setminus$	Differensmaengde.
U	Universalmængden.
$A^{\complement}$	Komplementærmængden af A.
$A \times B$	Det kartesiske produkt af A og B.
$f\colon A \to B$	Funktion/afbildning.
f^{-1}	Invers funktion.
$f(A)$	Billedet af f.
$f^{-1}(S)$	Urbilledet af $S \subseteq B$, hvor $f\colon A \to B$.

$f(x)$	f anvendt på x.
$f^{-1}(x)$	f^{-1} anvendt på x (ikke at forveksle med urbilledet af en mængde).
$x \mapsto y$	x afbilledes over i y.
$n!$	$n \cdot (n-1) \cdots 1$. Udtales "n-fakultet".

$\forall$	For alle.
$\exists$	Der eksisterer.
$\exists!$	Der eksisterer netop én.
$\neg$	Ikke.
$\Rightarrow$	Implikationspil.
$\Leftrightarrow$	Biimplikationspil.
$\lightning$	Modstrid.

Incidensgeometri

$x * y$	x er incident med y.
$U \trianglelefteq \mathcal{L}$	U er et underrum af $\mathcal{L}$.
$\pi(\ell)$	Parallelklassen af ℓ.
$\ell \parallel m$	ℓ er parallel med m.
G_f	Grafen af en funktion f.
$v(\ell)$	Antallet af punkter på linjen ℓ.
$b(p)$	Antallet af linjer gennem punktet p.
$\sum_{i=1}^{n} a_i$	Summen af tallene $a_1, \ldots, a_n$.
$B_d(x,r)$	Bolden af radius r med centrum i x med hensyn til metrikken d.

Talteori

$a \mid b$	a deler b.
$\gcd(n,m)$	Største fælles divisor af n og m.

$a \equiv b \pmod{n}$	a og b er kongruente modulo n.
$[a]_n$	Restklassen for a modulo n (kan også skrives $[a]$).
$\mathbb{Z}/n\mathbb{Z}$	Mængden af restklasser modulo n.
a^{-1}	Invers til $[a]_n$.
$\left(\frac{a}{p}\right)$	Legendresymbolet.
$\varphi(n)$	Eulers φ-funktion af et tal n.
$\lfloor x \rfloor$	Floor funktionen.
S_k	Mængden af summer af k kvadrater.
$\sigma(n)$	Divisorfunktionen.
d	Ofte en divisor.
n	Ofte et naturligt tal.
p	Ofte et primtal.
q	Ofte et primtal eller en kvotient.
r	Ofte en rest.

Ringteori

$-a$	Den additive inverse af a.
a^{-1}	Den multiplikative inverse af a.
$R[x]$	Polynomiumsring med koefficienter i R.
$R_1 \times R_2 \times \cdots \times R_n$	Produktringen af ringene R_1,R_2,..., R_n.
R^n	Produktringen af R med sig selv n gange.
$\varphi : R \to S$	Ringhomomorfi fra R til S.
$i : S \to R$	Inklusionsafbildningen.
$\ker \varphi$	Kernen af φ.
$\simeq$	Isomorfi.
aR	Idealet i R frembragt af a.
R/I	Kvotientringen af R med idealet I.
$\mathbb{C}$	De komplekse tal.
$\pi : R \to R/I$	Den kanoniske afbildning $\pi(a) = \overline{a}$.
$\deg p(x)$	Graden af polynomiet $p(x)$.

$\gcd(a,b)$	Største fælles divisor af a og b.

Gauge Symmetrier

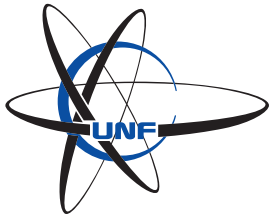
G	En gruppe.
g	Element i en gruppe.
g^{-1}	Det inverse gruppeelement til $g \in G$.
e	Identitetselementet i en gruppe.
λ_g	Gruppevirkningen af et element g.
γ	En-parameter undergruppe af Lie gruppe. Kurve i en mangfoldighed.
$\mathfrak{g}$	Lie algebra.
L	Element i en Lie algebra.
e^L	Eksponentialafbildningen.
$[\cdot,\cdot]$	Lie parentes. Kommutator.
$\mathcal{M}$	En mangfoldighed.
$T_p\mathcal{M}$	Tangentrummet til en mangfoldighed ved et punkt $p \in \mathcal{M}$.
$\mathcal{V}_p$	Tangentvektor i et punkt $p \in \mathcal{M}$.
$\mathcal{K}$	Konfigurationsrum.
$\mathcal{X}$	Positionsrum.
$\mathcal{P}$	Rum for kanonisk momentum.
k	Konfiguration.
x	Position.
p	Kanonisk momentum. Impuls.
t	Tid.
N	Antallet af frihedsgrader.
H	Hamiltonian.
$\mathcal{S}$	Et system.
Ω	Den symplektiske struktur.

d	Det totale differentiale.
∂	Det partielle differentiale.
∇	Gradienten.
D	Det covariante differentiale.
$\vec{V}$	Vektor.
$\mathcal{V}$	Vektorfelt.
$\mathcal{V}_X$	Hamilton vektorfelt af X.
$\Phi_{\mathcal{V}}$	Flowet af et vektorfelt.
$\widehat{\mu}$	Co-moment.
μ	Moment.
$\{\cdot,\cdot\}$	Poisson parentesen.
ξ	Generator af en Lie gruppe.
$\dim(\mathfrak{g})$	Dimensionen af en Lie algebra.
$\mathrm{GL}(N,\mathbb{R})$	Gruppen af invertible $N \times N$ matricer.
$\mathrm{Sp}(2N)$	Den symplektiske gruppe.
$\mathrm{SO}(N)$	Rotationsgruppen.
$\mathrm{U}(N)$	Den unitære gruppe.
$\mathrm{SU}(N)$	Den specielle unitære gruppe.

Græske bogstaver

Her er en tabel over det græske alfabet og bogstavernes navne, så I kan slå det op, hvis I får brug for det. Bogstavet længst til højre er det store bogstav, bogstavet længst til venstre er det lille bogstav, og bogstavet i midten er en alternativ udgave af det lille bogstav. Vær opmærksom på, at nogle bogstaver betyder noget bestemt i nogle afsnit – for eksempel Eulers φ-funktion.

$\alpha\ A$	$\beta\ B$	$\gamma\ \Gamma$	$\delta\ \Delta$	$\epsilon\ \varepsilon\ E$
alfa	beta	gamma	delta	epsilon
$\zeta\ Z$	$\eta\ H$	$\theta\ \vartheta\ \Theta$	$\iota\ I$	$\kappa\ K$
zeta	eta	theta	iota	kappa
$\lambda\ \Lambda$	$\mu\ M$	$\nu\ N$	$\xi\ \Xi$	$o\ O$
lambda	mu	nu	xi	omikron
$\pi\ \Pi$	$\rho\ \varrho\ P$	$\sigma\ \Sigma$	$\tau\ T$	$\upsilon\ \Upsilon$
pi	rho	sigma	tau	upsilon
$\phi\ \varphi\ \Phi$	$\chi\ X$	$\psi\ \Psi$	$\omega\ \Omega$	
phi	chi	psi	omega	



0 Faglig Introduktion

Der er stor forskel på måden, som man lærer matematik i folkeskolen og gymnasiet, og måden, som man lærer matematik på universitet. Hvorimod fokus i folkeskolen og gymnasiet har været at se på matematik i form af ligninger, som skal løses, og talværdier, der skal bestemmes, så er den *rene* matematik meget mere generel og dybdegående. Formålet er at definere og stille spørgsmål til alt, hvad vi tager for givet, når vi arbejder med matematik, som vi kender det. Hvad betyder det at "løse en ligning"? Hvad vil det sige, at noget er en funktion? Og hvad er et "tal" overhovedet? Før vi takler emnerne i dette kompendium, giver det derfor god mening at introducere nogle grundlæggende begreber og noget af tankegangen, som vi benytter til at udføre matematiske beviser.

I dette kapitel vil der af og til komme nye matematiske tegn, som ikke nødvendigvis er gennemgået endnu, men bare rolig, dem vender vi tilbage til.

1 Udsagn

Beviser spiller en helt central rolle i matematikken — de er beviset for, at de matematiske sætninger og formler, vi bruger, overhovedet fungerer! Senere i kapitlet vil I derfor blive præsenteret for forskellige strategier til at bevise påstande, men før vi når dertil, er det vigtigt at forstå, hvad et udsagn eller en påstand er, og hvordan vi kan sammensætte disse.

Definition 1.1 (Udsagn)**.** Et *udsagn* (eller en *påstand*) er en udtalelse, der enten er sand eller falsk.

Vi bruger typisk bogstaverne p, q og r til at repræsentere udsagn.

Eksempel 1.2. Her er nogle eksempler på udsagn:

- $1 + 1 = 2$
- 5 er et lige tal
- Månen er lavet af ost
- Kompendiet har ingen stavefjel

∘

Eksempel 1.3 (*Ikke*-eksempler)**.** Her er nogle eksempler, der *ikke* er udsagn:

- Det er godt, at kompendiet ikke har nogle stavefjel
- Bjørk og Kaare
- Skibidi
- $1 + 1 = x$

Det første punkt på listen er ikke objektivt sandt eller falskt, da det er et holdningsspørgsmål. Derfor er det ikke et udsagn. Selvom vi er glade for både vores koordinatorer og for UNF-danse, så er "Bjørk og Kaare" og "Skibidi" heller ikke udsagn, da de hverken kan være sande eller falske. Det sidste punkt er ikke et udsagn, fordi vi ikke har fastlagt værdien af x og derfor ikke kan afgøre, om $1 + 1 = x$ er sandt eller falskt. ∘

Nogle gange kan det være meningsfuldt at betragte det "modsatte" udsagn.

Definition 1.4 (Negation)**.** Lad p være en udsagn. Da er udsagn $\neg p$ givet ved *negationen* af p. Det betyder, at $\neg p$ er falsk, når p er sand, og sand, når p er falsk.

Bemærkning 1.5. Man kan ofte formulere en negation ved at sætte "ikke" ind i den oprindelige udsagn.

Eksempel 1.6. Lad p være udsagnet "vores chokoladekiks er blevet stjålet af måger". Da er $\neg p$ udsagnet "vores chokoladekiks er ikke blevet stjålet af måger". ∘

Vi kan sammensætte udsagn på flere forskellige måder. Når vi beviser sætninger, kan det for eksempel være nyttigt at sige noget om, at hvis en række antagelser er opfyldt, så gælder der noget sejt.

Definition 1.7 (Implikation)**.** Lad p og q være udsagn. En *implikation* er et udsagn på formen $p \implies q$. Vi kalder " $\implies$ " en *implikationspil*, og implikationen "$p \implies q$" udtales "p *medfører* q".

Hvis q er sand, når p er sand, så er udsagnet $p \implies q$ sand. Altså er $p \implies q$ kun falsk, hvis p er sand og q er falsk. I alle andre tilfælde er udsagnet sandt.

Bemærkning 1.8. Lad $p \implies q$ være et sandt udsagn. Da kan q godt være falsk, hvis p også er falsk.

Eksempel 1.9. Lad p være udsagnet "mågerne har stjålet vores chokoladekiks" og q være udsagnet "vi har ikke flere chokoladekiks". Da betyder implikationen $p \implies q$, at "mågerne har stjålet vores chokoladekiks" $\implies$ "vi har ikke flere chokoladekiks". Vi siger med dagligdagssprog, at "Hvis mågerne har stjålet vores chokoladekiks, så har vi ikke flere chokoladekiks". ∘

Definition 1.10. Lad p og q være udsagn. En *biimplikation* er en påstand på formen $p \iff q$. Vi kalder " $\iff$ " en *biimplikationspil*, og "$p \iff q$" udtales "*p hvis og kun hvis q*".
Påstanden $p \iff q$ er sand, hvis $p \implies q$ og $q \implies p$ begge er sande.

Bemærkning 1.11. Vi beviser ofte sætninger på formen $p \iff q$ ved at bevise $p \implies q$ og $q \implies p$ hver for sig.

Eksempel 1.12. Lad p være udsagnet "n er et lige tal", og lad q være udsagnet "2 går op i n". Da skriver vi biimplikationen $p \iff q$ som: "n er et lige tal" $\iff$ "2 går op i n". Vi siger med dagligdagssprog, at "n er et lige tal, hvis og kun hvis 2 går op i n". ∘

Det kan også være nyttigt at sammensætte udsagn på andre måder. Vi kan for eksempel forestille os en sætning, som kræver, at der skal være opfyldt mere end én antagelse, før den kan bruges. Her skal både antagelse 1, antagelse 2, ... *og* antagelse n være opfyldt.

Definition 1.13 (Konjunktion)**.** Lad p og q være udsagn. Da er $p \wedge q$ (udtales: "p og q") *konjunktionen* af p og q. $p \wedge q$ er sand når både p er sand og q er sand.

Eksempel 1.14. Lad p være udsagnet "storken så mågen," og q være udsagnet "mågen så storken". Da udtales konjunktionen $p \wedge q$ som "storken så mågen, og mågen så storken". Dette er sandt, hvis både storken så mågen, og mågen så storken — altså hvis de havde øjenkontakt. ∘

Definition 1.15 (Disjunktion)**.** Lad p og q være udsagn. Da er $p \vee q$ (udtales "p eller q") *disjunktionen* af p og q. $p \vee q$ er sand, når p eller q (eller begge) er sande.

Eksempel 1.16. Lad p være udsagnet "storken så mågen," og q være udsagnet "mågen så storken". Da udtales disjunktionen $p \vee q$ som "storken så mågen eller mågen så storken". Dette er sandt, hvis mågen så storken, eller storken så mågen. Det er også sandt, hvis de så hinanden. ∘

2 Mængder

En *mængde* er en samling af *elementer*. Vi skriver elementerne i en mængde i tuborg-klammer $\{\ldots\}$. Man kan forstå en mængde som en pose, der indeholder elementerne. Vi kan f.eks. forestille os en pose, der indeholder æbler, pærer og bananer. Som mængde vil vi skrive det som {æble,pære,banan}. Ligesom at det er ligemeget hvilken rækkefølge man har ting i en pose i, er det også ligemeget hvilken rækkefølge man skriver ting op i mængder. Altså er {æble,pære,banan} det samme som {pære,banan,æble}. Man behøver kun at skrive et element op en gang i en mængde. Det betyder at {æble,pære,banan,æble} er det samme som {æble,pære,banan}. Vi kan udtale os om hvorvidt et givent element er i en mængde. Hvis et element er i en mængde bruger vi symbolet $\in$. Hvis det ikke er i mængden bruger vi $\notin$.

$$\text{æble} \in \{\text{æble,pære,banan}\}$$

$$\text{måge} \notin \{\text{æble,pære,banan}\}$$

En mængde er givet ved de elementer der er i den. Hvis der er 2 mængder, der har præcis de samme elementer, så er de den samme mængde. Vi siger at en mængde er entydigt karakteriseret af sine elementer.

Eksempel 2.1. Følgende er eksempler på endelige mængder:

$$\{1,2,4\}$$
$$\{a,b\}$$
$$\{\text{☺},\xi\}$$
$$\{\}$$

Den sidste mængde er lidt speciel, da den netop ikke har nogle elementer. Vi kalder denne mængde for *den tomme mængde*, og vi skriver $\emptyset = \{\}$. ∘

Eksempel 2.2. Følgende er eksempler på uendelige mængder:

$$\begin{aligned} \mathbb{N} &= \{1,2,3,\dots\} && \text{(de naturlige tal)} \\ \mathbb{Z} &= \{\dots,-2,1,0,1,2,\dots\} && \text{(de hele tal)} \\ \mathbb{Q} &= \{\text{alle brøker}\} && \text{(de rationelle tal)} \\ \mathbb{R} &= \{\text{alle decimaltal}\} && \text{(de reele tal)} \end{aligned}$$

∘

Vi kan beskrive mængder ved at skrive alle elementerne i mængden. Vi kan også bruge *mængdebyggernotation*, hvor man skriver en mængde på formen

$$\{ \text{ elementer } \mid \text{ betingelser elementerne skal opfylde } \}.$$

Eksempel 2.3. Lad $a,b \in \mathbb{R}$. Da kan det lukkede interval fra a til b skrives med mængdebyggernotation som:

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

Denne mængde indeholder de reelle tal x, der opfylder, at $a \leq x \leq b$. ∘

Eksempel 2.4. Mængden af kvadrattal $\{1,4,9,16,25,\dots\}$ kan skrives som

$$\{n^2 \mid n \in \mathbb{N}\}.$$

Elementerne i denne mængde er de tal, der kan skrives på formen n^2, hvor n er et naturligt tal. ∘

Eksempel 2.5. Vi kan også skrive de rationelle tal med mængdebyggernotation:

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

Det er dog sjældent, at vi bruger de logiske konnektiver så udførligt, når vi skriver mængder. I praksis skriver vi i stedet, at

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

∘

Lad i det følgende A og B være to arbitrære mængder.

Definition 2.6. A og B er den samme mængde, $A = B$, hvis

$$x \in A \iff x \in B.$$

Det betyder, at mængderne A og B er den samme mængde, hvis og kun hvis ethver element i A også er et element i B, og ethvert element i B også er et element i A.

Definition 2.7 (Delmængde)**.** A er en *delmængde* af B, hvis og kun hvis alle elementer i A også er elementer i B, det vil sige

$$x \in A \implies x \in B.$$

Vi skriver, at $A \subseteq B$.

Bemærkning 2.8. Hvis $A = B$, er det naturligvis sandt, at $A \subseteq B$. Hvis $A \neq B$ skriver vi nogle gange $A \subsetneq B$, og vi siger, at A er en *ægte delmængde* af B

Eksempel 2.9. For alle mængder A gælder der, at $\emptyset \subseteq A$. Kan I regne ud hvorfor? ∘

Eksempel 2.10. I Eksempel 2.2 definerede vi nogle uendelige mængder, som opfylder, at $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$. ∘

Definition 2.11 (Foreningsmængden)**.** *Foreningsmængden* $A \cup B$ er mængden af de elementer, som findes i A og/eller B. Med mængdebyggernotation kan vi skrive, at

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

Definition 2.12 (Fællesmængden)**.** *Fællesmængden* (også kaldet *snittet*) $A \cap B$ er mængden af de elementer, som findes i både A og B. Med mængdebyggernotation kan vi skrive, at

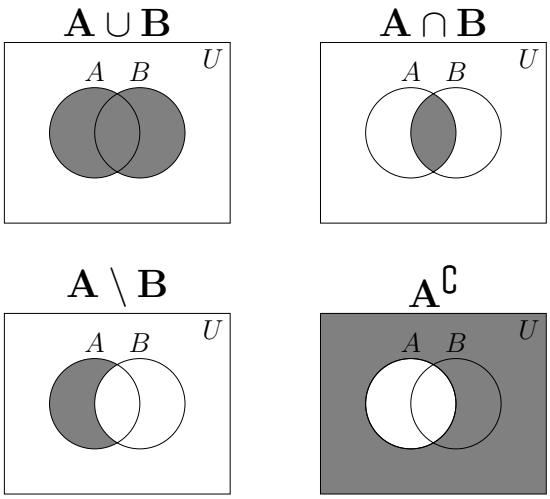
$$A \cap B = \{x \mid (x \in A) \wedge (x \in B)\}.$$

Definition 2.13 (Differensmængden)**.** *Differensmængden* $A \setminus B$ er mængden af de elementer i A, som ikke er elementer i B. Med mængdebyggernotation kan vi skrive, at

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

Definition 2.14 (Komplementærmængden)**.** Lad U være en mængde (ofte kaldet *universalmængden*) og lad $A \subseteq U$. Da er *komplementærmængden* til A (betegnet $A^{\complement}$) defineret som differensmængden mellem U og A. Altså er

$$A^{\complement} = U \setminus A.$$



Figur 0.1: Mængdeoperationer illustreret med Venn-diagrammer.

Eksempel 2.15. Lad $A = \{1, 2, 3, 4\}$, $B = \{3, 4, 5, 6\}$ og $U = \{1, 2, \ldots, 7, 8\}$. Da er

$$A \cup B = \{1,2,3,4,5,6\}$$
$$A \cap B = \{3,4\}$$
$$A \setminus B = \{1,2\}$$
$$B^{\complement} = \{1,2,7,8\}$$

$\circ$

Definition 2.16 (Tupler)**.** En *tuppel* er en ordnet mængde, hvor rækkefølgen af elementerne har betydning.

Bemærkning 2.17. Det er væsentligt, at to-tuplen (a,b) ikke er det samme som mængden $\{a,b\}$. Eksempelvis er $(a,b) \neq (b,a)$, hvis $a \neq b$, hvorimod $\{a,b\} = \{b,a\}$.

Definition 2.18. Lad $a \in A$ og $b \in B$. Det *kartesiske produkt* $A \times B$ er mængden af tupler (a,b), hvor $a \in A$ og $b \in B$. Med mængdebyggernotation kan vi skrive, at

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

Eksempel 2.19. Det (kartesiske) koordinatsystem, som I kender fra gymnasiet, er et kartesisk produkt:

$$\{(x,y) \mid x,y \in \mathbb{R}\} = \mathbb{R} \times \mathbb{R} = \mathbb{R}^2.$$

∘

Kvantorer

Hvis vi vil sige noget om elementerne i mængde, kan vi bruge kvantorer. Dette bør ikke generelt bruges i tekst, men er praktisk, når man for eksempel skriver på en tavle eller tager noter.

Definition 2.20. Symbolet $\forall$ kaldes *alkvantoren* og udtales "for alle".

Eksempel 2.21. Man kan skrive "Alle deltagere på MatCamp er seje" som "$\forall x \in \{$deltagere på MatCamp$\}$ gælder, at x er sej". Man bruger ":" til at skrive "gælder", så det kan skrives endnu kortere som "$\forall x \in \{$deltagere på MatCamp$\}$: x er sej". ∘

Definition 2.22. Symbolet $\exists$ kaldes *eksistenskvantoren* og udtales "eksisterer".

Når vi bruger $\exists$, siger vi, at der findes mindst én. Det betyder ikke, at der kun findes en, der kan sagtens findes flere.

Eksempel 2.23. Man kan skrive "Der er en måge, der vil stjæle UNF's chokoladekiks" som "$\exists x \in \{\text{måger}\} : x$ vil stjæle UNF's chokoladekiks". Det er værd at bemærke, at udsagnet betyder, at der findes mindst én måge, der vil stjæle UNF's chokoladekiks. ∘

Definition 2.24. Man skriver "der findes præcis én" med symbolet $\exists!$ som betyder *entydig eksistens*.

Eksempel 2.25. Man kan skrive "der findes kun en camp, der er den bedste" som "$\exists! x \in \{\text{camps}\} : x$ er bedst". ∘

Hvis der optræder flere kvantorer i et udsagn, så læser vi dem fra venstre mod højre.

Eksempel 2.26. Udsagnet

$$\forall x \in \{\text{chokoladekiks}\}\ \exists! y \in \{\text{mennesker}\} : y \text{ spiser } x$$

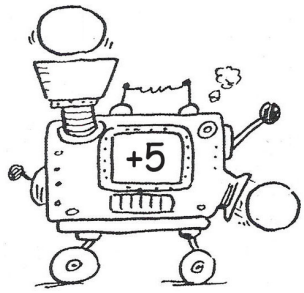
betyder, at for hver chokoladekiks findes (præcis) ét menneske, som spiser chokoladekiksen. Derimod betyder udsagnet

$$\exists! y \in \{\text{mennesker}\}\ \forall x \in \{\text{chokoladekiks}\} : y \text{ spiser } x,$$

at der findes (præcis) ét menneske, som spiser alle chokoladekiks. ∘

3 Funktioner

Inden vi præsenterer en mere stringent definition af funktioner, vil vi forsøge at skabe en intuition for begrebet. I folkeskolen og gymnasiet har I måske arbejdet med funktioner som grafer, der illustrerer en sammenhæng mellem to variable. Mere generelt vil vi dog tænke på en funktion som en "maskine", der sender elementer x fra A over i elementer $f(x)$ i B. Det vil vise sig, at dette funktionsbegreb er en del bredere end det, som I er vant til fra gymnasiet.



Figur 0.2: En funktion er en slags talmaskine.

Definition 3.1 (Funktion)**.** Lad A og B være ikke-tomme mængder. En *funktion* $f\colon A \to B$ (også kaldet en *afbildning*) er en relation mellem to mængder A og B, som til hvert element i A relaterer præcis ét element i B.

Bemærkning 3.2. Når vi har vist, at en relation faktisk er en funktion, så siger vi, at funktionen er *veldefineret*.

Vi bruger flere forskellige notationer til at definere funktioner. Hvis der er tale om en funktion mellem vores sædvanlige talmængder, benytter vi oftest en funktionsforskrift, som I er vant til fra gymnasiet. Hvis $A \subseteq B$ er talmængder, kan vi for eksempel skrive, at $f\colon A \to B$ er givet ved $f(x) = x$.

En anden notation er $a \mapsto b$, hvilket læses som funktionen, der sender a over i b. Denne notation bruges ofte, når vi ikke har brug for

at give et navn til vores funktioner, og/eller A og B er indforståede. Derudover er den også meget brugbar, når funktionen ikke er mellem vores vanlige talmængder.

Eksempel 3.3. Lad $f : \{1,2,3,\mathrm{k}\} \to \mathbb{R}$, hvor k er et bogstav, være givet ved $f(x) = x^2$. Her er f ikke en veldefineret funktion, fordi udtrykket k^2 ikke er defineret. ∘

Eksempel 3.4. Nogle gange kan det være nyttigt at konstruere en funktion, som går fra flere variable ind i en eller flere variable. Lad for eksempel $f\colon \mathbb{R}^2 \to \mathbb{R}$ være givet ved $f(x,y) = \sqrt{x^2 + y^2}$. Da er f en funktion, som returnerer afstanden fra origo til punktet (x,y) i planen. ∘

Definition 3.5 (Definitionsmængde og værdimængde)**.** Lad $f\colon A \to B$ være en funktion. Da kalder vi A *definitionsmængden* (eller domænet), og vi kalder B *værdimængden* (eller kodomænet).

Eksempel 3.6. Lad $f : \{a,b\} \to \{c,d\}$ være givet ved

$$f(x) = \begin{cases} c, & \text{hvis } x = a \\ d, & \text{hvis } x = b \end{cases}$$

Da er f en veldefineret funktion med definitionsmængde $\{a,b\}$ og værdimængde $\{c,d\}$. ∘

Eksempel 3.7. Lad $f : \mathbb{R} \to \mathbb{R}$ være en funktion givet ved $f(x) = x^2$. Da er $\mathbb{R}$ er både definitionsmængde og værdimængde for f. Bemærk at funktionen f er forskellig fra funktionen $g : \mathbb{Z} \to \mathbb{Z}$ givet ved $g(x) = x^2$. ∘

Definition 3.8 (Billede)**.** Lad $f : A \to B$ være en funktion og lad $T \subseteq A$. Vi definerer *billedet* af T (betegnet $f(T)$) som

$$f(T) = \{y \in B \mid \exists t \in T \mid f(t) = y\}.$$

Uformelt siger vi, at billedet af en delmængde er de elementer i værdimængden, som bliver ramt, når man bruger funktionen på delmængden.

Bemærkning 3.9. Selvom funktioner og billeder skrives på samme måde, er de ikke det samme. En vigtig forskel er, at en funktion forholder sig til elementer i mængder, mens billeder forholder sig til selve mængderne.

Eksempel 3.10. Lad $f : \{-2,-1,0,1,2\} \to \{0,1,2,3,4\}$ være givet ved $f(x) = x^2$. Vi kan finde billedet af $\{-2,0,1\}$ ved at undersøge elementerne enkeltvist. Da $f(-2) = 4$, $f(0) = 0$ og $f(1) = 1$, er billedet givet ved

$$f(\{-2,0,1\}) = \{0,1,4\}.$$

∘

Flere funktionsbegreber

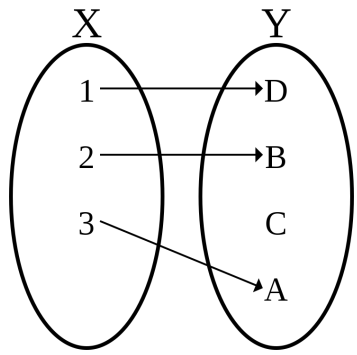
Funktioner kan have forskellige, nyttige egenskaber, og i det følgende vil vi opbygge et sprog herfor.

Definition 3.11 (Injektiv)**.** En funktion $f\colon A \to B$ er *injektiv*, hvis ethvert $b \in B$ bliver ramt af højst ét $a \in A$.

Uformelt siger vi, at en funktion er injektiv, hvis alle elementer i definitionsmængden sendes til forskellige elementer i værdimængden. Eller tilsvarende at $f(x) = f(y)$ medfører $x = y$.

Eksempel 3.12. Lad funktionen $f : \{1,2\} \to \{a\}$ være givet ved $f(x) = a$. Da er f ikke injektiv, fordi både 1 og 2 sendes til a. ∘

Eksempel 3.13. Lad nu $f\colon A \to B$ være en vilkårlig funktion. Hvis vi vil vise, at f er injektiv, så kan vi ofte bruge den samme fremgangsmåde: Vi antager, at der findes $x,x' \in A$, så $f(x) = f(x')$, og viser, at dette medfører $x = x'$.



Figur 0.3: Illustration af injektivitet

Lad os for eksempel vise, at funktionen $f\colon \mathbb{R} \to \mathbb{R}$ givet ved $f(x) = 5x - 6$ er injektiv. Da antager vi, at der findes x, $x' \in \mathbb{R}$, så $f(x) = f(x')$. Altså, at

$$5x - 6 = 5x' - 6.$$

Ved at lægge 6 til på begge sider af lighedstegnet får vi, at

$$5x = 5x',$$

og ved at dele resultatet med 5 får vi, at

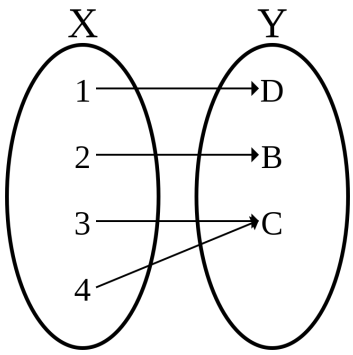
$$x = x',$$

som ønsket. Derfor er f injektiv. ∘

Definition 3.14 (Surjektiv)**.** En funktion $f : A \to B$ er *surjektiv*, hvis der for alle $b \in B$ findes $a \in A$, så $f(a) = b$.

Uformelt siger vi, at en funktion er surjektiv, hvis alle elementer i værdimængden bliver ramt mindst en gang.

Bemærkning 3.15. Man kan sikre, at en funktion er surjektiv ved at vælge værdimængden med omhu.



Figur 0.4: Illustration af surjektivitet

Eksempel 3.16. Lad $f\colon \{-1,\ 0,\ 1\} \to \{0,\ 1,\ 2,\ 3\}$ være givet ved $f(x) = 3x^2$. Da er f ikke surjektiv, fordi nogle af elementerne i værdimængden (f.eks. 2) ikke bliver ramt. Vi kan dog godt konstruere en funktion $g\colon \{-1,0,1\} \to \{0,3\}$ givet ved $g(x) = 3x^2$, som er surjektiv. ∘

Eksempel 3.17. Lad igen $f\colon A \to B$ være en vilkårlig funktion. Hvis vi vil vise, at f er surjektiv, skal vi vise, at der for ethvert $b \in B$ findes et $a \in A$, så $f(a) = b$. I praksis betyder dette ofte, at vi vælger et "smart" $a \in A$, som rammer et givet, men vilkårligt $b \in B$.

Som eksempel kan vi igen se på funktionen $f\colon \mathbb{R} \to \mathbb{R}$ givet ved $f(x) = 5x - 6$. Tag nu et vilkårligt $y \in \mathbb{R}$ og lad $x = (y + 6)/5$. Da er

$$f(x) = f\left(\frac{y+6}{5}\right) = 5\left(\frac{y+6}{5}\right) - 6 = y + 6 - 6 = y$$

Dermed har vi vist, at der for ethvert $y \in \mathbb{R}$ findes et $x \in \mathbb{R}$, så $y = f(x)$. Altså er f surjektiv.

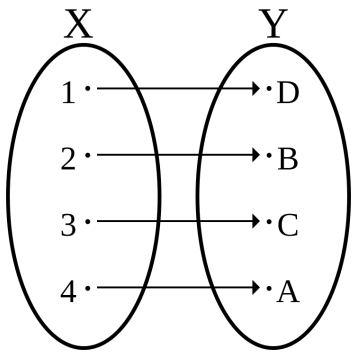
Bemærk at vi har fundet det "smarte" x ved at løse ligningen $y = 5x - 6$. ∘

Sætning 3.18. Lad $f\colon A \to B$ være en funktion. Da er f surjektiv, hvis og kun hvis $f(A) = B$.

Bevis. Se Eksempel 4.7. ■

Definition 3.19 (Bijektiv)**.** En funktion er *bijektiv*, hvis den er både injektiv og surjektiv.

Det betyder, at en funktion $f\colon A \to B$ er bijektiv, hvis og kun hvis ethvert $b \in B$ bliver ramt af præcist ét $a \in A$.



Figur 0.5: Illustration af bijektivitet

Eksempel 3.20. Vi har tidligere vist, at funktionen $f\colon \mathbb{R} \to \mathbb{R}$ givet ved $f(x) = 5x - 6$ både er injektiv og surjektiv. Derfor er f bijektiv. ∘

Det viser sig, at en bijektivitet er en stærk egenskab, som baner vejen for eksistensen af endnu en funktion.

Definition 3.21 (Invers funktion)**.** En funktion $g : B \to A$ er *invers* til $f : A \to B$, hvis det for alle $a \in A$ gælder, at $g(f(a)) = a$. Vi kan da skrive g som f^{-1}.

Når man skal finde en invers funktion kan det altså være en strategi at løse for a i ligningen $f(a) = b$ og lade f^{-1} være givet ved dette nye udtryk.

Sætning 3.22. Lad $f : A \to B$ være en funktion. Der findes en invers funktion $f^{-1} : B \to A$, hvis og kun hvis f er bijektiv.

Bevis. Opgave 5.45. ■

Eksempel 3.23. Da funktionen $f\colon \mathbb{R} \to \mathbb{R}$ givet ved $f(x) = 5x - 6$ er bijektiv, har den en invers. I Eksempel 3.17 så vi, at den inverse funktion $f^{-1}\colon \mathbb{R} \to \mathbb{R}$ er givet ved

$$f^{-1}(y) = (y + 6)/5.$$

∘

Definition 3.24 (Urbillede)**.** Lad $f : A \to B$ være en funktion og lad $S \subseteq B$. Da er *urbilledet* af S under f defineret som

$$f^{-1}(S) = \{a \in A \mid f(a) \in S\}.$$

Uformelt siger vi, at urbilledet er mængden af elementer i definitionsmængden, som rammer S.

Bemærkning 3.25. Selvom notationen for urbilledet minder om notationen for den inverse funktion, er de ikke det samme. Urbilledet adskiller sig ved at være en mængde − ikke en funktion.

Eksempel 3.26. Lad $f\colon \mathbb{R} \to \mathbb{R}$ være givet ved $f(x) = x^2$. Da er urbilledet af $\{4\}$ under f givet ved

$$f^{-1}(\{4\}) = \{x \in \mathbb{R} \mid x^2 = 4\} = \{-2,2\}.$$

∘

4 Beviser

For at vi kan vide, at noget i matematik er sandt, er vi nødt til at bevise det. Det er altså ikke nok bare at kaste dit udsagn op i luften og håbe, at ingen inden for høreafstand brokker sig. Derfor har vi[1] udviklet forskellige matematiske metoder til at bevise et udsagn. Vi kommer til at gennemgå en række forskellige besvimetoder, men først bliver vi dog nødt til at introducere logiske slutninger.

Definition 4.1 (Slutning)**.** Lad p_1,p_2, . . . ,p_n og q være udsagn. Vi siger, at q kan *sluttes* fra p_1,p_2, . . . ,p_n, hvis q er sand, når alle udsagnene p_1,p_2, . . . ,p_n er sande. Altså når p_1,p_2, . . . ,p_n til sammen medfører q.

Et matematisk bevis er faktisk bare en slutning af et udsagn q fra nogle andre udsagn p_1,p_2, . . . ,p_n, hvor alle udsagnene p_1,p_2, . . . ,p_n er tidligere beviste udsagn eller *aksiomer*.

Definition 4.2. Et *aksiom* er et udsagn, som er så banalt, at matematikere ikke kan bevise dem, men må antage det som sandt. Et eksempel på et aksiom er "hvis $a = c$ og $b = c$, så er $a = b$".

For at illustrere de forskellige bevistyper vil vi i løbet af dette afsnit bevise følgende resultat med hver af de viste metoder.

Proposition 4.3. Lad $n \in \mathbb{N}$. Hvis $3n + 2$ er ulige, så er n ulige.

Direkte bevis

De fleste sætninger er på formen $p \implies q$, hvor p er vores antagelser, og q er det, som vi gerne vil vise. Den mest direkte måde at bevise sådanne sætninger på kalder vi selvfølgelig et direkte bevis. Før vi når til definitionen, har vi et eksempel, som forhåbentlig vil hjælpe med at forklare strukturen i et direkte bevis.

[1] Ikke os faglige på Matematik Camp, men nogle kloge matematikere i tidernes morgen.

Eksempel 4.4. Lad p, q og r være tre udsagn. Hvis p medfører q, og q medfører r, så må p også medføre r. Mere formelt kan vi skrive, at

$$((p \Rightarrow q) \text{ og } (q \Rightarrow r)) \Rightarrow (p \Rightarrow r).$$

Hvis det virker lidt abstrakt, kan du erstatte "medfører" med "spiser": Lad os sige, at en ørn spiser en mus, som har spist et stykke ost. Så har ørnen også spist osten. På samme måde virker det med at medføre.

Vi kan udvide det fra at det gælder for 3 udsagn, til at det gælder for en hvilken som helst mængde af udsagn. Hvis vi har en masse udsagn p_1,p_2, ... ,p_n, hvor hvert udsagn medfører det næste, så medfører p_1 altså også p_n. ∘

Et *direkte bevis* er altså et bevis, hvor vi antager p og ved hjælp af en række logiske slutninger kommer frem til q. Formelt bruger vi følgende definition.

Definition 4.5 (Direkte bevis)**.** Et *direkte bevis* udnytter metoden vist i Eksempel 4.4 til at vise, at $p \Rightarrow q$, ved at vise, at

$$(\ldots((p \Rightarrow p_1) \Rightarrow p_2) \Rightarrow \cdots \Rightarrow p_n) \Rightarrow q.$$

Eksempel 4.6 (Bevis Proposition 4.3)**.** Vi vil nu bevise Proposition 4.3 med et direkte bevis.

Lad $n \in \mathbb{N}$. Antag, at $3n + 2$ er ulige. Da findes $k \in \mathbb{N}$, så $3n+2 = 2k+1$, per definition af ulige tal. Da har vi, at $3n = 2k-1$. Altså er $3n$ ulige, så n er også ulige. ∘

Eksempel 4.7 (Bevis Sætning 3.18)**.** Lad $f\colon A \to B$ være en funktion. Da siger Sætning 3.18, at f er surjektiv, hvis og kun hvis $f(A) = B$. Det vil vi nu bevise.

$\underline{\Longleftarrow}$: Ifølge vores antagelse er $f(A) = B$. Det vil sige, at

$$x \in f(A) \iff x \in B.$$

Altså bliver alle elementer i værdimængden ramt af funktionen. Så f er surjektiv.

$\underline{\Longrightarrow:}$ Vi skal vise, at ligheden $f(A) = B$ gælder. Dette gør vi ved at vise begge inklusioner. Lad $b_0 \in B$. Ifølge vores antagelse er f surjektiv, så

$$\forall b \in B,\, \exists a \in A\colon f(a) = b.$$

Altså findes et $a_0 \in A$, så $f(a_0) = b_0$. Da er $b_0 = f(a_0) \in f(A)$. Dermed er $B \subseteq f(A)$. Den anden inklusion gælder, per Definition 3.8.

Da vi har vist begge implikationer, har vi, at f er surjektiv, hvis og kun hvis $f(A) = B$. ∘

Modeksempler

Når vi skal vise, at noget *ikke* gælder, så kan det ofte være oplagt at finde et eksempel, hvor antagelserne er opfyldt, men konklusion ikke er.

Eksempel 4.8. Er det sandt, at hvis $3n + 2$ er ulige, så er n lige?

Fra Proposition 4.3 ved vi, at svaret er nej. Men hvordan kan vi bevise det? Betragt ligningen $3n + 2 = 11$. Vi ved, at 11 er et ulige tal. Ved at løse ligningen finder vi, at $n = 3$, hvilket er et ulige tal. Altså har vi fundet et eksempel, hvor $3n + 2$ er ulige, men n ikke er lige. ∘

Denne type eksempler kalder vi *modeksempler*.

Kontraposition

En anden bevismetode er bevis ved kontraposition. Hvis vi gerne vil vise, at $p \implies q$, så er det tilstrækkeligt at vise, at $\neg q \implies \neg p$. Vi vil nu gennemgå et eksempel på et simpelt kontrapositionsbevis.

Proposition 4.9. Hvis x^2 er ulige, så er x ulige.

Bevis. Vi vil gerne bruge kontraposition. Det kontraponerede udsagn er: Hvis x er lige, så er x^2 lige.

Hvis x er lige, da findes $n \in \mathbb{Z}$, så $x = 2n$. Da vil $x^2 = (2n)^2 = 2(2n^2)$. Altså er x^2 også lige. Per kontraposition har vi nu vist, at hvis x^2 er ulige, så er x ulige. ■

Følgende proposition siger, at bevismetoden faktisk er gyldig.

Proposition 4.10 (Kontraposition)**.** At bevise, at $p \implies q$, er ensbetydende med at bevise, at $\neg q \implies \neg p$.

Beviside. Det eneste tilfælde, hvor $p \implies q$ er falsk, er, når p er sand, og q er falsk (husk definitionen af implikation). I alle andre tilfælde er $p \implies q$ sand.

Tilsvarende er $\neg q \implies \neg p$ falsk, netop når $\neg q$ er sand, og $\neg p$ er falsk. Dette er ensbetydende med, at q er falsk, og p er sand. I alle andre tilfælde er $\neg q \implies \neg p$ sand.

Altså er de to udsagn henholdsvis sande og falske samtidigt. ■

Eksempel 4.11. Vi vil nu bevise Proposition 4.3 ved hjælp af kontraposition.

Det kontraponerede udsagn er: Hvis n er lige, så er $3n + 2$ også lige.

Antag, at $n \in \mathbb{N}$ er lige. Da findes $k \in \mathbb{N}$, så $n = 2k$. Da er $3n + 2 = 6k + 2 = 2(3k + 1)$. Altså er $3n + 2$ lige.

Per kontraposition er Proposition 4.3 bevist. ○

Modstrid

Generelt vil vi ikke have modstrid i vores matematik – vi vil altså ikke sige, at en påstand er sand, hvis det i så fald medfører, at en anden påstand er både sand og falsk på samme tid (mere formelt vil vi ikke have, at hvis udsagn p er sandt, så er både udsagnet q og udsagnet "ikke-q" sande). Det kan vi udnytte til at lave *modstridsbeviser*.

Lad os sige, at vi gerne ville bevise, at $p \implies q$. Negationen af dette ville være, at $p \wedge \neg q$ – altså at både p og "ikke-q" gælder. Så hvis vi kan vise, at $p \wedge \neg q$ medfører noget vrøvl, så ved vi, at $p \implies q$ er sandt.

Bemærkning 4.12. Bemærk, at negationen af et udsagn IKKE er det samme som det kontraponerede udsagn. Hvor det kontraponerede udsagn er ensbetydende med det oprindelige udsagn, så er negationen det "modsatte" af det oprindelige udsagn.

Følgende proposition siger, at bevismetoden faktisk er gyldig.

Proposition 4.13 (Modstrid)**.** Lad p, q og r være udsagn. Hvis $(p \wedge \neg q) \implies (r \wedge \neg r)$ er sandt, så er $p \implies q$ sandt.

Beviside. Bemærk, at $r \wedge \neg r$ aldrig kan være sandt. Så $(p \wedge \neg q) \implies (r \wedge \neg r)$ er kun sandt, hvis $p \wedge \neg q$ er falsk (husk definitionen af implikation). Altså bliver $p \wedge \neg q$ nødt til at være falsk. Da $p \implies q$ er det modsatte (altså negationen) af $p \wedge \neg q$, så er $p \implies q$ sand. ■

Eksempel 4.14. Vi vil nu bevise Proposition 4.3 ved hjælp af modstrid.

Antag, at $3n+2$ er ulige, og at n er lige. Bemærk, at $2 = 2(1)$, så 2 er et lige tal. Da n er lige, så er $3n$ også lige. Hvis vi trækker et lige tal fra et ulige tal, så får vi et ulige tal. Altså er $(3n+2) - 3n = 2$ et ulige tal. Altså er 2 både et lige og et ulige tal, hvilket er en modstrid. ↯ ○

Det er nemt at komme til at lave et såkaldt "falsk modstridsbevis". Man forsøger måske først at lave et modstridsbevis, men ender med at lave et direkte bevis eller et bevis ved kontraposition. De kan altså se ud på to måder:

- Direkte bevis forklædt som modstridsbevis: $p \wedge \neg q \implies q$ (hvor man typisk ikke bruger $\neg q$)
- Kontrapositionsbevis forklædt som modstridsbevis: $p \wedge \neg q \implies \neg p$ (hvor man typisk ikke bruger p)

Eksempel 4.15 (Falsk modstrid)**.** Følgende er et falsk modstridsbevis for Proposition 4.3.

Antag, at n er lige. Altså er $n = 2k$. Det vil sige, at

$$\begin{aligned} 3n + 2 &= 3 \cdot 2 \cdot k + 2 \\ &= 2(3k + 1). \end{aligned}$$

Altså er $3n + 2$ lige. Men vi ved at $3n + 2$ er ulige! Altså kan n ikke være lige.

Beviset er sådan set godt nok indtil sætningen "Men vi ved at $3n + 2$ er ulige!". Hvis vi fjerner de sidste to sætninger, så er det et korrekt bevis ved kontraposition. ∘

Induktion

Eksempel 4.16 (Konstruktion af $\mathbb{N}$)**.** For at få ideen til *induktionsprincippet* vil vi starte med at *konstruere* de naturlige tal $\mathbb{N}$. Dette gør vi i to skridt:

(1) Definér det første naturlige tal til at være tallet 1.

(2) For hvert naturligt tal n definerer vi det næste naturlige tal til at være $n + 1$.

Fra det første skridt har vi, at 1 er det første tal i $\mathbb{N}$. Ved hjælp af det andet skridt kan vi nu lade $n = 1$. Da er $n + 1 = 2$, så 2 er det næste tal i $\mathbb{N}$. Ved at anvende det andet skridt igen og igen kan vi på denne måde konstruere alle de naturlige tal. ∘

Dette giver os ideen til følgende sætning.

Sætning 4.17 (Induktionsprincippet)**.** *Induktionsprincippet* siger, at hvis et udsagn opfylder de to nedenstående punkter, da gælder udsagnet for alle tal i $\mathbb{N}$.

- *Induktionsstarten*: Udsagnet gælder for $n = 1$.
- *Induktionsskridtet*: Hvis udsagnet gælder for n, så gælder det også for $n + 1$.

Før vi gennemgår eksempler, vil vi forsøge at gøre det lidt tydeligere, at induktion i virkeligheden "bare" er en slags algoritme – induktion er trods alt et ret abstrakt koncept.

Bemærkning 4.18. Lad $p(n)$ være et udsagn, som indeholder variablen $n \in \mathbb{N}$ (for eksempel kunne $p(n)$ være udsagnet $n \leq n^2$). Et induktionsbevis er et bevis, hvor vi beviser noget ved at vise, at det gælder for $n = 1$ og, at hvis det gælder for et bestemt n, så gælder det også for det næste. Dette deler vi op i en slags "algoritme":

- *Induktionsstarten* ($n = 1$): Bevis, at udsagnet $p(1)$ er sandt (for at fortsætte eksemplet: $p(1)$ er udsagnet $1 \leq 1^2$).
- *Induktionsskridtet*: Bevis, at hvis $p(m)$ er sand (dette kalder vi *induktionsantagelsen*), så er $p(m+1)$ også sand (for eksempel hvis $p(m)$ er udsagnet $m \leq m^2$, så er $p(m+1)$ udsagnet $m+1 \leq (m+1)^2$).

Når både induktionsstarten og induktionsskridtet er bevist, så siger *induktionsprincippet*, at $p(n)$ er sand for alle $n \in \mathbb{N}$.

På mange måder minder induktionsbeviser om dominobrikker. Forestil dig en række dominobrikker. Så kan man med induktion vise, at hvis man vælter den første brik, så vælter alle brikkerne.

Induktionsstarten er, at man vælter den første dominobrik. Induktionsskridtet er at vise, at hvis en brik vælter, så vælter den næste brik også.

Med dette vist ved vi nu, at alle brikkerne vælter:

- Vi vælter den første (induktionsstarten).
- Fordi den første brik er væltet, så vælter den anden brik også (induktionsskridtet).
- Fordi den anden brik er væltet, så vælter den tredje brik også (induktionsskridtet).
- Og så videre...Altså vælter alle dominobrikkerne.

Eksempel 4.19. Vi vil gerne bevise, at summen af de første n ulige tal er lig n^2. Vi fører beviset ved induktion.

Induktionsstart: Summen af det første ulige tal er lig med 1^2. Det er sandt, da det første ulige tal er 1, og $1 = 1^2$.

Induktionsantagelse: Antag, at summen af de første n ulige tal er lig n^2.

Induktionsskridt: Bemærk, at det n'te ulige tal er $2n-1$ (se Opgave 5.43). Ifølge induktionsantagelsen er $1+3+5+\cdots+(2n-1) = n^2$. Det $(n+1)$'te ulige tal er $2(n+1)-1$, så summen af de første $n+1$ ulige tal er $n^2 + (2(n+1)-1) = n^2 + (2n+1) = (n+1)^2$, hvilket var præcis det, vi ville vise.

Per induktionsprincippet har vi, at summen af de første n ulige tal er lig n^2. ∘

5 Opgaver

Påstande

- **Opgave 5.1:**
  Diskuter følgende spørgsmål med hinanden: Hvad er et udsagn?

- **Opgave 5.2:**
  Hvilke af nedenstående er udsagn:

  1) 2024

  2) Kompendiet er skrevet af en måge.

  3) Jorden er rund.

  4) Chokoladeis er godt.

- **Opgave 5.3:**
  Lad p være udsagnet "hesten spiser et æble" og lad q være udsagnet "det regner".

  1) Hvad betyder $\neg p$?

  2) Hvad betyder $\neg q$?

  3) Hvad betyder $p \implies q$?

  4) Hvad betyder $q \implies q$?

  5) Hvad betyder $p \iff q$?

  6) Hvad betyder $p \wedge q$?

  7) Hvad betyder $p \vee q$?

Mængder

- **Opgave 5.4:**
  Diskuter følgende spørgsmål med hinanden: Hvad er en mængde?

- **Opgave 5.5:**
  Hvilke elementer har følgende mængder?

  1) {pære, banan, æble}

2) $\{\heartsuit, \diamondsuit, \spadesuit, \clubsuit\}$

3) [2,13)

4) $\{1, \{1\}, \{1,\{1\}\}\}$

- **Opgave 5.6:**
  Hvilke af følgende mængder er lig hinanden?

  1) {kuglepen, blyant, kridt}

  2) {kuglepen, kuglepen, kuglepen}

  3) {kridt, blyant, kridt, kuglepen, kridt}

  4) {kuglepen, blyant, kridt,☺}

•• **Opgave 5.7:**
Hvilke af følgende mængder er lig hinanden?

1) $\emptyset$

2) $\{\emptyset\}$

3) $\{\}$

4) $\{\{\},\emptyset\}$

- **Opgave 5.8:**
  Lad $A = \{1,2,3\}$, $B = \{1,2\}$ og $C = \{2,3\}$. Hvilke af følgende er sande?

  1) $A \subseteq B$

  2) $A \subseteq A$

  3) $B \subseteq A$

  4) $B \subseteq C$

  5) $C \subseteq A$

  6) $C \subsetneq A$

  7) $A \subsetneq A$

- **Opgave 5.9:**
  Lad $A = \{2, 3, 5, 7\}$, $B = \{1, 3, 5, 8\}$ og universalmængden $U = \{1, 2, \ldots, 7, 8\}$. Bestem følgende:

1) $A \cup B$

2) $A \cap B$

3) $A \setminus B$

4) $B \setminus A$

5) $U \setminus (A \cup B)$

6) $A^{\complement}$

7) $B^{\complement}$

- **Opgave 5.10:**
  Skriv følgende mængder med mængdebyggernotation:

  1) $\{2,4,6,8,10,\dots\}$

  2) $\{\dots,-5,-3,-1,1,3,5,\dots\}$

  3) $\{1,\frac{1}{2},\frac{1}{4},\dots\}$

- **Opgave 5.11:**
  Lad $a,b \in \mathbb{R}$. Udtryk det åbne interval (a,b) og de halvåbne intervaller $(a,b]$ og $[a,\infty)$ med mængdebyggernotation.

- **Opgave 5.12:**
  Lad $A = \mathbb{N}$, $B = \mathbb{Z}$ og $U = \mathbb{Z}$, hvor U er universalmængden. Bestem følgende:

  1) $A \cup B$

  2) $A \cap B$

  3) $A \setminus B$

  4) $A^{\complement}$

  5) $B^{\complement}$

  6) $U^{\complement}$

- **Opgave 5.13:**
  Lad A, B og C være mængder. Er det altid tilfældet, at $(A \cup B) \cap C = A \cup (B \cap C)$? Hvis nej, så giv et eksempel, hvor det ikke er tilfældet.

•• **Opgave 5.14:**
Opskriv følgende mængder på elementform:

1) $\{x \in \mathbb{R} \mid x^2 = 2\}$

2) $\{x \in \mathbb{Z} \mid x^2 = 2\}$

3) $\{x \in \mathbb{R} \mid x^2 - 6x = 0\}$

4) $\{x \in \mathbb{Z} \mid x^2 < 16\}$

•• **Opgave 5.15:**
Vi siger, at A og B er disjunkte, hvis $A \cap B = \emptyset$. Hvilke af følgende mængder er disjunkte?

1) $\emptyset \cap A$

2) $\mathbb{N} \setminus \mathbb{Z}$

3) $\mathbb{Q} \setminus \mathbb{Z}$

4) $[0,1) \cap (1,2]$

5) $\{(x,y) \in \mathbb{R}^2 \mid x^2 \leq y\} \cap \{(x,y) \in \mathbb{R}^2 \mid -x^2 \geq y\}$

•• **Opgave 5.16:**
Skitser følgende mængder:

1) $[-1,1)$

2) $[-1,1] \cup (2,3)$

3) $[-1,1] \cap (0,2)$

4) $\{(x,y) \in \mathbb{R}^2 \mid -1 \leq x \leq 1,\ 0 < y < 2\}$

5) $\{(x,y) \in \mathbb{R}^2 \mid x \leq y\}$

6) $\{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\} \cup \{(x,y) \in \mathbb{R}^2 \mid x < y\}$

7) $\{(x,y) \in \mathbb{R}^2 \mid x^2 \leq y \leq 1 - x^2\}$

• **Opgave 5.17:**
Lad $A = \{2,3,5\}$ og $B = \{\diamondsuit, \clubsuit\}$. Hvad er $A \times B$?

•• **Opgave 5.18: Tupler**
Hvilke af følgende er sande?

1) $\{a,b\} = \{b,a\}$

2) $(a,b) = (b,a)$

3) $(a,a) = (a)$

4) $(a,a) = (a,a)$

•• **Opgave 5.19:**
Skriv følgende udsagn med kvantorer.

1) Alle måger er onde

2) Der findes mindst en ond måge

3) Der findes en og kun en god måge

4) Der findes mindst et kvadrattal som er et heltal

5) Alle kvadrattal er heltal

6) Alle kvadrattal har et tilhørende naturligt tal som er dets kvadratrod

•• **Opgave 5.20:**
Skriv følgende udsagn med hverdagssprog.

1) $\forall m \in \{2n \mid n \in \mathbb{Z}\} : \frac{m}{2} \in \mathbb{Z}$

2) $\exists n \in \{\text{primtal}\} : n$ er lige

3) $\forall n \in \mathbb{Z} \exists - n \in \mathbb{Z} : n + (-n) = 0$

••• **Opgave 5.21:**
Udregn

$$\left[\frac{-1}{1}, 1 + \frac{1}{1}\right] \cap \left[\frac{-1}{2}, 1 + \frac{1}{2}\right] \cap \left[\frac{-1}{3}, 1 + \frac{1}{3}\right]$$

og gæt på, hvad vi får, hvis vi fortsætter følgen

$$\left[\frac{-1}{1}, 1 + \frac{1}{1}\right] \cap \left[\frac{-1}{2}, 1 + \frac{1}{2}\right] \cap \left[\frac{-1}{3}, 1 + \frac{1}{3}\right] \cap \left[\frac{-1}{4}, 1 + \frac{1}{4}\right] \cap \ldots$$

Funktioner

- **Opgave 5.22:**
  Diskuter følgende spørgsmål med hinanden:
  1) Hvad er en funktion?
  2) Hvornår er en funktion injektiv?
  3) Hvornår er en funktion surjektiv?
  4) Hvornår er en funktion bijektiv?

- **Opgave 5.23:**
  Lad $f : \mathbb{R} \to \mathbb{R}$ være givet ved $f(x) = x^2$.
  1) Find billedet af $[-1,1]$ under f.
  2) Find billedet af $\mathbb{R}$ under f.

•• **Opgave 5.24:**
  Hvorfor er følgende funktioner *ikke* veldefinerede?
  1) $f\colon \mathbb{R} \to \mathbb{N}$ givet ved $f(x) = x$
  2) $f\colon \mathbb{R} \to \mathbb{R}$ givet ved $f(x) = 1/x$

•• **Opgave 5.25:**
  Lad $f\colon \mathbb{N} \to \mathbb{Z}$. Find eksempler, hvor f er
  1) injektiv,
  2) surjektiv,
  3) bijektiv,
  4) ingen af delene.

•• **Opgave 5.26:**
  Lad $f\colon \mathbb{R} \to \mathbb{R}$ være givet ved $f(x) = 3x + 6$. Vis, at f er bijektiv og find inversfunktionen til f.

- **Opgave 5.27:**
  Lad $f\colon \{a\ ,b,\ c\} \to \{1,\ 2,\ 3\}$ være givet ved $f(a) = 1$, $f(b) = 2$, $f(c) = 3$. Er f bijektiv? Hvis ja, find inversfunktionen til f.

•• **Opgave 5.28:**
Lad $f\colon \mathbb{Z} \to \mathbb{Z}$ være givet ved $f(n) = -n$. Vis, at f er bijektiv. Hvad er f^{-1}?

•• **Opgave 5.29:**
Lad $f : \mathbb{R} \to \mathbb{R}$ være givet ved $f(x) = x^2$.

1) Er f injektiv?

2) Er f surjektiv?

3) Hvordan kan vi vælge definitionsmængde og værdimængde, så f er bijektiv?

4) Lad nu $f\colon A \to B$ være givet ved $f(x) = x^2$ og antag, at f er bijektiv. Find den inverse funktion til f.

• **Opgave 5.30:**
Lad $f\colon \mathbb{Z} \setminus \{0\} \to \mathbb{N}$ være givet ved $f(x) = |x|$, hvor $|x|$ er absolutværdien af x, altså tallets afstand fra 0.

1) Er f injektiv?

2) Er f surjektiv?

3) Er $f\colon \mathbb{R} \to \mathbb{R}$ givet ved samme forskrift surjektiv?

•• **Opgave 5.31:**
Hvorfor er følgende funktioner ikke bijektive?

1) $f\colon \mathbb{N} \to \mathbb{Z}$ givet ved $f(x) = x$

2) $f\colon \mathbb{Z} \setminus \{0\} \to \mathbb{N}$ givet ved $f(x) = x^2$

3) $f\colon \mathbb{R} \to [-1{,}1]$ givet $f(x) = \sin x$

4) $f\colon [-\pi/2{,}\pi/2] \to \mathbb{R}$ givet ved $f(x) = \sin x$

5) $f\colon \mathbb{R} \to \mathbb{R}$ givet ved $f(x) = x^3 - x$

•• **Opgave 5.32:**
Lad $f\colon A \to B$ være givet ved følgende forskifter. Bestem (hvis muligt) A og B, så f er henholdsvis injektiv, surjektiv, bijektiv eller ingen af delene.

1) $f(x) = 2^x$

2) $f(x) = \sin(2x)$

•• **Opgave 5.33:**
Lad $f\colon \mathbb{R} \to \mathbb{R}$ være givet ved $f(x) = \cos x$.

1) Bestem billedet af $\mathbb{R}_+$ og $\mathbb{R}_-$ under f.

2) Bestem urbilledet af [0,1].

Beviser

•• **Opgave 5.34:**
Vis, at hvis et heltal $n \in \mathbb{Z}$ er ulige (så der findes et $k \in \mathbb{Z}$ så $n = 2k + 1$), så er n^2 også ulige.

•• **Opgave 5.35:**
Bevis, at hvis to vilkårlige heltal m,$n \in \mathbb{Z}$ er ulige, så er $n + m$ lige.

•• **Opgave 5.36:**
Vis, ved brug af modstrid, at summen af to lige tal er lige.

•• **Opgave 5.37:**
Vis, ved brug af kontraposition, at hvis n er ulige, så er n^2 også ulige.

•• **Opgave 5.38:**
Husk, at $A \subseteq B$, hvis $x \in A \implies x \in B$. Vis, at

1) $\{1,2\} \subseteq \{x^2 + 2x \geq 3\}$

2) $\{x \in \mathbb{Z} \mid 5x + 1 \leq 11\} \subseteq \{x \in \mathbb{Z} \mid x \leq 2\}$

3) $\{x \in \mathbb{R} \mid x^2 < 2\} \subseteq \{x \in \mathbb{R} \mid x < 6\}$

4) $(A \cap B) \subseteq A$

•• **Opgave 5.39:**
Vis, at $A = B \iff (A \subseteq B) \wedge (B \subseteq A)$.

•• **Opgave 5.40:**
Vis, at $A \setminus B = A \cap B^{\complement}$.

••• **Opgave 5.41: De Morgans love**
Vis, at $(A \cup B)^{\complement} = A^{\complement} \cap B^{\complement}$ og $(A \cap B)^{\complement} = A^{\complement} \cup B^{\complement}$.
[Hint: $x \in A^{\complement} \iff x \notin A$.]

•• **Opgave 5.42:**
Brug induktionsprincippet til at vise, at lige tal nummer n er lig $2n$.
[Hint: Prøv først at regne det for $n = 1$, $n = 2$ og $n = 3$ for at få en god ide.]

•• **Opgave 5.43:**
Brug induktionsprincippet til at vise, at ulige tal nummer n er lig $2n - 1$.

•• **Opgave 5.44:**
Brug induktionsprincippet til at vise, at en $(n + 2)$-kant har en vinkelsum på $n \cdot 180°$.
[Hint: Regn først nogle eksempler med små værdier af n.]

••• **Opgave 5.45: Bevis for Sætning 3.22**
Lad $f\colon A \to B$ være en funktion. Vi skal nu vise, at der findes en invers funktion $f^{-1} : B \to A$ hvis og kun hvis f er bijektiv.

1) Antag, at f^{-1} findes og vis, at f er injektiv.

2) Antag, at f^{-1} findes og vis, at f også er surjektiv (og dermed bijektiv).

3) Antag, at f er bijektiv og vis, at f^{-1} findes.

••• **Opgave 5.46:**
Brug induktion til at bevise følgende:

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

2) $2^n < 3^{n-1}$ for alle $n \geq 3$.

3) $2^n < n!$ for alle $n \geq 4$

•• **Opgave 5.47:**
Brug induktionsprincippet til at vise, at $1+2+4+\cdots+2^n = 2^{n+1}-1$.

•• **Opgave 5.48:**
Vis ved induktion, at 5 går op i $11^n - 6$, for alle $n \in \mathbb{N}$.

•• **Opgave 5.49:**
Vis ved induktion, at $1^2 + 2^2 + 3^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$ for alle $n \in \mathbb{N}$

Definition 5.1 (Geometrisk sum)**.** En *geometrisk sum* er en sum på formen $x^0 + x^1 + \cdots + x^n$ for $x \neq 1$ og $n \in \mathbb{N}$.

•• **Opgave 5.50:**
Vis ved brug af induktion, at $x^0 + x^1 + \cdots + x^n = \frac{1-x^{n+1}}{1-x}$.
[Hint: Bemærk, at $x^0 = 1$ for alle tal x.]

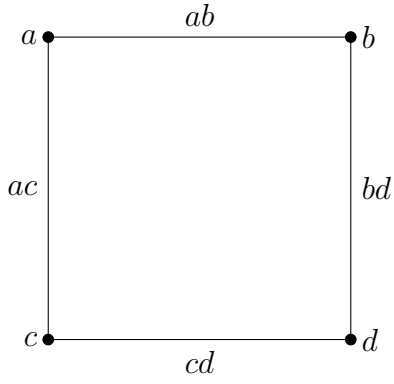
••• **Opgave 5.51:**
Vis *uden* brug af induktion, at $x^0 + x^1 + \cdots + x^n = \frac{1-x^{n+1}}{1-x}$.
[Hint: Prøv at gange summen med $(1 - x)$.]

Incidensgeometri 1

1 Introduktion til incidens

Når man tænker på geometri, er det oftest den Euklidiske af slagsen. Med den slags geometri beskæftiger man sig med mange komplicerede begreber, såsom afstande, vinkler og mellemliggenhed. Dette hjælper os ved at benytte sig af den intuitive forståelse, som mange har om verden omkring sig. Men det tilføjer også en masse kompleksitet. Kompleksitet som vi kan være foruden.

Når vi fjerner sådanne begreber, og kun lader incidens stå tilbage, ender vi med et simpelt udgangspunkt. Objekter og incidens. Incidens beskriver sammentræf mellem to objekter. Det er altså et forhold mellem objekter, og dette forhold gøres mere tydeligt med en ordentlig definition lidt senere. Lad os starte med et eksempel:



Foroven ses noget, som vi ville kalde en firkant. Nogle ville måske

kalde dette et rektangel eller et kvadrat, men sådanne ord giver ikke rigtig mening her, givet vores manglende interesse i længder og vinkler. Men vi kan tydeligt se, at der er linjer og punkter. Disse linjer og punkter er vores objekter. Kigger vi nærmere, ser vi, at nogle af disse linjer og punkter træffer hinanden i firkantens hjørner. Navngivningen af linjerne og punkterne gør det let at beskrive incidensen i denne firkant. Linjerne er navngivet efter de punkter, som de træffer. Det ses, for eksempel, at linjen ab træffer punkterne a og b.

Dette kapitel kommer til at handle om *incidensstrukturer* som firkanten overfor, og nogle af de interessante resultater for og forhold mellem forskellige incidensstrukturer, matematikere har opdaget gennem tiden.

2 Prægeometrier og geometrier

Nu da vi har en intuition for, hvad incidens er for en størrelse, vil vi gå videre til de mere stringente definitioner, der ligger til grund for de emner, vi vil beskæftige os med i resten af forløbet. Vi starter med en definition af de strukturer, vi arbejder med.

Definition 2.1 (Prægeometri)**.** Lad I være en mængde af elementer, som vi kalder *typer*. En prægeometri Γ *over* I er en 3-tupel $(X, *, \mathbf{type})$, som opfylder:

1. X er en ikke-tom mængde. Vi kalder dens elementer for *elementerne* i Γ.

2. $\mathbf{type} : X \to I$ er en surjektiv funktion – den fortæller os *typen* af de enkelte objekter af X.

3. $* : X \times X \to \{\text{Sandt,Falsk}\}$ er en funktion,[1] kaldet *incidensfunktionen*, som opfylder

   - $*$ er refleksiv: Hvis $a \in X$, så har vi, at $*(a,a) = \text{Sandt}$.
   - $*$ er symmetrisk: Hvis $a,b \in X$, så har vi at $*(a,b) = \text{Sandt}$ hvis og kun hvis $*(b,a) = \text{Sandt}$.
   - Hvis $a,b \in X$ opfylder, at $*(a,b) = \text{Sandt}$, og $\mathbf{type}(a) = \mathbf{type}(b)$, så er $a = b$.

Hvis $*(a,b) = \text{Sandt}$ siger vi, at a inciderer med b, og skriver ofte $a * b$.

Definition 2.2. Lad Γ være prægeometri over typemængden I. *Rangen* af Γ er defineret som kardinaliteten af I.

Det er umiddelbart ret meget at få smidt i hovedet på én gang, så lad os tage et eksempel.

[1] Særligt er $*$ det, vi kalder for en binær relation.

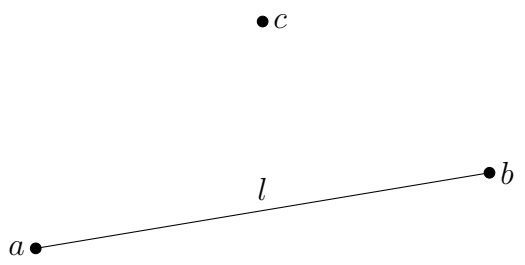
Eksempel 2.3. Lad $I = \{punkt, linje\}$, og definer $X = \{a, b, c, l\}$, med følgende typefunktion

$$\begin{aligned} \mathbf{type}(a) = \mathbf{type}(b) = \mathbf{type}(c) &= linje, \\ \mathbf{type}(l) &= linje, \end{aligned}$$

og hvor incidens er givet således

$$a * l,\ b * l.$$

Det kan vi illustrere:



prægeometrien (X,**type**,Γ) har rang 2. ∘

Vi vil i dette kapitel beskæftige os med prægeometrier, og særligt også *geometrier* – som nok kan læses på vores sprogbrug, så er en prægeometri ikke *helt* en geometri for sig selv. Det er rettere en struktur, som vi sætter ekstra krav på, for at det kan være en geometri.

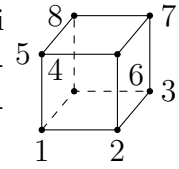
For at kunne definere en geometri, bliver vi nødt til at introducere *flag*.

Definition 2.4 (Flag)**.** Lad $\Gamma = (X, *, \mathbf{type})$ være en prægeometri over typemængden I. Et *flag* F er en mængde af *parvist incidente* elementer af Γ – altså, alle elementer i F inciderer ethvert andet punkt i F.

Definition 2.5 (Maksimalt flag)**.** Lad Γ være en prægeometri over typemængden I.

- Et *maksimalt flag* M er et flag i Γ med $\mathbf{type}(M) = I$.
- Et flag kaldes *co-maksimalt*, hvis der findes et element $x \in X \setminus F$, så $F \cup \{x\}$ er et maksimalt flag.

Eksempel 2.6. Lad dine øjne beskue denne kube. Vi opskriver al den information, vi har om kuben. Vi bruger notation som i figuren til højre til at beskrive punkter, linjer og flader:



Punkter: $\{1, 2, 3, 4, 5, 6, 7, 8\}$,

Linjer: {{1,2}, {1,4},{1,5}, {2,3}, {2,6}, {3,4},
{3,7}, {4,8}, {5,6}, {5,8}, {6,7}, {7,8}},

Flader: {{1,2,3,4}, {1,2,5,6}, {1,4,5,8},
{2,3,6,7}, {3,4,7,8}, {5,6,7,8}}.

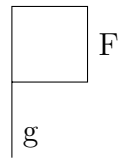
Vi kan på baggrund af dette identificere alle flag, der indeholder punktet 1:

{1}, {1, {1,2}}, {1, {1,4}}, {1, {1,5}},
{1, {1,2,3,4}},{1, {1,2,5,6}}, {1,{1,4,5,8}},
{1,{1,2},{1,2,3,4}},{1,{1,2},{1,2,5,6}}, {1,{1,4},{1,2,3,4}},
{1,{1,4},{1,4,5,8}},{1,{1,5},{1,2,5,6}}, {1,{1,5},{1,4,5,8}}.

Man kan gøre dette for alle punkter, og man ville ende med at have samtlige flag i kuben, men det er faktisk ikke super interessant eller stimulerende at skrive ud, så det gør vi ikke.

Læg dog mærke til, at der iblandt alle flagene foroven er nogle maksimale og co-maksimale flag. For kuben er de maksimale flag alle de flag, der består af et punkt, en linje og en flade. ∘

Hvis en linje g og en flade F inciderer, danner de et flag $\{g, F\}$. Tegnes dette som i figuren til højre, bliver det tydeligt, hvor navnet **flag** kommer fra.



Definition 2.7. En prægeometri Γ over typemængden I kaldes en *geometri*, hvis hvert flag i Γ er indeholdt i et maksimalt flag i Γ.

Eksempel 2.8. Eksempel 2.3 er et eksempel på en prægeometri, som *ikke* er en geometri: c er en del af flaget $\{c\}$, men inciderer ikke med nogen linje, så kan ikke være en del af et maksimalt flag. ∘

Med vores nuværende viden kan vi kigge tilbage på kuben og konkludere, at den udgør en geometri. I kommer senere til at se nogle væsentligt mere spændende geometrier end kuben.

3 Lineære rum

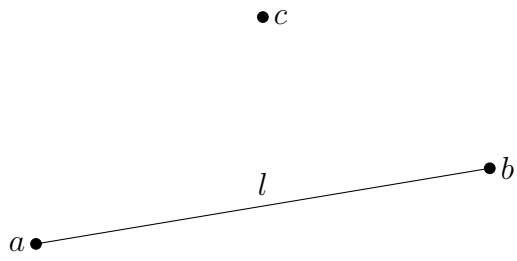
Nu kan vi begynde at konstruere nogle interessante strukturer, inklusiv nogle vi kender fra Euklidisk geometri. Vi skal dog først bruge et par definitioner.

Definition 3.1 (Nær-lineære rum)**.** Et *nær-lineært rum* er en prægeometri $\mathcal{L}$ med rang $= 2$ over typemængden $\{punkt, linje\}$, der opfylder:

NL_1 Enhver linje inciderer med mindst to punkter,

NL_2 Ethvert par af punkter ligger højst på en linje.

Eksempel 3.2. Prægeometrien fra Eksempel 2.3 er et nær-lineært rum.



○

Definition 3.3 (Lineære rum)**.** Et *lineært rum* er en geometri $\mathcal{L}$ med rang $= 2$ over typemængden $\{punkt,linje\}$, der opfylder:

$\mathcal{L}_1$ For hvert par punkter x og y i $\mathcal{L}$ er der præcis én linje, som inciderer med både x og y.

$\mathcal{L}_2$ Hver linje inciderer med mindst to punkter.

Bemærkning 3.4. Ethvert lineært rum er et nær-lineært rum. Dette gælder dog ikke nødvendigvis den anden vej.

Eksempel 3.5. Det nær-lineære rum fra Eksempel 3.2 er *ikke* et lineært rum. ∘

Lad os nu kigge på nogle eksempler på lineære rum. Vi kan starte med et pænt og behageligt lineært rum, som vi alle kender (og har et variende følelsesmæssigt forhold til).

Eksempel 3.6 (Det Euklidiske Plan)**.** Lad E være det *Euklidiske plan*. Vi påstår, at E er et lineært rum. Det betyder, at vi skal tjekke tre ting. Nemlig at E er en geometri, og at begge aksiomer for et lineære rum holder. Det er umiddelbart en stor opgave, men frygt ej! Vi tager den sammen.

E er en geometri!

Først vil vi gerne konstruere E, så vi ved hvad det overhovedet er vi skal vise er en geometri. Lad $I = \{punkt,linje\}$. Så vil vi gerne lave vores mængde af elementer i E, som skal indeholde punkter og linjer. Punkterne skal bestå af alle tupler, der tager indgange i de reelle tal $\mathbb{R}$. Altså,

$$P := \{(x,y) \mid x,y \in \mathbb{R}\}.$$

Men E skal også indeholde linjer, så vi introducerer et nyt begreb. Lad f være en funktion. Så defineres grafen af f ved $G_f := \{(x,f(x)\}$. Det beskriver hver ret linje, som har en funktionsforskrift. Vi kalder mængden af graferne af alle funktioner $f : \mathbb{R} \to \mathbb{R}$ hvor $f(x) = ax + b$ med $a,b \in \mathbb{R}$ for $\mathbf{G}$. Men hvad med de "lodrette" linjer? De kan beskrives ved $\Lambda := \{\{x\} \times \mathbb{R} \mid x \in \mathbb{R}\}$. Så kan vores mængde af linjer skrives som

$$L := \mathbf{G} \cup \Lambda.$$

Definér $X := P \cup L$.

Med alle vores elementer på plads, skal vi nu have en måde at definere incidens mellem elementer. Det gør vi ved en incidens funktion $* : X \times X \to \{\text{Sandt}, \text{Falsk}\}$, hvorom der gælder følgende:

- $l * p$ og $p * l$, hvis og kun hvis $p \in l$.
- $l * l'$, hvis og kun hvis $l = l'$.
- $p * p'$, hvis og kun hvis $p = p'$.

Definitionen garanterer at incidensen foroven er refleksiv, at den er symmetrisk, og at to incidente elementer af samme type altid er ens.

Vi skal også bruge en type funktion. Vi definerer **type** i E ved $\mathbf{type}^{-1}(punkt) = P$ og $\mathbf{type}^{-1}(linje) = L$. Dermed opnås en prægeometri $E := (X, *, \mathbf{type})$.

Det mangler at verificeres, at denne konstruktion faktisk også er en geometri. Det skal derfor vises, at alle flag er indeholdt i et maksimalt flag. Da E er en prægeometri med rang $= 2$, er det nemmest at vise ved at demonstrere, at alle punkter inciderer med en linje, og alle linjer inciderer med et punkt.

Lad $p \in P$. Så er $p = (x,y)$ med $x,y \in \mathbb{R}$. Der findes et $l \in \Lambda$, og særligt $l \in L$ så $(x,y) \in l$. Så alle punkter inciderer med en linje.

Lad $l \in \mathbf{G}$. Så findes der reelle tal a,b, så $l = G_f$ for funktionen $f : \mathbb{R} \to \mathbb{R}$ givet ved $f(x) = ax + b$. Da har vi at $(0,f(0)$ for eksempel er et element i G_f, og derfor i . Dermed inciderer alle linjer med et punkt. Så E er en geometri.

$\mathcal{L}_1$ Vi bliver givet to punkter. Vi skal vise, at der findes en linje, der går gennem begge punkter, og at denne linje er unik.

Vi starter med at vise eksistens - at linjen findes.

Lad $x := (x_1,x_2), y := (y_1,y_2) \in E$ være to forskellige punkter. Da har vi to forskellige tilfælde. I det første gælder $x_1 = y_1$. Da ligger begge punkter på linjen $\{x_1\} \times \mathbb{R} \in L$, siden denne linje går gennem alle punkter med x_1 i første indgang. Særligt går den igennem x og y.

I den anden situation gælder $x_1 \neq y_1$. Da findes en linje funktion $f : \mathbb{R} \to \mathbb{R}$ givet ved $f(t) = at + b$ med $a,b \in \mathbb{R}$ så $f(x_1) = x_2$. Dette ved vi, da vi, givet to reelle tal (kald dem ax_1 og b), ved, at de har en differens. Så vi kan garantere, at G_f går gennem x, ligemeget hvad a er. Vi kan derfor finde et a, der vil medføre, at G_f også går gennem y med dette simple trick. Definér $a := \frac{y_2-x_2}{y_1-x_1}$. Bemærk, at nævneren ikke bliver 0 da $x_1 \neq y_1$. Da $f(x_1) = ax_1 + b = \frac{y_2-x_2}{y_1-x_1}x_1 + b = x_2$ ved vi at $b = x_2 - \frac{y_2-x_2}{y_1-x_1}x_1$. Dette medfører

$$\begin{aligned} f(t) &= \frac{y_2 - x_2}{y_1 - x_1}t + x_2 - \frac{y_2 - x_2}{y_1 - x_1}x_1 \\ \Rightarrow f(y_1) &= \frac{y_2 - x_2}{y_1 - x_1}y_1 + x_2 - \frac{y_2 - x_2}{y_1 - x_1}x_1 \\ &= \frac{y_2 - x_2}{y_1 - x_1}(y_1 - x_1) + x_2 \\ &= y_2 - x_2 + x_2 \\ &= y_2. \end{aligned}$$

Altså, givet to punkter med forskellige værdier i første indgang vil funktionen f altid have en graf, der går gennem begge punkter.

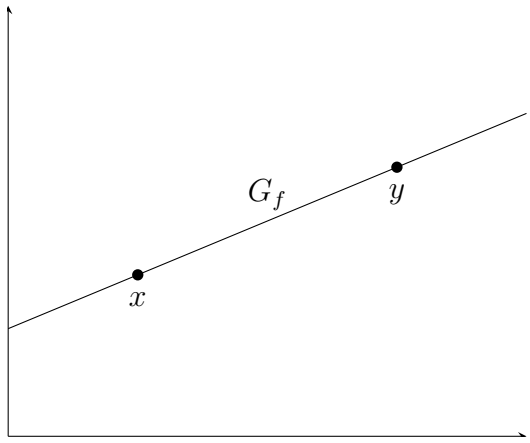
Vi viser nu unikhed.

Vi har nu to forskellige situationer her. Nemlig situationen hvor linjerne er på formen $\{x\} \times \mathbb{R}$ hvor $x \in \mathbb{R}$, og situationen hvor linjerne har en graf. Vi starter med at overveje den første situation.

Lad $x, y \in \{x_1\} \times \mathbb{R}$ være to forskellig punkter. Da gælder $x_1 = y_1$. Det følger, at der ikke eksisterer en funktion $f : \mathbb{R} \to \mathbb{R}$ hvorom det gælder at $x \in G_f$ og $y \in G_f$. Bemærk, at den "lodrette" linje $\{x_1\} \times \mathbb{R}$ er unikt afgjort af x_1.

Nu overvejer vi den anden situation.

Lad x,$y \in G_f$ være forskellige punkter med $f : \mathbb{R} \to \mathbb{R}$, $f(t) = at + b$ med a,$b \in \mathbb{R}$ (se Figur 1.1). Bemærk, at x og y ikke kan ligge på samme "lodrette" linje, da de begge ligger på en graf, og da gælder $x_1 \neq y_1$. Antag nu, at der eksisterer en funktion $g : \mathbb{R} \to \mathbb{R}, g(t) = ct + d$ med c,$d \in \mathbb{R}$, således at x,$y \in G_g$.



Figur 1.1: Illustration af situationen hvor x,$y \in G_f$.

Da gælder at $x_2 = cx_1 + d$ og $y_2 = cy_1 + d$. Særligt gælder $d = x_2 - cx_1$. Da $y \in G_g$ følger det, at $y_2 = cy_1 + x_2 - cx_1$, og ligeledes da $y \in G_f$ har vi $y_2 = ay_1 + x_2 - ax_1$. Dermed

$$\begin{aligned} ay_1 + x_2 - ax_1 &= cy_1 + x_2 - cx_1 \\ \Leftrightarrow a(y_1 - x_1) + x_2 &= c(y_1 - x_1) + x_2 \\ \Leftrightarrow a(y_1 - x_1) &= c(y_1 - x_1) \Leftrightarrow a = c. \end{aligned}$$

Det følger herfra at $d = x_2 - cx_1 = x_2 - ax_1 = b$. Altså, gælder $f(x) = ax + b = cx + d = g(x)$ for alle $x \in \mathbb{R}$ og det medfører $G_f = G_g$.

$\mathcal{L}_2$ Igen har vi to forskellige situationer. I den ene situation har vi at gøre med linjer på formen $\{x\} \times \mathbb{R}$ med $x \in \mathbb{R}$, og i den anden situation gælder det linjer med en graf.

Vi starter med den første situation.

Lad $l \in \Lambda$ med $x \in \mathbb{R}$. Fiksér et $x_0 \in \mathbb{R}$, så $l = \{x_0\} \times \mathbb{R}$. Så inciderer alle punkter med x_0 i første indgang med l. Det er samme antal, som antallet af reelle tal, altså overtælleligt mange, og særligt er der mindst to. Alle de "lodrette" linjer inciderer altså med mindst to punkter.

Nu takler vi den anden situation med linjer, der kan beskrives med grafer.

Lad $f : \mathbb{R} \to \mathbb{R}$ med $f(x) = ax + b$ hvor a,$b \in \mathbb{R}$. Nu prøver vi at finde to punkter, som inciderer med grafen G_f. Bemærk, at punktet $(0,b) \in G_f$, og at $(1, a+b) \in G_f$. Disse er også bestemt elementer i P. Dermed inciderer alle linjer, der kan beskrives med grafer, med mindst to punkter.

Det Euklidiske plan er dermed et lineært rum, da det er en geometri, som opfylder $\mathcal{L}_1$ og $\mathcal{L}_2$. ∘

Bemærk, at den svære del er at konstruere E på en måde, så det bliver en geometri. Ofte kan man regne med at blive givet en geometri.

Da vi nu har set et eksempel på et lineært rum, skal vi også opbygge en smule viden om lineære rum.

Først skal vi lige kende nogle begreber.

For lineære rum, hvis et punkt x inciderer med en linje g, siger vi, at x *ligger på* g, eller at g *går igennem* x.

En linje, som går igennem to punkter x og y, hedder linjen *gennem* x *og* y og skrives xy.

Et punkt, som ligger på to linjer g og h, kaldes et *skæringspunkt* af g og h. Det skrives $g \cap h$. Vi viser nu, at dette punkt er entydigt, hvis det findes.

Sætning 3.7 ([1, Sætning 1.3.1])**.** Lad $\mathcal{L}$ være et lineært rum. For alle par af linjer g og h er der højst ét punkt, som ligger på både g og h.

Bevis. Lad g og h være forskellige linjer i $\mathcal{L}$. Antag for modstrid, at der eksisterer to punkter som inciderer med både g og h. Men så opfylder $\mathcal{L}$ ikke $\mathcal{L}_1$, og er derfor ikke et lineært rum, så vi har en modstrid. ■

Definition 3.8 (Underrum)**.** Lad $\mathcal{L}$ være et nær-lineært rum, og lad U være en mængde af punkter i $\mathcal{L}$. U kaldes et *underrum* af $\mathcal{L}$ hvis for alle par af punkter x og y i U, indeholder U alle punkterne på linjen xy. Det skrives $U \trianglelefteq \mathcal{L}$.

- Hvis $U \trianglelefteq \mathcal{L}$ og $U \neq \mathcal{L}$, siger vi, at U er et *ægte* underrum af $\mathcal{L}$. Vi skriver $U \triangleleft \mathcal{L}$.

- Lad $U \triangleleft \mathcal{L}$. Vi siger, at U er *maksimalt* i $\mathcal{L}$, hvis der ikke findes et ægte underrum $V \triangleleft \mathcal{L}$, så $U \triangleleft V \triangleleft \mathcal{L}$.

Eksempel 3.9. I et lineært rum $\mathcal{L}$ er den tomme mængde, et enkelt punkt, en linje og mængden af alle punkter i $\mathcal{L}$ alle sammen eksempler på underrum. Som en matematikentusiast vil læseren nok gerne se et bevis på det. Så det må læseren jo lave (se Opgave 8.11). ∘

I nær-lineære rum identificeres punktmængden af et underrum med underrummet selv. Hvis alle punkter på en linje g er indeholdt i en punktmængde M, siger vi, at g er indeholdt i M.

Bemærkning 3.10. Da linjer også er underrum, identificeres de også med deres punktmængder. Hvis punktmængden M består af alle punkter der ligger på en linje g, siger vi at punktmængden M er linjen g, skrevet $M = g$. Altså betragter vi i dette kompendie altid linjer som mængder af punkter.

Da vi på denne måde kan identificere linjer med mængden af de punkter, de går igennem, så er det ofte naturligt at *definere* linjerne som mængder af punkter fra starten af, og definere incidens så et punkt inciderer med en linje hvis og kun hvis punktet er et element i linjen.

Sætning 3.11 ([1, Sætning 1.3.2])**.** Lad $\mathcal{L}$ være et lineært rum. Fællesmængden af en arbitrær samling af underrum af $\mathcal{L}$ er et underrum af $\mathcal{L}$.

Bevis. Lad $(U_j)_{j\in J}$ være en samling af underrum i $\mathcal{L}$. Hvis $\bigcap_{j\in J} U_j = \emptyset$ er vi færdige, fordi den tomme mængde altid er et underrum i $\mathcal{L}$. Hvis fællesmængden er et enkelt punkt, er vi igen færdige, da enkelte punkter er underrum i $\mathcal{L}$.

Lad derfor x og y være to punkter indeholdt i $\bigcap_{j\in J} U_j$. Det betyder at x,$y \in U_j$ for alle $j \in J$. Siden U_j er et underrum for alle $j \in J$, er linjen xy indeholdt i U_j for alle $j \in J$. Hvorfor er vi så færdige? ■

Definition 3.12. Lad $\mathcal{L}$ være et lineært rum.

(a) En punktmængde M af $\mathcal{L}$ er *kollineær*, hvis alle punkter i M ligger på en fælles linje.

(b) Lad M være en punktmængde af $\mathcal{L}$, og lad

$$\langle M\rangle := \bigcap_{j\in J} U_j$$

hvor $(U_j)_{j\in J}$ er mængden af underrum af $\mathcal{L}$ der indeholder M. $\langle M\rangle$ er det mindste underrum af $\mathcal{L}$, der indeholder M og kaldes *spannet* af M.

(c) Lad x, y og z være tre ikke-kollineære punkter i $\mathcal{L}$. Underrummet $\langle x, y, z\rangle$ kaldes et plan i $\mathcal{L}$.

Lemma 3.13 ([1, Lemma 1.3.3])**.** Lad $\mathcal{L}$ være et lineært rum, og lad U være et underrum af $\mathcal{L}$. Lad derudover M være en punktmængde indeholdt i U. Da er $\langle M\rangle$ indeholdt i U.

Bevis. Bemærk, at U selv er et underrum. Særligt er U et underrum, der indeholder M. Det følger dermed af Definition 3.12(b), at $\langle M\rangle$ er indeholdt i U. ■

Definition 3.14. Lad $\mathcal{L}$ være et lineært rum. Et maksimalt ægte underrum af $\mathcal{L}$ kaldes et *hyperplan* af $\mathcal{L}$.

Sætning 3.15 ([1, Sætning 1.3.4])**.** Lad $\mathcal{L}$ være et lineært rum, og lad H være et ægte underrum af $\mathcal{L}$, hvor hver linje i $\mathcal{L}$ har mindst ét punkt til fælles med H. Da er H et hyperplan i $\mathcal{L}$.

Bevis. Vi skal vise at H er maksimal, altså at der ikke findes et ægte underrum U af $\mathcal{L}$, så H er et ægte underrum af U.

Antag for modstrid, at U er et ægte underrum af $\mathcal{L}$, således H er et ægte underrum af U. Så har vi $H \lhd U \lhd \mathcal{L}$. Lad x være et punkt i U, som ligger udenfor H, og lad y være et punkt i $\mathcal{L}$, der ligger udenfor U. Så findes en linje g som går gennem x og y pr. Definition 3.3 ($\mathcal{L}_2$). Pr. antagelse mødes linjen g og H i et punkt $z \neq x$. Det følger fra dette og Definition 3.8, at linjen $g = xz$ er indeholdt i U. Da $y \in g$, modstrider dette at $y \notin U$. Altså findes der ikke et U, så $H \lhd U \lhd \mathcal{L}$. ■

4 Projektive og affine planer

Nu vil vi gerne i gang med at tilføje noget mere struktur til vores lineære rum. Vi kommer til at se på nogle særtilfælde af lineære rum, nemlig projektive og affine planer. Formålet er at opbygge noget viden om disse strukturer, og sidst i afsnittet ser vi den tætte kobling mellem dem.

Men vi starter, som ethvert andet godt afsnit i en matematiktekst, med en definition.

Definition 4.1. Et *projektivt plan* er en geometri $\mathbf{P}$ med rang $= 2$ over typemængden $\{punkt, linje\}$, som opfylder følgende.

PP_1 Hvert par af forskellige punkter i $\mathbf{P}$ inciderer med præcis én linje.

PP_2 Hvert par af forskellige linjer har netop ét skæringspunkt.

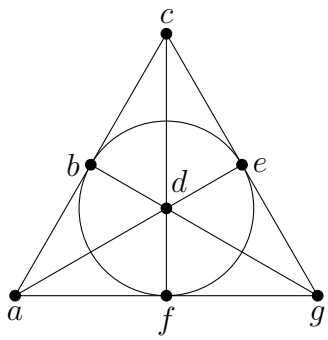
PP_3 Hver linje inciderer med mindst tre forskellige punkter. Der er mindst to forskellige linjer.

Bemærkning 4.2. Hvis $\mathbf{P}$ er et projektivt plan, så er $\mathbf{P}$ et lineært rum. Det betyder, at alt der gælder for lineære rum også gælder for projektive planer.

Eksempel 4.3. Et eksempel på et projektivt plan[2] er *Fano-planet*[3]. Det har syv punkter som vi kalder a,b,c,d,e,f,g og syv linjer som vi kalder $\{a,b,c\}$, $\{a,f,g\}$, $\{a,d,e\}$, $\{c,e,g\}$, $\{c,d,f\}$, $\{b,d,g\}$, $\{b,e,f\}$ (se Figur 1.2). Lad ikke cirkelformen snyde dig! Cirklen repræsenterer linjen $\{b,e,f\}$ i Fano-planet.

[2]Nok *det* mest berømte eksempel.

[3]Opkaldt efter den italienske matematiker Gino Fano.



Figur 1.2: Fano-planet

○

Projektive planer ligner ikke helt den geometri, vi er vant til – Euklidisk geometri – og derfor er den måske heller ikke helt intuitiv. Hvad vi mister i intuition, vinder vi til gengæld tilbage i rare egenskaber. En af de mest nyttige egenskaber ved projektive planer er, at hvis et udsagn om et projektivt plan er sandt, så er *dualen* af udtrykket også sandt.

Definition 4.4 (Dualudsagn)**.** Lad p være et udsagn om punkterne og linjerne i et lineært rum. Det *duale udsagn* til p er det udsagn q, hvor man ændrer rollerne af linjerne og punkterne i geometrien.

Eksempel 4.5. Lad p være udsagnet "ethvert projektivt plan har mindst 5 punkter." Da er det duale udsagn til p "ethvert projektivt plan har mindst 5 linjer." ○

Som sagt er det dejlige ved projektive planer, at ethvert udsagn om dem medfører de tilsvarende duale udsagn.

Sætning 4.6 (Dualitetssætningen, [2, Lemma 3.1.5])**.** Lad $\mathbf{P}$ være et projektivt plan. Hvis q er et sandt udsagn om $\mathbf{P}$, så er det duale af q også sandt.

Bevis. Vi vil vise dette ved at vise at aksiomerne for projektive planer – PP_1, PP_2 og PP_3 medfører deres duale. At q er sandt vil sige

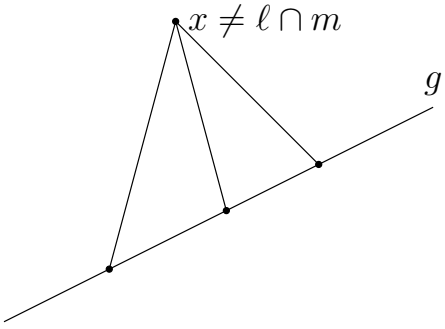
at q følger af aksiomerne for $\mathbf{P}$. Derfor ville det duale af q følge, hvis de duale af aksiomerne er sande.

PP_1 og PP_2 er hinandens duale – altså er de duale af PP_1 og PP_2 sande. At vise at det duale af PP_3 er sandt kræver lidt mere omtanke. Det duale af PP_3 er følgende udsagn.

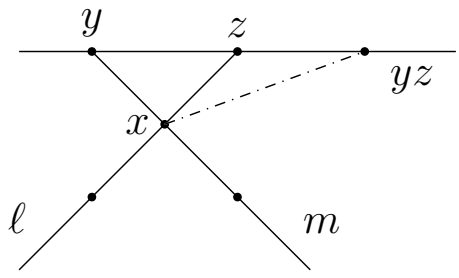
> Hvert punkt inciderer med mindst 3 linjer. Der er mindst to forskellige punkter.

Ud fra PP_3 ved vi, at der findes mindst 2 linjer ℓ og m, som hver især går gennem mindst 3 punkter. Da ℓ og m *skal* skære hinanden i ét punkt jævnfør PP_2, kan vi antage at de skærer hinanden i et af de punkter, som PP_3 garanterer. Altså er der mindst 5 forskellige punkter i $\mathbf{P}$ (hvilket er mere end 2!).

Tag et punkt x i $\mathbf{P}$. Hvis x ikke er skæringspunktet mellem ℓ og m, findes der en linje i $\mathbf{P}$, som x ikke ligger på, da x så ikke både kan være på ℓ og m. Kald denne linje g. g har mindst 3 punkter (PP_3), og pr. PP_1 er der præcis én linje mellem x og hvert af punkterne på g. Ingen af disse linjer kan være ens, da det ville medføre at en linje forskellig fra g har mere end ét skæringspunkt med g, i modstrid med PP_2. Derfor ligger x på mindst 3 forskellige linjer (se Figur 1.3).



Figur 1.3: Situation 1



Figur 1.4: Situation 2

Hvis x *er* skæringspunktet mellem ℓ og m findes der et punkt $y \in \ell \setminus m$ og et punkt $z \in m \setminus \ell$, da både ℓ og m har mindst 3

punkter, og kun deler ét (PP_3 og PP_2). Linjen yz eksisterer grundet PP_1, og x ligger ikke på yz. Ellers ville yz have to punkter til fælles med både ℓ og m, i modstrid med PP_2.

yz har mindst 3 forskellige punkter (PP_3, og derfor er der mindst 3 forskellige linjer gennem x pr. PP_1. Se Figur 1.4 for en illustration af situationen. Dermed har vi vist dualen af PP_3. ■

Sætning 4.6 er meget vigtig og fortjener derfor klart at være en sætning: Den siger essentielt set, at vi egentlig kun behøver at bevise 50% af alle sætninger om projektive planer – de resterende 50% får vi gratis pr. dualitet.

Vi vil nu introducere affine planer. De er umiddelbart en mere intuitiv type af geometri – vi vil se, at det Euklidiske plan eksempelvis er et affint rum – men til gengæld mangler de den dejlige dualitet som vi kender fra projektive planer. Senere vil vi dog se, at projektive og affine planer i virkeligheden er meget tæt relaterede.

Definition 4.7.

(a) Et *affint plan* er en geometri $\mathbf{A}$ med rang = 2 over typemængden $\{punkt, linje\}$, der opfylder følgende.

- AP_1 Hvert par af punkter i $\mathbf{A}$ inciderer med præcis én linje.
- AP_2 **Parallel aksiom**. Lad ℓ være en linje, og lad x være et punkt hvor $x \notin \ell$. Der findes netop én linje m gennem x som ikke har noget punkt til fælles med ℓ.
- AP_3 Hver linje inciderer med mindst to punkter. Der er mindst to forskellige linjer.

(b) To linjer ℓ og m i et affint plan kaldes *parallelle* hvis $\ell = m$ eller hvis ℓ og m ingen fælles punkter har. Hvis ℓ og m er parallelle, skriver vi $\ell \parallel m$.

(c) Lad ℓ være en linje i et affint plan $\mathbf{A}$, og lad $\pi(\ell)$ være mængden af linjer, der er parallelle med ℓ. Da kaldes $\pi(\ell)$ en *parallelklasse* i $\mathbf{A}$.

Ethvert affint plan er også et lineært rum. Hvis man kigger på definitionerne af affine planer og lineære rum, kan man se, at vi blot pålægger linære rum nogle ekstra krav, for at de kan være affine planer.

Eksempel 4.8 (Det Euklidiske Plan)**.** Vi vender igen tilbage til vores gode ven, det Euklidiske Plan, E. Vi ved i forvejen, at E er et lineært rum, så vi skal kun vise at AP_2 holder for E, og at E indeholder mindst to forskellige linjer.

Vi beviser først (AP_2). Lad ℓ være en linje i E, og lad $x := (x_1, x_2)$ være et punkt i E, der ikke ligger på ℓ.

Hvis ℓ er en "lodret" linje, så kan den skrives på formen $\{y\} \times \mathbb{R}$ for et $y \in \mathbb{R}$, hvor $y \neq x_1$. Da går linjen $m := \{x_1\} \times \mathbb{R}$ gennem x og er disjunkt fra ℓ, da $x_1 \neq y$. Desuden er denne linje unik: Alle andre lodrette linjer indeholder ikke x, og uanset hvilken funktion $f : \mathbb{R} \to \mathbb{R}$ vi vælger, skærer G_f ℓ, da $(y,f(y)) \in \ell$.

Hvis ℓ kan beskrives med en graf, har vi at $\ell = G_f$ for en funktion $f : \mathbb{R} \to \mathbb{R}$, hvor $f(t) = at + b$ med $a,b \in \mathbb{R}$. Vælg $c := x_2 - f(x_1)$. c er ikke lig med 0, da $c = 0$ ville medføre at $f(x_1) = x_2$, altså ville $x = (x_1, f(x_1))$, men så ville x ligge i ℓ.

Skriv da $h(t) := f(t) + c$. Vi har at $x \in G_h$, da

$$h(x_1) = f(x_1) + c = f(x_1) + x_2 - f(x_1) = x_2,$$

så $x = (x_1,h(x_1)) \in G_h$. Vi påstår at G_h ikke skærer G_f. Men det er blot en påstand. Det skal bevises.

Antag for modstrid at $G_f \cap G_h \neq \emptyset$. Det betyder, at $f(t) = h(t)$ for et eller andet $t \in \mathbb{R}$. Da gælder $at + b = at + b + (x_2 - f(x_1)) \Rightarrow 0 = (x_2 - f(x_1)) = c$. Men $c \neq 0$, så det er en modstrid. Så G_h skærer ikke g og går igennem punktet x. Nu mangler vi at vise at G_h er *den unikke* parallelle linje til G_f. For det første er ingen lodret linje parallel til G_f, så vi behøver kun at vise at der ikke findes andre lineære funktioner g, hvis graf går gennem x og ikke skærer G_g.

Antag at $g : \mathbb{R} \to \mathbb{R}$ givet ved $g(t) = \alpha t + \beta$ er sådan at $x \in G_g$ og $G_g \cap G_f = \emptyset$. Hvis $\alpha \neq a$, så kan vi prøve at se om G_f og G_g skærer,

altså om der er et $t_0 \in \mathbb{R}$, således at $f(t) = g(t)$. Vi får ligningen

$$\begin{aligned} at_0 + b &= \alpha t_0 + \beta \\ \Leftrightarrow at_0 - \alpha t_0 &= \beta - b \\ \Leftrightarrow (a - \alpha)t_0 &= \beta - b \\ \Leftrightarrow t_0 &= \frac{\beta - b}{a - \alpha}. \end{aligned}$$

Altså findes sådan et t_0, så $f(t_0) = g(t_0)$. Det vil sige, at hvis G_f og G_g *ikke skærer*, så skal $a = \alpha$. I udregningen ovenfor bruger vi *meget specifikt* at $a - \alpha \neq 0$, da $\alpha \neq a$.

Nu har vi vist at en hvilken som helst parallel linje til G_f skal have hældning a. Vi mangler altså kun at vise at β skal være lig med $(b + c)$, for at G_g indeholder x. Hvis $\beta \neq c$ har vi at

$$g(x_1) = ax_1 + \beta \neq ax_1 + b + c,$$

men da $ax_1 + c = h(x_1) = x_2$ får vi altså at $g(x_1) \neq x_2$, så $x \notin G_g$. Altså *skal* $\beta = b+c$, og vi får at $g = h$. Altså er h den unikke lineære funktion så $x \in G_h$ og $G_h \parallel G_f$, og altså er G_h den eneste linje i E, som går gennem x og ikke skærer ℓ.

Vi viser nu, at der er mindst to linjer i E. Husk, at linjerne i E blev defineret ved $L := \mathbf{G} \cup \Lambda$, hvor $\mathbf{G}$ var mængden af alle grafer for funktionerne givet ved $f : \mathbb{R} \to \mathbb{R}$ hvor $f(x) = ax+b$ med a,$b \in \mathbb{R}$, og Λ er givet ved $\{\{x\} \times \mathbb{R} \mid x \in \mathbb{R}\}$. $\mathbf{G}$ er ikke tom, da den i hvert fald indeholder grafen for $f(x) = 0$ for alle $x \in \mathbb{R}$. Derudover indeholder Λ bestemt $\{0\} \times \mathbb{R}$. Man kan måske genkende disse som det, vi ofte kalder x- og y-akserne. Disse er bestemt forskellige, da de findes i to disjunkte mængder. Det betyder, at L er foreningen af to disjunkte ikke-tomme mængder. Da er der mindst to forskellige linjer i L. ∘

Vi introducerer nu en ny slags funktion, nemlig en *ækvivalensrelation*. Idéen med ækvivalensrelationer er at kunne kigge på to elementer og tænke for sig selv, "men de er jo ens". Vi definerer en ækvivalensrelation som følgende.

Definition 4.9. Lad X være en mængde og lad $\sim$: $X \times X \to$ {Sandt,Falsk} være en binær relation[4]. Vi siger, at $\sim$ er en *ækvivalensrelation*, hvis den opfylder følgende:

1. $\sim$ er refleksiv. Altså, $\sim$ (a,a) = Sandt for alle $a \in X$. Da skriver vi $a \sim a$.

2. $\sim$ er symmetrisk. Det vil sige, at hvis $a \sim b$, så er $b \sim a$ for alle $a, b \in X$.

3. $\sim$ er transitiv. Altså, hvis $a \sim b$ og $b \sim c$, så er $a \sim c$ for alle $a, b, c \in X$.

Det var ret meget definition, men nu skal vi se, at det var det hele værd.

Sætning 4.10 ([1, Sætning 1.3.6])**.** Lad **A** være et affint plan. $\parallel$ er en ækvivalensrelation.

Bevis. Siden alle linjer er parallelle med sig selv, er $\parallel$ en refleksiv relation.

Det ses tydeligt, at hvis ℓ ingen punkter deler med m, så deler m ingen punkter med ℓ. Så $\parallel$ er også symmetrisk.

Lad $\ell \parallel m$ og $m \parallel k$. Antag at ℓ og k ikke er parallelle. Da har ℓ og k et skæringspunkt x. Dermed ligger x på to linjer, der begge er parallelle med m, som modstrider Definition 4.7(AP_2). ■

Som det blev nævnt over dette bevis, betyder det, at vi i nogle henseende kan betragte parallelle linjer i affine planer som "ens". Dette betyder også, at en parallel-klasse $\pi(\ell)$ ikke er afhængig af vores valg af linjen ℓ, altså $\pi(\ell) = \pi(m)$ for alle $m \in \pi(\ell)$.

Lemma 4.11. Lad **P** være et projektivt plan, og lad ℓ,m være to forskellige linjer i **P**. Der findes et punkt $z \in \mathbf{P} \setminus (\ell \cup m)$– altså et punkt i **P** som hverken ligger på ℓ eller m.

[4]Vi har tidligere set et eksempel på en binær relation, nemlig incidensfunktionen.

Bevis. Lad $x \in \ell \setminus m$ og $y \in m \setminus \ell$. Sådanne punkter findes, da både ℓ og m har mindst tre punkter (PP_3) og kun skærer hinanden i ét (PP_2). Der findes en unik linje xy gennem x og y (PP_1). Den er forskellig fra både ℓ og m, da $x \in xy$ men $x \notin m$, mens $y \in xy$ men $y \notin l$. xy skærer kun ℓ og m i henholdsvis x og y (PP_2), så der findes et tredje punkt z på xy, som hverken er x eller y (PP_3), og derfor hverken ligger på ℓ eller m. Altså er z hverken i ℓ eller m, og dermed har vi $z \notin \ell \cup m$. ■

Korollar 4.12. Et projektivt plan indeholder mindst tre linjer.

Nu skal vi se koblingen mellem projektive og affine planer.

Sætning 4.13 ([1, Sætning 1.3.7])**.** Lad $\mathbf{P}$ være et projektivt plan og lad $\ell \in \mathbf{P}$ være en linje. Lad $\mathbf{A}$ være geometrien med rang = 2 over typemængden $\{punkt, linje\}$, defineret som følgende:

> Punkterne i $\mathbf{A}$ er punkterne i $\mathbf{P}$, der ikke ligger på ℓ.
>
> Linjerne i $\mathbf{A}$ er linjerne i $\mathbf{P}$ foruden ℓ.
>
> Incidensen og typefunktionen i $\mathbf{A}$ er de samme som dem i $\mathbf{P}$, bare restringeret til $\mathbf{A}$, så deres domæner ikke indeholder de fjernede punkter og den fjernede linje.

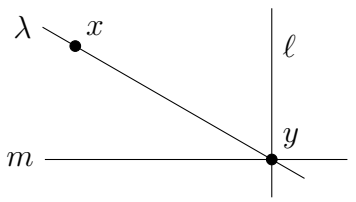
Geometrien $\mathbf{A}$ er et affint plan.

Bevis. Det, vi her skal vise, er, at ovenstående antagelser medfører, at $\mathbf{A}$ opfylder aksiomerne for affine planer. Vi tager aksiomerne et ad gangen.

AP_1 Dette aksiom følger direkte af PP_1.

AP_2 Givet en linje m og et punkt x ikke på m skal vi nu vise eksistens og entydighed af den parallelle linje til m gennem x.

> Lad m være en linje i $\mathbf{A}$ og lad x være et punkt i $\mathbf{A}$, der ikke ligger på m. Lad $y := m \cap \ell$ være skæringspunktet mellem m og ℓ i $\mathbf{P}$, og lad $k := xy$ være linjen gennem x og y.



Eksistens: Pr. konstruktion ligger punktet x på linjen k, og m og k har ingen fælles punkter **i A**, da deres oprindelige skæringspunkt, y, er blevet fjernet, da vi fjernede ℓ.

Entydighed: Lad g være en linje gennem x, som ikke har noget skæringspunkt med m **i A**. Siden g og m skærer hinanden i **P**, ligger deres skæringspunkt på linjen ℓ. Det følger, at $g \cap m = m \cap \ell = y$. Dermed gælder $g = xy = k$.

AP_3 Pr. Korollar 4.12 er vi garanteret mindst tre linjer i **P**. Dermed er vi garanteret mindst to linjer i **A**.

■

Konsekvensen af denne sætning er, at vi har en måde at konstruere affine planer fra projektive planer, nemlig ved at fjerne en linje fra det projektive plan. Helt generelt, hvis vi lader **P** være et projektivt plan og ℓ være en linje i **P**, så er $\mathbf{P} \setminus \ell$ det affine plan, der skabes ved at fjerne ℓ fra **P**[5].

Denne konstruktion giver altid et affint plan, men valget af linjen ℓ påvirker det affine plan, man ender med. Det er generelt ikke nødvendigvis tilfældet, at $\mathbf{P} \setminus \ell$ og $\mathbf{P} \setminus m$ er ens for et projektivt plan **P** og to forskellige linjer ℓ,m i **P**.[6]

Hver parallelklasse i $\mathbf{P} \setminus \ell$ svarer netop til mængden af linjer gennem et punkt i ℓ. Dette giver anledning til en ny definition, en form for "omvendt procedure" til den vi netop har set.

[5]Husk at vi identificerede linjer med deres punktmængder, så alle punkterne på linjen fjernes også, idet linjen fjernes.

[6]"Ens" betyder i dette tilfælde "isomorfe". Se Afsnit 5.

Definition 4.14. Lad $\mathbf{A}$ være et affint plan, og lad $\mathbf{P}(\mathbf{A})$ være geometrien af rang 2 over typemængden $\{punkt, linje\}$ defineret som følgende:

- Punkterne i $\mathbf{P}(\mathbf{A})$ er punkterne i $\mathbf{A}$ og parallel-klasserne i $\mathbf{A}$.
- Linjerne i $\mathbf{P}(\mathbf{A})$ er linjerne i $\mathbf{A}$ og en ekstra linje l. Denne linje kaldes for *horisontlinjen*.
- Incidensen i $\mathbf{P}(\mathbf{A})$ er defineret som følgende:

	Linje $k \in \mathbf{A}$	Horisontlinjen ℓ
Punkt $x \in \mathbf{A}$	Incidens som i $\mathbf{A}$	x og ℓ inciderer ikke
Klasse $\pi(m)$	$\pi(m)$ og k inciderer hvis og kun hvis $k \in \pi(m)$	$\pi(m)$ og ℓ inciderer for alle linjer m i $\mathbf{A}$ (punkterne på ℓ er netop parallelklasserne i $\mathbf{A}$)

Tabel 1.1: Incidens i $\mathbf{P}(\mathbf{A})$

$\mathbf{P}(\mathbf{A})$ kaldes den *projektive lukning* af $\mathbf{A}$.

Sætning 4.15 ([1, Sætning 1.3.8])**.** Lad $\mathbf{A}$ være et affint plan. Den projektive lukning $\mathbf{P}(\mathbf{A})$ af $\mathbf{A}$ er et projektivt plan.

Bevis. Vi verificerer, at $\mathbf{P}(\mathbf{A})$ opfylder PP_1, PP_2 og PP_3.

PP_1 Lad x,y være to forskellige punkter $\mathbf{P}(\mathbf{A})$.

Hvis x og y begge er i $\mathbf{A}$, så er linjen xy i $\mathbf{A}$ den unikke linje mellem x og y i $\mathbf{P}(\mathbf{A})$.

Hvis $x \in \mathbf{A}$ og $y = \pi(m)$ er en parallel-klasse i $\mathbf{A}$, findes der netop én linje $k \in \pi(m)$ gennem x ifølge AP_2. k er den unikke linje mellem x og y.

Hvis $x = \pi(m)$ for en linje m i $\mathbf{A}$ og $y = \pi(k)$ for en anden linje k i $\mathbf{A}$, så er ℓ den unikke linje gennem x og y.

PP_2 Lad k,m være to forskellige linjer i $\mathbf{P}(\mathbf{A})$.

Hvis k,m begge oprindeligt var linjer i $\mathbf{A}$ opstår to forskellige tilfælde. I det første tilfælde er k og m ikke parallelle. Så findes et skæringspunkt mellem k og m i $\mathbf{A}$, og dermed findes det samme skæringspunkt i $\mathbf{P}(\mathbf{A})$. I det andet tilfælde har vi, at $k \parallel m$. Da gælder både at $\pi(k) = \pi(m)$, og k og m skærer da hinanden i $\mathbf{P}(\mathbf{A})$ i punktet $\pi(m)$.

Hvis m er en linje i $\mathbf{A}$, og $k = \ell$ er horisontlinjen i $\mathbf{P}(\mathbf{A})$, da skærer m og k hinanden i $\pi(m)$ i $\mathbf{P}(\mathbf{A})$.

PP_3 Siden der er mindst to linjer i $\mathbf{A}$, er der også mindst to linjer i $\mathbf{P}(\mathbf{A})$. Hvis m er en linje i $\mathbf{A}$, findes der mindst to punkter i $\mathbf{A}$ på m (AP_3). Der findes et yderligere punkt $\pi(m)$ på m i $\mathbf{P}(\mathbf{A})$. g inciderer dermed med mindst tre punkter.

Det skal også vises, at horisontlinjen ℓ indeholder minimum tre punkter. Lad m,k være linjer i $\mathbf{A}$, således at $m \cap k = x$. Lad yderligere $y \in m$ og $z \in k$ være punkter forskellige fra x. Så findes en linje $h = yz$. Da har vi tre ikke-parallelle linjer, m, k og h, og dermed tre parvist disjunkte parallel-klasser, $\pi(m)$, $\pi(k)$ og $\pi(h)$. Da parallel-klasserne i $\mathbf{A}$ netop er punkterne på ℓ, har ℓ mindst tre punkter.

■

5 Lineære funktioner og isomorfier

I mange matematiske grene har man en eller anden form for begreb om en strukturbevarende afbildning – en eller anden slags funktion, som bevarer den struktur, man studerer. I gruppeteori har man for eksempel homomorfier; en homomorfi $\varphi : G \to H$, hvor G og H er grupper, er en funktion, så $\varphi(ab) = \varphi(a)\varphi(b)$ for alle a,$b \in G$. Denne slags funktioner er især vigtige, når det kommer til at afgøre, om to strukturer er "essentielt ens." Altså hvornår kan vi sige, at to strukturer (her geometrier) er de samme?

I incidensgeometri har man et par forskellige begreber om, hvad det vil sige at bevare struktur. Nogle definitioner af strukturbevarende funktioner er meget generelle og antager ingenting om de underliggende strukturer, andet end at de er geometrier [1, Kapitel 2]. Andre er mere specifikke og antager for eksempel, at vi arbejder med (nær-)lineære rum [2].

Vi kommer fremover primært til at arbejde med nær-lineære rum, så vi tager den anden tilgang til strukturbevarende funktioner. Når det kommer til spørgsmålet om, hvornår to (nær-)lineære rum er essentielt ens, vil vi forsøge at tackle dette spørgsmål ved at introducere begrebet *isomorfi*.

Lineære funktioner

Vores begreb for "strukturbevarende funktion" bliver *lineær funktion*.

Definition 5.1 (Lineær funktion)**.** Lad $\mathcal{L} = (X, *, \mathbf{type})$ og $\mathcal{M} = (Y, \star, \mathbf{type}')$ være to nær-lineære rum. Lad $P = \mathbf{type}^{-1}(punkt)$ og $Q = \mathbf{type}'^{-1}(punkt)$ være mængderne af punkter i hhv. $\mathcal{L}$ og $\mathcal{M}$. En funktion $\lambda : P \to Q$ kaldes *lineær* når

$$\lambda(\ell) \in \mathbf{type}'^{-1}(linje) \text{for alle} \ell \in \mathbf{type}^{-1}(linje).^{7}$$

[7] Her betragter vi som altid linjer som værende lig med mængden af de punkter, som de inciderer med

Altså sender λ linjer i $\mathcal{L}$ til linjer i $\mathcal{M}$.

Når vi prøver at definere en form for strukturbevarende funktion mellem to geometrier, ville det være meget naturligt at forvente, at den bevarer incidens. Heldigvis er det tilfældet for lineære funktioner. Det er en af de egenskaber ved lineære funktioner, som gør, at det ikke virker helt dumt at bruge netop dem som definitionen på "strukturbevarende afbildninger mellem lineære rum."

Lemma 5.2. Lad $\mathcal{L} = (X, *, \mathbf{type})$ og $\mathcal{M} = (Y, \star, \tilde{\mathbf{type}})$ være to nær-lineære rum med punktmængder hhv. P og Q. Lad derudover $\lambda : P \to Q$ være en lineær funktion.

(a) Lad x og y være punkter i $\mathcal{L}$, så $x * y$. Så har vi at $\lambda(x) \star \lambda(y)$.

(b) Lad ℓ og m være linjer i $\mathcal{L}$, så $\ell * m$. Så har vi at $\lambda(\ell) \star \lambda(m)$.

(c) Lad x være et punkt i $\mathcal{L}$ og ℓ en linje, så $x * \ell$. Da inciderer $\lambda(x)$ med $\lambda(\ell)$.

Bevis. (a) Da $\mathbf{type}(x) = \mathbf{type}(y)$, medfører $x * y$ at $x = y$ pr. definition af incidens. Derfor er $\lambda(x) = \lambda(y)$, og vi har, igen pr. definition af incidens, at $\lambda(x) \star \lambda(y)$.

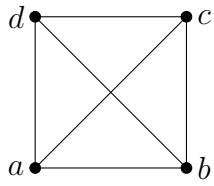
(b) Erstat x og y i del (a) med ℓ og m.

(c) Hvis $x * \ell$ i $\mathcal{L}$, vil det sige at $x \in \ell$, da vi betragter linjerne i $\mathcal{L}$ som mængderne af de punkter, som de inciderer med. Det vil sige at $\lambda(x) \in \lambda(\ell)$. Fordi $\lambda(\ell)$ er en linje i $\mathcal{M}$ pr. Definition 5.1 giver det mening at snakke om incidens med $\lambda(\ell)$. Da $\lambda(x) \in \lambda(\ell)$ har vi at $\lambda(x) \star \lambda(\ell)$.

■

Altså har vi, at hvis noget inciderer i $\mathcal{L}$, så gør de det også efter, man har brugt λ på dem.

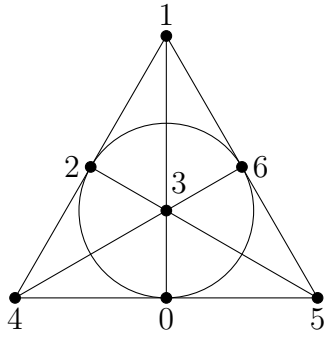
Eksempel 5.3. Betragt det lineære rum $\mathcal{T} = (X, *, \mathbf{type})$,



med punkterne a,b,c,d og hvor vi definerer linjerne som

$$L = \big\{\{x,y\} \,\big|\, x,y \in \mathbf{type}^{-1}(punkt)\big\}.$$

Her er incidens ved at et punkt x inciderer med en linje ℓ hvis og kun hvis $x \in \ell$. Lad derudover $\mathcal{P} = (Y, *, \mathbf{type}')$ være Fano-planet:



Med punkterne 0,1,2,3,4,5,6 og linjerne {1,2,4}, {1,6,5}, {4,0,5}, {4,3,6}, {2,3,5}, {1,3,0} og {2,0,6}. Her betragtes linjerne igen som mængder af punkter, så hvor en linje sendes hen bestemmes fuldstændig af hvor punkterne sendes hen.
Lad funktionen

$$\lambda : \mathbf{type}'^{-1}(punkt) \to \mathbf{type}^{-1}(punkt)$$

være givet ved

$$\begin{aligned} &\lambda(0) = d, \quad \lambda(1) = b, \\ &\lambda(2) = b, \quad \lambda(3) = b, \\ &\lambda(4) = d, \quad \lambda(5) = a, \\ &\lambda(6) = b, \end{aligned}$$

så den opfylder, at

$$\begin{aligned}
&\lambda(\{1,2,4\}) = \{b,d\}, \quad \lambda(\{0,4,5\}) = \{d,a\},\\
&\lambda(\{1,5,6\}) = \{b,a\}, \quad \lambda(\{0,1,3\}) = \{b,d\},\\
&\lambda(\{2,3,5\}) = \{b,a\}, \quad \lambda(\{3,4,6\}) = \{b,d\},\\
&\lambda(\{0,2,6\}) = \{b,d\}.
\end{aligned}$$

Denne funktion er en lineær funktion, fordi uanset hvilken linje i Fano-planet man vælger, så er billedet af den linje under λ en linje i $\mathcal{T}$. ∘

Isomorfi

I introduktionen til dette afsnit nævnte vi, at det ofte har interesse at vide, hvornår to strukturer er essentielt ens. Det er jo i virkeligheden lidt fjollet at skelne mellem to geometrier, som egentlig har præcis den samme struktur. Tag for eksempel Fano-planet fra Eksempel 4.3 og Fano-planet fra Eksempel 5.3. Selvom deres punkter og linjer ikke hedder det samme, kan vi se, at de er præcis den samme konstruktion, hvis bare vi ændrer punkternes navne. Det giver altså ikke mening at skelne mellem dem.

Til at beskrive idéen om enshed indfører vi begrebet *isomorfi*.

Definition 5.4. Lad $\mathcal{L}$ og $\mathcal{M}$ være to nær-lineære rum, og lad P og Q være mængden af punkter i henholdsvis $\mathcal{L}$ og $\mathcal{M}$. Da er $\iota : P \to \mathfrak{Q}$ en *isomorfi*, hvis ι er en bijektiv lineær funktion, og ι^{-1} også er lineær.

Vi siger, at $\mathcal{L}$ og $\mathcal{M}$ er *isomorfe* hvis der findes en isomorfi fra $\mathcal{L}$ til $\mathcal{M}$, og vi skriver i så fald $\mathcal{L} \simeq \mathcal{M}$.

Lineære rum opfører sig særlig pænt med hensyn til isomorfier: Vi skal se, at for at tjekke om en funktion er en isomorfi mellem lineære rum, så er det faktisk nok at tjekke, om den er lineær og bijektiv.

Lemma 5.5 ([2, Lemma 2.6.2])**.** Lad $\mathcal{L}$ og $\mathcal{M}$ være lineære rum. Hvis λ er en lineær bijektiv funktion fra $\mathcal{L}$ til $\mathcal{M}$, så er λ^{-1} lineær.

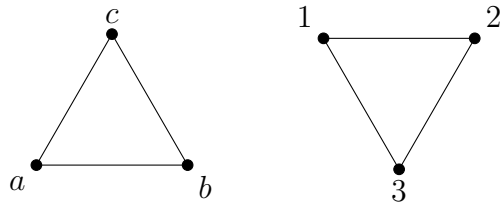
Bevis. Lad x og y være to forskellige punkter i $\mathcal{M}$. Siden λ er bijektiv, så findes der entydige punkter v og w i $\mathcal{L}$, så $\lambda(v) = x$ og $\lambda(w) = y$. Vi har så at $\lambda(vw) = xy$, da λ er lineær, og pr. Definition 3.3 $\mathcal{L}_1$ findes én og kun en linje gennem v og w, og det samme med x og y. Men så er $\lambda^{-1}(xy) = vw$.

Da alle linjer kan defineres ud fra 2 punkter, som ligger på dem ($\mathcal{L}_1$), vil det sige at λ^{-1} sender enhver linje i $\mathcal{M}$ til en linje i $\mathcal{L}$, og den er altså lineær. ■

Altså er enhver bijektiv lineær funktion mellem lineære rum en isomorfi.

Eksempel 5.6. λ fra Eksempel 5.3 er *ikke* en isomorfi – den er hverken surjektiv eller injektiv. Faktisk findes der ingen isomorfi mellem de to lineære rum i eksemplet, da de har et forskelligt antal punkter, og der derfor ikke kan eksistere en bijektion mellem deres punkter.

Til gengæld er de følgende to lineære rum isomorfe:



Prøv at finde en isomorfi mellem dem.

Desuden er *identiteten* $Id : \mathcal{T} \to \mathcal{T}$, hvor $\mathcal{T}$ er det lineære rum $\mathcal{T}$ fra Eksempel 5.3, givet ved $Id(x) = x$ for alle punkter x i $\mathcal{T}$ er en isomorfi. Faktisk er det generelt sandt: et nær-lineært rum er altid trivielt isomorft til sig selv. ○

At der altid findes isomorfier fra et nær-lineært rum til sig selv, inspirerer os til at lave følgende definition.

Definition 5.7 (Kollination)**.** Lad $\mathcal{L}$ være et nær-lineært rum. Hvis ι er en isomorfi fra $\mathcal{L}$ til sig selv, kalder vi ι en *kollination* på $\mathcal{L}$. Vi definerer desuden $\mathrm{Aut}(\mathcal{L})$ som mængden af alle kollinationer på $\mathcal{L}$.

Lemma 5.8. Lad $\mathcal{L}$, $\mathcal{M}$ og $\mathcal{N}$ være nær-lineære rum, og lad λ være en isomorfi mellem $\mathcal{L}$ og $\mathcal{M}$, mens μ er en isomorfi mellem $\mathcal{M}$ og $\mathcal{N}$. Da er $\mu \circ \lambda$ en isomorfi fra $\mathcal{L}$ til $\mathcal{N}$.

Desuden er λ^{-1} en isomorfi.

Bevis. Vi ved, at $\mu \circ \lambda$ er bijektiv, da den er en sammensætning af bijektive funktioner. Hvis ℓ er en linje i $\mathcal{L}$, så er $\lambda(\ell)$ også en linje i $\mathcal{M}$, da λ er lineær. Dermed er $\mu(\lambda(\ell))$ også en linje i $\mathcal{N}$, og dermed er $\mu \circ \lambda$ lineær. Vi kan lave samme argument for at vise at $(\mu \circ \lambda)^{-1} = \lambda^{-1} \circ \mu^{-1}$ også er lineær, så $\mu \circ \lambda$ er en isomorfi.

Da λ er en isomorfi, er λ^{-1} også lineær og bijektiv med lineær invers og dermed en isomorfi. ■

Korollar 5.9 ([2, Lemma 1.7.3])**.** Hvis λ og μ er to kollineationer på et nær-lineært rum $\mathcal{L}$, så er både λ^{-1} og $\lambda \circ \mu$ også kollineationer.

Bevis. Da λ er en kollineation, er den en isomorfi fra $\mathcal{L}$ til sig selv. Altså er λ^{-1} også en isomorfi (Lemma 5.8). Dermed er λ^{-1} en kollineation.

$\lambda \circ \mu$ er en isomorfi (Lemma 5.8), og da både domænet og kodomænet er $\mathcal{L}$, er den en kollineation. ■

Den interesserede læser kan nu verificere, at $\mathrm{Aut}(\mathcal{L})$ er en gruppe med funktionssammensætning som den binære operation og Id som identiteten.

Vi vil slutte afsnittet af med at vise at isomorfi er en ækvivalensrelation på nær-lineære rum.

Sætning 5.10. $\simeq$ (isomorfi) er en ækvivalensrelation på nær-lineære rum.

Bevis. Lad $\mathcal{L}$, $\mathcal{M}$, $\mathcal{N}$ være nær-lineære rum.

Refleksivitet: Identiteten $Id_{\mathcal{L}}$ er en isomorfi mellem $\mathcal{L}$ og $\mathcal{L}$, så $\mathcal{L} \simeq \mathcal{L}$.

Symmetri: Hvis $\mathcal{L} \simeq \mathcal{M}$, findes der en isomorfi ι fra $\mathcal{L}$ til $\mathcal{M}$. Dermed er ι^{-1} en isomorfi (Lemma 5.8), så $\mathcal{M} \simeq \mathcal{L}$.

Transitivitet: Hvis $\mathcal{L} \simeq \mathcal{M}$ og $\mathcal{M} \simeq \mathcal{N}$ findes der en isomorfi λ fra $\mathcal{L}$ til $\mathcal{M}$ og en isomorfi μ fra $\mathcal{M}$ til $\mathcal{N}$. Dermed er $\mu \circ \lambda$ en isomorfi fra $\mathcal{L}$ til $\mathcal{N}$ (Lemma 5.8), så $\mathcal{L} \simeq \mathcal{N}$.

■

Bemærkning 5.11. Lad os slutteligt diskutere, hvorfor vi mener, at isomorfe nær-lineære rum er ens.

Antag, at vi har en isomorfi ι fra $\mathcal{L}$ til $\mathcal{M}$. Hvis ℓ er en linje i $\mathcal{L}$, så er $\iota(\ell)$ en unik linje i $\mathcal{M}$. Desuden sendes hvert punkt på ℓ til et unikt punkt på $\iota(\ell)$. Hvis en linje m skærer ℓ i punktet x i $\mathcal{L}$, så skærer $\iota(m)$ $\iota(\ell)$ i netop $\iota(x)$.

I lyset af Lemma 5.2 har vi, at a inciderer med b i $\mathcal{L}$ hvis og kun hvis $\iota(a)$ inciderer med $\iota(b)$, og samtidig er $\iota(a)$ af samme type som a for alle a i $\mathcal{L}$.

Alt i alt kan $\mathcal{M}$ betragtes som en tro kopi af $\mathcal{L}$, da den knytter ethvert punkt og enhver linje i $\mathcal{L}$ til et entydigt bestemt element i $\mathcal{M}$ med præcis de samme egenskaber.

På grund af Sætning 5.10 ved vi, at vi kan dele nær-lineære rum op i ækvivalensklasser, og givet at der ikke er nogen meningsfuld forskel mellem isomorfe rum, er vi ikke interesserede i at skelne mellem to nær-lineære rum i samme ækvivalensklasse.

Det er derfor en meget interessant problemstilling, om to rum er isomorfe, eller endnu vigtigere, om man kan klassificere alle (nær-)lineære rum *op til isomorfi*.

Bemærk slutteligt, at det kan være ret svært at vise, at to rum *ikke* er isomorfe – hvis man vil vise, at to rum er isomorfe, er det så "simpelt" som at finde en isomorfi mellem rummene. Hvis man skal vise, at to rum *ikke* er isomorfe, gælder det om at vise, at der *ikke findes en eneste isomorfi* mellem dem. Det er ret let at vise, hvis rummene ikke har samme antal punkter, men ellers bliver det lidt mere involveret. Derfor handler det ofte om at finde *invarianter*:

Ting, som ikke kan være forskellige mellem isomorfe objekter. Derudover er det også smart at kende en masse nødvendige betingelser for at to lineære rum er isomorfe.

6 Endelige geometrier

Resten af dette kapitel om incidensgeometri handler om de tilfælde, hvor geometrien er *endelig*. Særligt kommer vi til at drøfte endelige lineære rum og deres egenskaber.

Det særlige ved endelige geometrier fremfor uendelige er at man rent faktisk kan have tal på ting. Vi kan begynde og tælle antal punkter, linjer, og antal punkter på linjer, etc. Det kommer vi til at benytte os af i stor stil, og vi kommer til at gennemgå en masse tricks og sætninger, som vi kan bruge som værktøj til at arbejde med disse endelige størrelser. I afsnittet "Det projektive plan af orden 2" kommer vi endda til at bruge nogle af tælle-egenskaberne til at klassificere alle projektive planer med 7 punkter op til isomorfi. Endelige geometrier har desuden den fordel, at de er nemme at forestille sig og illustrere.

Vi starter med at definere hvad en endelig (præ)geometri er.

Definition 6.1 (Endelig prægeometri)**.** Vi kalder en prægeometri $\Gamma = (X, *, \mathbf{type})$ *endelig* hvis mængden X er endelig. Ligeledes definerer vi endelig geometri.

Vi kommer særligt til at have med endelige pregeometrier over typemængden $\{punkt, linje\}$ at gøre i dette afsnit. Fordi vi derfor kommer til at tælle en del punkter og linjer, giver det god mening at indføre lidt notation vedrørende antal af punkter og linjer.

Definition 6.2. Lad Γ være en prægeometri over typemængden $I = \{punkt, linje\}$.

For en linje ℓ i Γ skriver vi $v(\ell)$ for antallet af punkter, som inciderer med ℓ. For et punkt p skriver vi $b(p)$ for antallet af linjer gennem p.

Når vi ved hvilken prægeometri vi arbejder i, skriver vi b for antallet af linjer i alt, og v for antallet af punkter i alt.

Sumnotation

I dette afsnit kommer vi til at skulle lægge en masse tal sammen. Derfor vil vi i stor grad bruge *sumnotation*. Vi må derfor definere sumnotation, da det generelt anses som "dårlig skik" i matematik at bruge udtryk og notation, som endnu ikke er blevet introduceret.

Definition 6.3 (Dårlig skik)**.** Lad T være en matematisk tekst. Når T benytter sig af endnu ikke introducerede begreber og notation, siger vi, at T har dårlig skik, og skriver >:(.

Definition 6.4 (Sumnotation)**.** Lad n være et naturligt tal, og lad $(a_1, \ldots, a_n) \subseteq \mathbb{R}$ være en endelig tupel af reelle tal. Vi definerer

$$\sum_{i=1}^{n} a_n := a_1 + \ldots + a_n.$$

Eksempel 6.5. Lad $k \in \mathbb{R}$. Da er

$$\sum_{n=1}^{3}(n \cdot k) = (1 \cdot k) + (2 \cdot k) + (3 \cdot k) = (1 + 2 + 3) \cdot k = 6k.$$

○

Incidensmatricer

Som nævnt i introduktionen er endelige geometrier nemme at illustrere – det at de er endelige gør dem *mulige* at illustrere korrekt i første omgang.[8] Tegninger er dog ikke altid den nemmeste måde at forstå en geometri på, så nogle gange er det nyttigt at bruge en *incidensmatrix*. Desuden har de den fordel, at de er nemmere at lave matematik på end tegninger er.

Definition 6.6 (Incidensmatrix)**.** Lad Γ være en endelig endelig prægeometri over typemængden $I = \{punkt, linje\}$. Da der er endeligt mange punkter i Γ kan vi give dem hvert et indeks, så de hedder

[8] Læseren udfordres til at illustrere hele det euklidiske plan på et stykke papir.

p_1,p_2,...,p_n, hvor n, og vi kan kalde linjerne ℓ_1,ℓ_2,...,ℓ_k. Vi definerer *incidensmatricen* for $\mathcal{L}$ således.

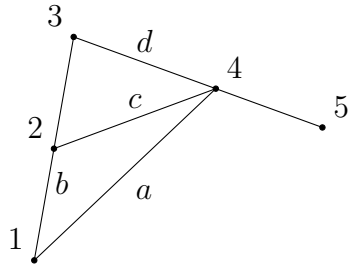
$$\begin{array}{c} \\ p_1 \\ p_2 \\ \vdots \\ p_n \end{array} \begin{array}{c} \begin{array}{cccc} \ell_1 & \ell_2 & \dots & \ell_k \end{array} \\ \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1k} \\ r_{21} & r_{22} & \dots & r_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & \dots & r_{nk} \end{pmatrix} \end{array}$$

hvor

$$r_{ij} = \begin{cases} 1 & \text{hvis } p_i * \ell_j \\ 0 & \text{ellers} \end{cases}.$$

Vi skriver også $v_i = v(\ell_i)$ og $b_i = b(p_i)$, når først vi har etableret en indeksering.

Eksempel 6.7. Betragt følgende geometri.



Incidensen i denne geometri kan beskrives med følgende matrix.

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

$\circ$

Proposition 6.8. Lad Γ være en endelig prægeometri over typemængden $\{\text{punkt}, \text{linje}\}$. Givet en incidensmatrix, gælder følgende identiteter.

$$\sum_{i=1}^{v} r_{ij} = v_j$$
$$\sum_{j=1}^{b} r_{ij} = b_i$$
$$\sum_{j=1}^{b} v_j = \sum_{j=1}^{b}\sum_{i=1}^{v} r_{ij} = \sum_{i=1}^{v}\sum_{j=1}^{b} r_{ij} = \sum_{i=1}^{v} b_i$$

Bevis. Vi er givet en incidensmatrix. Tælles alle 1-tallerne i en kolonne får vi antallet af punkter, der ligger en linje, altså $\sum_{i=1}^{b} r_{ij} = v_j$. Tælles alle 1-tallerne i en række, får vi antallet af linjer, der går gennem et punkt, og dermed $\sum_{j=1}^{v} r_{ij} = b_i$. Det giver os de første to identiteter.

Tæller vi for hver kolonne, kolonne for kolonne, får vi $\sum_{j=1}^{v} v_j$. Denne sum er antallet af 1-taller i incidensmatricen, altså antallet af gange et punkt inciderer med en linje. Ligeledes kan vi tælle for hver række, række for række, og vi får $\sum_{i=1}^{b} b_i$. Her har vi igen talt alle 1-taller i incidensmatricen, men bare række for række, altså antallet af gange en linje inciderer med et punkt.

Den sidste identitet følger direkte af de første to identiteter og observationerne foroven. ■

Bemærkning 6.9. Bemærk at incidensmatricen for en endelig prægeometri ikke er entydig: Den afhænger netop af hvilket nummer vi giver hvilket punkt og hvilken linje. Altså kan to forskellige incidensmatricer godt repræsentere den samme geometri. Det er en relativt ligetil opgave at regne ud hvor mange måder man kan omindeksere punkterne og linjerne på, og dermed hvor mange måder incidensmatricen kan skrives op på.

Altså, hvis vi bytter om på to rækker eller to kolonner i en incidensmatrix, repræsenterer den nye incidensmatrix den samme geometri.

Proposition 6.10. Lad $\mathcal{L}$, $\mathcal{M}$ være to lineære rum med incindensmatricer henholdsvist L, M. Hvis $L = M$, altså hvis alle indgangene i L og M er ens, så har vi $\mathcal{L} \simeq \mathcal{M}$.

Bevis. Lad v være antallet af punkter i $\mathcal{L}$. $\mathcal{M}$ har det samme antal punkter, da L og M er lige store. Definer ligeledes b som antallet af linjer i $\mathcal{L}$. Lad hhv. $\{p_1, \ldots, p_v\}$ og $\{q_1, \ldots q_v\}$ være punkterne i $\mathcal{L}$ og $\mathcal{M}$, i den rækkefølge som de opstår i hhv. L og M. Definer ligeledes $\{\ell_1, \ldots, \ell_b\}$ i $\mathcal{L}$ og $\{m_1, \ldots, m_b\}$ i $\mathcal{M}$.

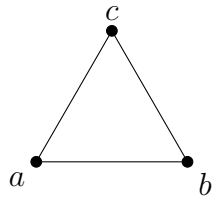
Hvis P er mængden af punkter i $\mathcal{L}$ og Q er mængden af punkter i $\mathcal{M}$, definer $\iota : P \to Q$ ved $\iota(p_k) = q_k$, $k \in \{1, \ldots, v\}$. Vi påstår, at ι er en isomorfi.

ι er lineær, da $\iota(\ell_k) = m_k$ for alle $k \in \{1, \ldots, b\}$. Det følger af, at hvis $p_i \in \ell_k$, så er den ik'te indgang i L lig med 1, og dermed er den ik'te indgang i M lig med 1, så $\iota(p_i) = q_i \in \ell_k$, og hvis $p_j \notin \ell_k$, så er $q_j \notin m_k$. Desuden er ι bijektiv, så ι er en isomorfi (Lemma 5.5). ■

Endelige lineære rum

Definition 6.11. Et *endeligt lineært rum* er en endelig geometri som også er et lineært rum. Ligeledes definerer vi endelige projektive og affine planer, samt endelig nær-lineære rum.

Eksempel 6.12. Fano planet fra Eksempel 4.3 er et endeligt lineært rum. Et andet eksempel er trekanten.



∘

Proposition 6.13 ([2, Udvalgte dele af Sætning 1.6.4 og tilhørende korollar])**.** Lad $\mathcal{L}$ være et endeligt nær-lineært rum. Da har vi at

$$\sum_{j=1}^{b} v_j(v_j - 1) = v(v-1),$$

hvis og kun hvis $\mathcal{L}$ er et lineært rum.

Bevis. Vi viser kun "hvis"-delen, da "kun hvis"-delen ikke er nødvendigt i dette afsnit, og er relativt mere indviklet, uden at være spændende eller vellystinspirerende nok til at gøre op for det. Den kan dog stadig være nyttig i opgaveløsning, og derfor har vi valgt at skrive sætningen i sin fulde magt alligevel. Et komplet bevis findes i [2, s. 15].

Vi starter med at tælle antallet af unikke par af punkter i $\mathcal{L}$ – altså hvor mange forskellige måder vi kan udvælge to punkter fra $\mathcal{L}$ på. Her vil parret (x,y) og (y,x) tælle som *det samme par*. I kombinatorik findes der en formel for at regne netop dette antal. Hvis man har en mængde med A med n punkter, og $k \in \{0,1,\ldots,n\}$ så er

$$C(n,k) = \frac{n!}{k!(n-k)!},$$

antallet af unikke delmængder af A, som indeholder k elementer, hvor

$$n! := n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 1$$

kaldes n *fakultet* (se evt. [3, Sætning 6.3.1] eller søg på *binomialformlen*). Altså må der være

$$C(v,2) = \frac{v!}{2!(v-2)!} = \frac{v(v-1)(v-2)!}{2(v-2)!} = \frac{v(v-1)}{2}$$

unikke par af punkter i $\mathcal{L}$. Men da $\mathcal{L}_1$ giver os at ethvert par af punkter i $\mathcal{L}$ bestemmer én og kun én linje i $\mathcal{L}$, kan vi også tælle antallet af punkter i $\mathcal{L}$ på følgende måde.

Éthvert par af punkter er garanteret at have en linje til fælles. Desuden, hvis et par af punkter ligger på én linje, kan det samme par ikke ligge på en forskellig linje, da det ville stride med $\mathcal{L}_1$. Det vil sige, at hvis vi finder antallet af unikke par af punkter på hver linje ℓ i $\mathcal{L}$ og så lægger alle disse antal sammen, så skal det være præcis lig antallet af unikke par af punkter i $\mathcal{L}$ i alt. Men på en given linje ℓ_i har vi, pr. samme argument som ovenfor, at antallet af unikke par af punkter er $\frac{v_i(v_i-1)}{2}$. Altså får vi

$$\begin{aligned} \sum_{i=1}^{b} \frac{v_i(v_i-1)}{2} &= \frac{v(v-1)}{2} \\ \Leftrightarrow \frac{1}{2}\sum_{i=1}^{b} v_i(v_i-1) &= \frac{v(v-1)}{2} \\ \Leftrightarrow \sum_{i=1}^{b} v_i(v_i-1) &= v(v-1), \end{aligned}$$

præcis som hævdet. ■

Lineære funktioner mellem endelige nær-lineære rum

Vi har i det tidlige afsnit om lineære funktioner diskuteret generelle egenskaber om den slags funktioner, samt isomorfi. Når det kommer til endelige nær-lineære rum har vi dog et par ekstra ting at skrive om dem – ting, som kun giver mere retfærdiggørelse for hvorfor vi mener, at isomorfi er et godt begreb for enshed.

Lemma 6.14 ([2, Lemma 1.7.1])**.** Lad $\mathcal{L}$ og $\mathcal{M}$ være nær-lineære rum med punktmængderne henholdsvis P og Q. lad ydermere $f : P \to Q$ være en injektiv lineær funktion. Så er $v(\ell) = v(f(\ell))$ for alle linjer ℓ i $\mathcal{L}$.

Bevis. Hvert punkt i ℓ sendes til et unikt punkt i $\mathcal{L}$, da f er injektiv. For hvert punkt i ℓ er der altså et unikt punkt i $f(\ell)$. Altså er der

mindst lige så mange punkter i $f(\ell)$ som i ℓ. Det ville til gengæld være absurd hvis $v(f(\ell)) > v(\ell)$. Det kan nemlig kun lade sig gøre hvis f ikke er en funktion. ■

Proposition 6.15 ([2, Lemma 1.7.2])**.** Hvis f er en isomorfi fra $\mathcal{L}$ til $\mathcal{M}$, hvor $\mathcal{L}$,$\mathcal{M}$ er nær-lineære rum, så er $v(\ell) = v(f(\ell))$ og $b(p) = b(f(p))$ for alle linjer ℓ og punkter p i $\mathcal{L}$.

Bevis. Den første del af propositionen er bare Lemma 6.14.

Bemærk at hvis $p \in \ell$ så er $f(p) \in f(\ell)$, så $b(p) \leq b(f(p))$ (Lemma 5.2). Siden f^{-1} også er en isomorfi, navnlig fra $\mathcal{M}$ til $\mathcal{L}$ har vi også den modsatte ulighed. ■

Bemærk at dette giver os en nødvendig betingelse for at to nær-lineære rum kan være isomorfe, og derfor noget vi kan bruge til at modbevise at to nær-lineære rum er isomorfe: Hvis $\mathcal{L}$ og $\mathcal{M}$ er nær-lineære rum, så *skal* følgende være sandt.

Proposition 6.16. Skriv $v(\mathcal{L})$ for antallet af punkter i $\mathcal{L}$, og skriv ligeledes $b(\mathcal{L})$ for antallet af linjer. Gør det tilsvarende for $\mathcal{M}$. Lad L være en incidensmatrix for $\mathcal{L}$ og M være en incidensmatrix for $\mathcal{M}$. Skriv λ_{ij} for den ij'te indgang i L og m_{ij} for den ij'te indgang i M.

Lad $i \in \{1, \ldots, b(\mathcal{L})\}$. Så gælder det, at der skal findes et $j \in \{1, \ldots, b(\mathcal{M})\}$ så

$$\sum_{k=1}^{v(\mathcal{L})} \lambda_{ki} = \sum_{k=1}^{v(\mathcal{M})} m_{kj}.$$

Ydermere skal vi kunne finde sådan et j *uden gentagelse* for ethvert $i \in \{1, \ldots, b(\mathcal{L})\}$. Dvs. at for hver linje i $\mathcal{L}$ skal vi *uden gentagelse* kunne finde en linje i $\mathcal{M}$, som går gennem lige så mange punkter. Det duale udsagn gælder også (se Definition 4.4), og præcis det samme skal gælde, hvis vi bytter om på $\mathcal{L}$ og $\mathcal{M}$, da isomorfi er en symetrisk relation.

Endelige projektive planer

I dette delafsnit kommer vi til at udforske nogle af de kombinatoriske egenskaber, som er særlige for endelige projektive rum. Undervejs vil vi vise De Bruijn-Erdős sætningen, som er en sætning, der blandt andet giver os et godt kriterie for at identificere hvorvidt et givet lineært rum er et projektivt plan – et kriterie, som kan være betydeligt mindre tungt at verificere end aksiomerne for et projektivt plan. Desuden vil vi vise, at dette kriterie ikke kun er tilstrækkeligt for at et lineært rum er et projektivt plan, men at det faktisk også er nødvendigt. Det lader os formulere følgende sætning: Et lineært rum $\mathcal{L}$ er et projektivt plan hvis og kun hvis $v = b$ og der findes mindst to linjer som hver inciderer med mindst 3 punkter i $\mathcal{L}$ (Sætning 6.26).

De Bruijn-Erdős sætningen

De Bruijn-Erdős sætningen er et vigtigt første skridt i retningen af at kunne bakke vores modige påstand op. Særligt giver den os den ene retning: Hvis $v = b$ og der findes mindst to linjer med mindst 3 punkter i det lineære rum $\mathcal{L}$, så er det et projektivt plan.

Paul Erdős, ham, som giver navn til halvdelen af af dette bemærkelsesværdige resultat, var en ret velkendt matematiker, og han var mindst ligeså interessant, som han var velkendt. Derfor finder vi det passende at dedikere et lille paragraf til ham. Erdős er kendt for en del forskellige ting. Erdős var praktisk talt hjemløs det meste af sit liv, af eget valg. Han rejste fra sted til sted, og besøgte forskellige matematikere for at samarbejde, hvor han så boede hos vedkommende under samarbejdet. Efter opholdet ville han rejse videre til den næste. Han brugte på den måde rigtig meget tid på at skrive matematiske artikler, og han har i sin levetid udgivet omtrent 1500

Figur 1.5: Paul Erdős til et seminar i Budapest (efterår 1992). Kilde: [4]

matematiske artikler, og samarbejdet med over 500 medforfattere. [5] Inspireret af mængden af medforfattere, han har haft, har man defineret Erdős-tallet.

Definition 6.17 (Erdős-tallet)**.** En persons Erdős-tal defineres induktivt. Paul Erdős har Erdős-tal 0. Hvis man har skrevet en artikel med en person, som har Erdős-tal n, får man selv Erdős-tal $n+1$, med mindre man har skrevet en artikel med en person med et lavere Erdős-tal. Man bliver altid tildelt det lavest mulige Erdős-tal. Hvis ikke man kan gives noget Erdős-tal på denne måde, får man Erdős-tal ∞.

Begge forfattere af dette kapitel har Erdős-tal ∞.

Lemma 6.18 ([2, Lemma 2.2.1])**.** Lad $\mathcal{L}$ være et lineært rum med v linjer og b punkter. Da gælder $\sum_{i=1}^{v} b_i(b_i-1) \leq b(b-1)$.

Bevis. Ser vi på størrelsen $b_i(b_i-1)$, opdager vi at $b_i(b_i-1) = \frac{b_i!}{(b_i-2)!}$. Denne størrelse er $C(b_i,2)$, og på lignende vis som i beviset for Proposition 6.13, beskriver den antallet af ordnede linjepar gennem punktet p_i, pr. [3, Sætning 6.3.1]. Altså tæller venstresiden af uligheden alle ordnede par af linjer, der skærer hinanden. På lignende vis ser vi, at højresiden tæller alle ordnede linjepar i $\mathcal{L}$. Der er bestemt mindst lige så mange ordnede linjepar i alt, som der er ordnede par af linjer, der skærer hinanden. ■

Korollar 6.19. Lad $\mathcal{L}$ være et lineært rum. Alle par af linjer i $\mathcal{L}$ har et skæringspunkt hvis og kun hvis $\sum_{i=1}^{v} b_i(b_i-1) = b(b-1)$.

Bevis. Vi har allerede $\sum_{i=1}^{v} b_i(b_i-1) \leq b(b-1)$, pr. Lemma 6.18. Ligheden kommer til gengæld af, at da alle linjer skærer hinanden i ét punkt, så må der være præcis så mange ordnede par af linjer, der skærer hinanden, som der er ordnede par af linjer i alt. ■

Sætning 6.20 (de Bruijn-Erdős, [2, Sætning 2.2.2])**.** Lad $\mathcal{L}$ være et endeligt lineært rum med $b > 1$. Det gælder at

(a) $b \geq v$,

(b) hvis $b = v$ har alle par af linjer et skæringspunkt.

I tilfælde (b) kan der opstå to situationer: Enten har én linje $v-1$ punkt, mens alle andre linjer har 2 punkter (alle andre linjer end den "store" linje ville så bestå af et punkt på den store linje og det ene punkt, som ikke ligger på den) eller også har hver linje $k+1$ punkter, mens hvert punkt ligger på $k+1$ linjer. Her skal $k \geq 2$ være et naturligt tal.

Bevis. (a) Lad $M := \min\{b_i \mid 1 \leq i \leq v\}$.

Bemærk at $M > 1$ – ellers ville samtlige punkter ligge på den samme linje, hvilket ville medføre $b = 1$, i strid med vores antagelser. Lad p_v være et punkt med M linjer gennem sig, og kald dem $\ell_1, \ldots, \ell_m$. Lad $p_i \in \ell_i$, for alle $1 \leq i \leq M$, så $p_i \neq p_v$.

Så har vi at $i \neq j \Rightarrow p_i \notin \ell_j$.

Derfor har vi

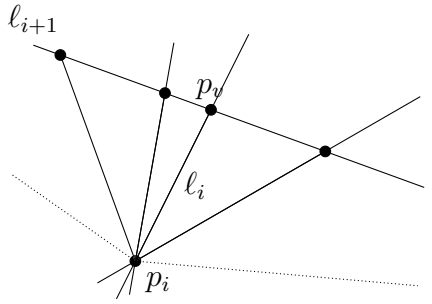
$$v(\ell_{i+1}) \leq b(p_i), \forall i \in \{1,2,\ldots,M-1\}. \tag{1.1}$$

(se figur Figur 1.6). Vi har nemlig, at der for alle punkter på ℓ_{i+1} findes en linje fra p_i til det punkt, da $p_i \notin \ell_{i+1}$. Ingen af disse linjer kan være lig hinanden, så der er mindst $v(\ell_i)$ linjer gennem p_i. Vi har også at $v_1 \leq b_m$ af samme årsag. Da $b(p_v) = M$ er minimummet af alle b_i-er, har vi $b(p) \geq b(p_v)$ for alle punkter p. Samtidig har vi $b(p_v) \geq v(\ell)$ for alle linjer ℓ, som ikke går gennem p_v – situationen er meget lig Figur 1.6. Altså har vi

$$b(p) \geq b(p_v) \geq v(\ell), \text{ når } p_v \notin \ell. \tag{1.2}$$

Vi ved fra Proposition 6.8 at $\sum_{j=1}^{b} v_j = \sum_{i=1}^{v} b_i$. Vi kan splitte summen op, så vi får

$$\sum_{j=1}^{M} v_j + \sum_{j=M+1}^{b} v_j = \sum_{i=1}^{M} b_i + \sum_{i=M+1}^{v} b_i.$$



Figur 1.6: Illustration af p_i og ℓ_{i+1}.

Vi har fra diskussionen vedrørende Ligning (1.1) at

$$\sum_{j=1}^{M} v_j \leq \sum_{i=1}^{M} b_i.$$

Desuden har vi for alle $j \in \{M+1, M+2, \ldots, b\}$ at $p_v \notin \ell_j$, så Ligning (1.2) giver os at $v(\ell_j) \leq b(p_v) \leq b(p)$ for alle punkter p.

Hvis $b < v$ medfører det at $v_j \leq b_j$ for alle $j \in \{M+1, M+2, \ldots, b\}$. Derfor har vi

$$\begin{array}{lcl} v_1 & \leq & b_M \\ v_2 & \leq & b_1 \\ v_3 & \leq & b_2 \\ & \vdots & \\ v_M & \leq & b_{M-1} \\ v_{M+1} & \leq & b_{M+1} \\ v_{M+2} & \leq & b_{M+2} \\ & \vdots & \\ v_b & \leq & b_b \end{array} .$$

Dermed er

$$\sum_{i=1}^{v} b_i = \sum_{j=1}^{b} v_j \leq \sum_{i=1}^{b} b_i.$$

Men under vores antagelse at $b < v$ burde dette være umuligt: Der går mindst én linje gennem hvert punkt,[9] så det medfører at

$$\sum_{i=1}^{v} b_i > \sum_{i=1}^{b} b_i,$$

og altså må antagelsen at $b < v$ være forkert, og vi har vist (a).

(b) Hvis $v = b$, er $v_j = b_j$ for alle $j \in \{M+1, M+2, \ldots, b\}$: Vi har allerede $v_j \leq b_j$, og $b_j \leq v_j$, fordi hvis $b_j > v_j$ ville vi have

$$\begin{aligned} \sum_{i=1}^{v} b_i = \sum_{j=1}^{b} v_j &< \sum_{i=1}^{b} b_i, \\ \Leftrightarrow \sum_{i=1}^{b} b_i &< \sum_{i=1}^{b} b_i \text{ da } v = b, \end{aligned}$$

hvilket er umuligt. På præcis samme vis har vi $b_i = v_{i+1}$ for alle $i \in \{1, \ldots, M-1\}$, og $b_m = v_1$. Fra Proposition 6.13 har vi at

$$\begin{aligned} \sum_{j=1}^{b} v_j(v_j - 1) &= v(v-1) \\ \Rightarrow \sum_{i=1}^{v} b_i(b_i - 1) &= b(b-1), \end{aligned}$$

hvor den anden lighed følger af at $v = b$ og diskussionen ovenfor. Men så giver Korollar 6.19 at alle par af linjer i $\mathcal{L}$ har et skæringspunkt. Der er nu to forskellige muligheder.

Mulighed 1: Antag at der er to linjer ℓ og ℓ' i $\mathcal{L}$, så alle punkter ligger på enten ℓ eller ℓ'. Da de skærer hinanden har vi at der findes et $p \in \ell \cap \ell'$. Hvis vi tager et $x \in \ell \setminus \{p\}$

[9] Da $b > 1$ har vi at der minimum fire punkter, og alle par af punkter ligger på én linje.

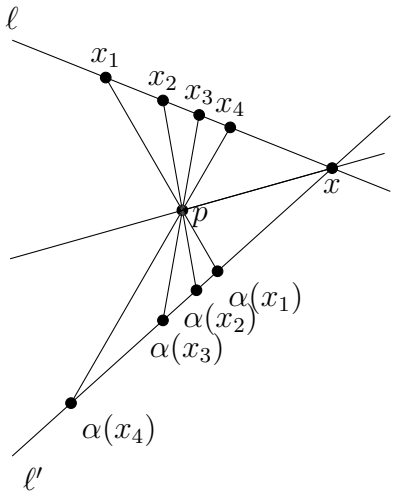
og et $x' \in \ell' \setminus \{p\}$[10], så har xx' kun to punkter – hvis xx' havde mere end to punkter ville den have mindst et punkt $y \neq x$ til fælles med enten ℓ eller ℓ' – antag at $y \in \ell$ uden tab af generalitet. Men det medfører pr. $\mathcal{L}_2$ at $xx' = \ell$, hvilket er umuligt, at $x' \in xx'$ men $x' \notin \ell$. Altså har xx' kun to punkter.

Hvis $z \in \ell \setminus \{p\}$ og $z' \in \ell' \setminus \{p\}$, så $z \neq x$ og $z' \neq x'$, så må zz' også kun have to punkter, z og z'. Men det vil sige at den ikke skærer xx', hvilket modstrider at alle linjer har et skæringspunkt, så det er umuligt.

Altså har én af linjerne kun to punkter. Antag uden tab af generalitet at det er ℓ. Da ℓ skærer ℓ' i ét enkelt punkt må der være $v - 1$ punkter på ℓ'.

Mulighed 2: Lad ℓ og ℓ' være to forskellige linjer i $\mathcal{L}$, og antag at $p \notin \ell \cup \ell'$.[11] Definér $\alpha : \ell \to \ell'$ på følgende måde.

For $x \in \ell$, lad $\alpha(x) = px \cap \ell'$ være skæringspunktet mellem px og ℓ'. Bemærk at hvis $x = \ell \cap \ell'$, så er $\alpha(x) = x$.



[10]Det må vi godt givet aksiomerne for lineære rum!

[11]Hvis ikke sådan et p fandtes, ville vi i virkeligheden være i mulighed 1.

α er surjektiv: Hvis $x' \in \ell'$, så er $\alpha(x'p \cap \ell) = x'$. Desuden er α injektiv: Hvis $\alpha(x) = \alpha(y)$ vil det sige at $xp \cap \ell' = yp \cap \ell'$. Men det vil sige at xp og yp begge indeholder p og $\alpha(y)$, så de må være den samme linje ($\mathcal{L}_2$). Da $xp = yp$ må $x = xp \cap \ell = yp \cap \ell = y$. Det vil sige at $\alpha : \ell \to \ell'$ er en bijektion, så $v(\ell) = v(\ell')$. Siden ℓ og ℓ' bare var to arbitrære forskellige linjer har vi vist at alle linjer har $v(\ell)$ punkter. Skriv $k + 1 := v(\ell)$.

Hvis p er et punkt i $\mathcal{L}$ findes der mindst en linje, som p ikke er på: Der findes mindst to linjer, m og m' pr. antagelse. Hvis $p = m \cap m'$ kan vi vælge $x \in m \setminus \{p\}$ og $x' \in m' \setminus \{p\}$. Da er p ikke på xx'. Hvis $p \neq m \cap m'$, ligger p enten ikke på m eller ikke på m'. Vælg en linje g så $p \notin g$. Der findes $v(g) = v(\ell) = k + 1$ forskellige linjer gennem p som skærer g ($\mathcal{L}_1$). Da alle linjer gennem p *skal* skære g, må dette være antallet af linjer gennem p i alt. Altså går der for alle punkter p i $\mathcal{L}$ præcis $k + 1$ linjer gennem p.

$k \geq 1$, pr. $\mathcal{L}_2$. Hvis $k = 1$ får vi en trekant som den fra Eksempel 6.12 – men trekanten falder under mulighed 1, så altså må $k \geq 2$.

Beviset er nu færdigt. ■

Korollar 6.21. Ethvert endeligt lineært rum med $b = v$, og hvor der findes mindst 2 forskellige linjer med mindst 3 punkter er et projektivt plan.

Bevis. Vi skal vise, at hvis vi har et lineært rum $\mathcal{L}$ opfylder, at $b = v$, og at der findes $i,j \in \{1, \ldots, b\}$, $i \neq j$, så både $v_i \geq 3$ og $v_j \geq 3$, så er $\mathcal{L}$ automatisk et projektivt plan. Det gør vi ved at gennemgå aksiomerne for et projektivt plan ét for ét, og tjekke om $\mathcal{L}$ opfylder dem.

PP_1 Dette aksiom er opfyldt, da det er identisk med $\mathcal{L}_1$, som $\mathcal{L}$ opfylder, da det er et lineært rum.

PP_2 $\mathcal{L}$ opfylder dette aksiom, pr. Sætning 6.20(b).

PP_3 Der vil altid være mindst 2 forskellige linjer i $\mathcal{L}$ pr. antagelse. Da mindst to af disse linjer hver især inciderer med mindst 3 punkter, giver Sætning 6.20(b), situation 2, at *samtlige* linjer i $\mathcal{L}$ går gennem mindst 3 punkter, så derfor er dette aksiom også opfyldt.

■

Sætning 6.20 giver os altså et værktøj til at identificere projektive planer. Hvis vi først ved at vi arbejder i et endeligt lineært rum, er det nu ret let at tjekke om rummet er et projektivt plan, bare ved at kigge på incidensmatricen.

En alternativ karakterisering af endelige projektive planer

I det foregående delafsnit formåede vi at gøre halvdelen af det hårde benarbejde for at bevise vores modige påstand fra starten af delafsnittet om endelige projektive planer: Et lineært rum $\mathcal{L}$ er et projektivt plan hvis og kun hvis $v = b$ og der findes mindst to linjer med mindst 3 punkter i $\mathcal{L}$. Nu er der altså en smule håb for at det vi nu engang har hævdet rent faktisk er sandt. Dette delafsnit kommer til at udføre den anden halvdel af benarbejdet, i form af at bevise følgende sætning.

Sætning 6.22 ([6, Sætning 6.3])**.** Lad $\mathbf{P} = (X, *, \mathbf{type})$ være et endeligt projektivt plan. Så er der lige mange punkter og linjer i $\mathbf{P}$, altså $b = v$. Derudover ligger der lige mange punkter på alle linjer, og dette antal er lig antallet af linjer gennem ethvert punkt. Hvis $k + 1$ er antallet af punkter på hver linje (eller, ækvivalent, antallet af linjer gennem ethvert punkt), så er $v = b = k^2 + k + 1$.

Bemærkning 6.23. At der er lige mange punkter på alle linjer, og at dette antal er lig antallet af linjer gennem ethvert punkt, er ikke noget særligt for *endelige* projektive planer, som vi kommer til at

se i beviset for sætningen. Derfor kunne sætningen formuleres på en løsere måde.

Herfra kan vi definere orden.

Definition 6.24 (Orden af et projektivt plan)**.** Vi kalder tallet k fra Sætning 6.22 for *ordenen* af det projektive plan.

Det viser sig at Sætning 6.22 ikke er helt trivielt at vise. Faktisk skal vi først bruge følgende Lemma.

Lemma 6.25. Lad $\mathbf{P}$ være et projektivt plan, og lad $x \in \mathbf{P}$ være et vilkårligt punkt. Der findes en linje i $\mathbf{P}$, som ikke går gennem x.

Bevis. Hvis ℓ er en linje med $x \in \ell$, så findes der en forskellig linje m pr. PP_3. Da både ℓ og m indeholder minimum 3 punkter, findes der to punkter på m, som ikke er lig x, og det samme med ℓ. Lad $y \in m \setminus \ell$ og $z \in \ell \setminus m$ begge være forskellige fra x. Der findes sådanne punkter, da m og ℓ kun skærer hinanden i ét punkt. Så indeholder yz ikke x: Hvis $x \in yz$, så er $yz = zx = \ell$, pr PP_1. Men det er umuligt, da y ikke ligger på ℓ, men ligger på yz, så vi får en modstrid. Altså indeholder yz ikke x. ■

Med denne viden kan vi nu bevise Sætning 6.22.

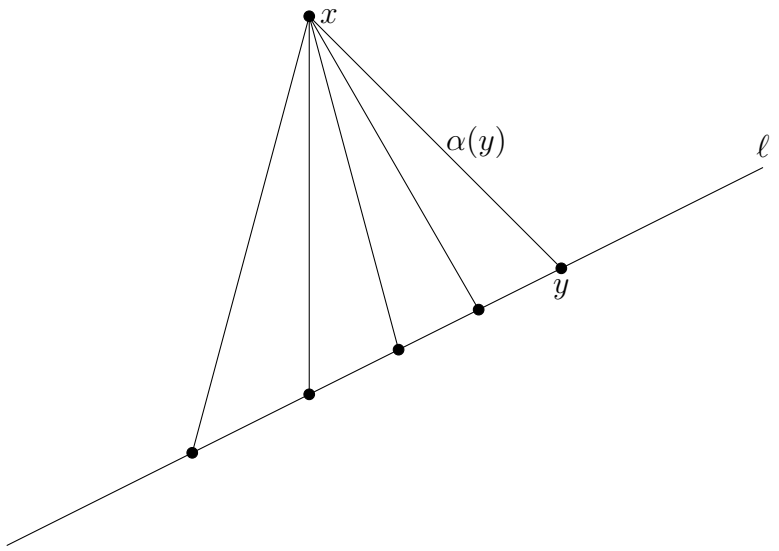
Bevis for Sætning 6.22. Lad $\mathbf{P}$ være et projektivt plan, og lad ℓ være en linje i $\mathbf{P}$. For alle punkter $p \in \mathbf{P}$ skriver vi $[p]$ for mængden af punkter, linjer, som går gennem p.

Hvis x er et punkt, som ikke inciderer ℓ, findes der en bijektion $\alpha : \ell \to [x]$, defineret på følgende måde.

$$\alpha(y) = yx.$$

α er injektiv, da ℓ og yx skærer i ét bestemt punkt, som konsekvens af PP_2. Dette punkt er netop y. Altså, hvis $\alpha(y) = \alpha(z)$, så har vi at $yx = zx$, så yx og zx skærer ℓ i samme punkt. Da y og z begge ligger på ℓ, må de derfor være det samme punkt. α er også surjektiv:

Hvis $m \in [x]$, så skærer den ℓ pr. PP_2; lad os kalde skæringspunktet y. Dermed er $m = yx = \alpha(y)$.



Figur 1.7: Illustration af hvad α gør ved et punkt $y \in [\ell]$.

Da α er en bijektion har vi altså, at $b(x) = v(\ell)$.

Lad ℓ og m være to forskellige linjer i $\mathbf{P}$. Der findes et punkt $z \notin \ell \cup m$ (Lemma 4.11). Da $z \notin m \cup \ell$ har vi, pr. den tidligere argumentation med bijektionen, at $v(m) = v(\ell) = b(z)$. Derfor, hvis man har to forskellige linjer, så har de det samme antal punkter. Kald dette antal $k + 1$.

Lad nu x være et vilkårligt punkt i $\mathbf{P}$, og lad ℓ være en linje, som ikke går gennem x (Lemma 6.25). Så er $b(x) = v(\ell) = k + 1$. Med det har vi vist, at alle punkter x ligger på $k + 1$ linjer, og alle linjer ℓ går gennem $k + 1$ punkter.

Det var lidt af en mundfuld, men nu mangler vi så også kun at vise, at $v = b$, og at der er k^2+k+1 punkter og linjer i alt. Vi starter med punkterne.

Lad x være et punkt i $\mathbf{P}$. Der findes $k+1$ linjer gennem x, og de går alle gennem x. Ethvert par af linjer gennem x skærer til gengæld kun hinanden i x. Det vil sige, at hver linje indeholder k unikke

punkter, og så x. Ethvert punkt $y \neq x$ i $\mathbf{P}$ ligger på præcis én linje, som går gennem x (PP_1), så vi har at antallet af punkter må være antallet af alle punkter på alle linjer gennem x. Det bliver altså til $b(x) \cdot k + 1$, da der for hver linje gennem x er k unikke punkter, og vi har $b(x)$ linjer gennem x – det giver $b(x) \cdot q$ punkter. Derudover har vi punktet x, så det i alt bliver til $b(x) \cdot k + 1 = (k+1) \cdot k + 1 = q^2 + q + 1$ punkter i $\mathbf{P}$.

Ligeledes kan vi tage en vilkårlig linje ℓ. Der er $k+1$ punkter på ℓ. Gennem hvert punkt er der $k+1$ linjer, hvoraf én er ℓ. Derfra får vi på samme måde som før, at hvert punkt på ℓ ligger på k unikke linjer, og så ℓ, da enhver linje i $\mathbf{P}$ skærer ℓ i et og kun et punkt (PP_2). Derfor er der $k^2 + k + 1$ linjer, som skærer ℓ i et punkt, og fordi alle linjer skærer ℓ i et eller andet punkt, må der i alt være $k^2 + k + 1$ linjer i $\mathbf{P}$

Altså har vi at $v = b = k^2 + k + 1$. ■

I lyset af Sætning 6.22, samt Sætning 6.20 kan vi nu efterhånden sige en del om endelige projektive planer, og dette leder os til hovedresultatet for dette delafsnit om endelige projektive planer.

Sætning 6.26 (Hovedsætnignen om endelige projektive planer)**.** Lad $\mathcal{L}$ være et endeligt lineært rum. Da har vi at

a) $\mathcal{L}$ er et projektivt plan hvis og kun hvis $v = b$ og der findes mindst to linjer med mindst 3 punkter i $\mathcal{L}$.

b) Hvis $\mathcal{L}$ er et projektivt plan, så går der præcis lige mange linjer gennem hvert punkt i $\mathcal{L}$, og dette antal er præcis lig antallet af punkter på hver linje i $\mathcal{L}$, og er derfor større end eller lig 3 (PP_3).

c) Hvis $\mathcal{L}$ er et projektivt plan, og k er ordenen af $\mathcal{L}$, så $v_i = k+1$ for alle $i \in \{1, \ldots, b\}$, så er der i alt $k^2 + k + 1$ linjer i $\mathcal{L}$ og det samme gælder antallet af punkter.

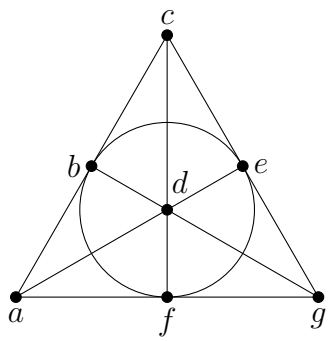
Bevis. Sætningen følger direkte af Sætning 6.22 og Korollar 6.21. ■

Med denne hovedsætning har vi en gang for alle vist at vi altså ikke løj, dengang vi i starten af delafsnittet om endelige projektive planer hævdede at vi havde fundet en ækvivalent definition af hvad det vil sige at være et endeligt projektivt plan, og vi har desuden fået nogle belejlige tælleværktøjer med på vejen.

Det projektive plan af orden 2

Dette delafsnit bliver et case-study af en særlig slags endelig geometri, nemlig *det* projektive plan af *orden* 2. Når vi lægger vægt på ordet *det*, er det fordi der faktisk kun findes et enkelt et (op til isomorfi) – og det er netop det, vi har tænkt os at vise i dette afsnit.

Sætning 6.27 ([1, Sætning 1.8.7])**.** Ethvert endeligt projektivt plan **P** af orden 2 er Fano planet:



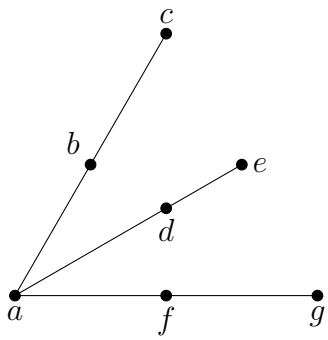
Figur 1.8: Det endelige projektive plan af orden 2

Altså har **P** syv punkter og syv linjer, og hvis vi giver punkterne navnene $\{a,b,c,d,e,f,g\}$, så kan linjerne skrives som $\{a,b,c\}$, $\{a,d,e\}$, $\{a,f,g\}$, $\{b,d,g\}$, $\{b,e,f\}$, $\{c,d,f\}$ og $\{c,e,g\}$.

Bevis. Vi betragter en endelig geometri af orden 2, og så ser vi hvilke ting, der er nødt til at gælde om den.

Bemærk først, at i kraft af Sætning 6.22, må **P** have 7 punkter, da dens orden er 2, så vi får at antallet af punkter i **P** er $2^2+2+1=7$.

Der må være tre linjer, som går gennem a, da $\mathbf{P}$ har orden 2. Da vi endnu ikke har lavet nogen antagelser om de andre punkter kan vi antage, at de tre linjer er $\{a,b,c\}$, $\{a,d,e\}$ og $\{a,f,g\}$. Vi får følgende konstruktion indtil videre



Da $\mathbf{P}$ er et projektivt plan, skærer linjen gennem c og e linjen $\{a,f,g\}$. Bemærk at skæringspunktet ikke kan være a, da vi allerede har 3 linjer gennem a, men ellers er der indtil videre ikke noget, som gør at *ce skal* skære specifikt f eller specifikt g, så vi vælger at antage, at *ce* skærer $\{a,f,g\}$ i punktet g, uden tab af generalitet. Så *skal cd* skære $\{a,f,g\}$ i f; ellers ville der være to linjer gennem c og f eller c og a.

På samme måde skal *bd* skære $\{c,e,g\}$ i g, og *be* møder $\{a,f,g\}$ i f.

Med alt dette in mente får vi linjerne $\{a,b,c\}$, $\{a,d,e\}$, $\{a,f,g\}$, $\{b,d,g\}$, $\{b,e,f\}$, $\{c,d,f\}$ og $\{c,e,g\}$. ■

7 Perspektiver og videre læsning

Der er meget mere at sige (og skrive) om incidensgeometri end vi nogensinde ville have tid til at formidle på en sommerlejr, selv hvis lejren varede en hel måned. I dette kapitel har vi introduceret nogle af de grundlæggende begreber og koncepter der skal til for at begynde at lave incidensgeometri, og derefter har vi benyttet disse koncepter til at fokusere på de tilfælde hvor man har endelige projektive og affine planer.

Vi har derfor i stor grad haft et fokus på at "tælle ting"; et kombinatorisk perspektiv. Hvis dette underemne af incidensgeometri interesserer læseren, er [2] en skøn tekst, som klart kan anbefales – det kan nok også ses på hvor stor en vidensmæssig gæld vi skylder netop denne bog, når det kommer til sætningerne i den anden halvdel af vores kapitel. Man kan læse store dele af den uden at have nogen anden baggrund inden for universitetsmatematik end den, man får på Matematik Camp, og der er masser af gode eksempler og opgaver. Der er dog dele af bogen, for eksempel [2, Afsnit 3.7, 3.10, 4.6], hvor man skal bruge en vis baggrund inden for lineær algebra.[12] Særligt skal man kende aksiomerne for et vektorrum over et arbitrært legeme, samt nogle grundlæggende begreber og konstruktioner, som underrum og span. Denne viden skal man finde andetsteds. Bemærk dog, at den nødvendige viden om ringteori er at finde i kapitel 3.

Der er masser af typer af incidensstrukturer, som vi slet ikke kom ind på i dette kapitel; vi har valgt en ret snæver tilgang, og stort set kun benyttet os af to slags lineære rum. Der findes mange flere slags geometrier med deres egne spændende egenskaber og nytte – og der findes mange andre resultater om dem, end de kombinatoriske resultater vi har præsenteret.

Ønsker man en mere generel introduktion til incidensgeometri som bredere emne end det, vi har fremlagt, kan [1] anbefales. Den giver et godt overblik over de forskellige områder af incidensgeome-

[12] De kan dog springes over.

tri på kortfattet manér. Den er god til at give et generelt overblik, og præsenterer en masse imponerende, vigtige (og seje!) sætninger indenfor incidensgeometri. Eksempelvis viser den, at alle projektive rum af dimension højere end eller lig med 3 alle kan konstrueres ud fra en helt speciel slags vektorrum, som gør det muligt at betragte dem som mere end blot incidensstrukturen. Det gør det også lettere at klassificere strukturerne op til isomorfi, og sætningen i sin helhed kendes som incidensgeometriens første fundamentalsætning. Bemærk at det i denne bog bliver nødvendigt at have en (i hvert fald overfladisk) baggrund inden for gruppeteori, ringteori og lineær algebra, hvis man ønsker at gå i dybden med denne bog, selvom det er muligt at komme igennem de to første kapitler uden.

Hvor [1] giver relativt få eksempler, giver [6] mange gode eksempler og perspektiver. Det er dog en meget avanceret tekst.

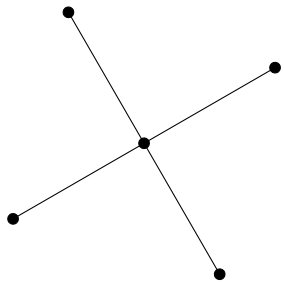
Til sidst vil vi nævne [7], som er en god ressource, hvis man undrer sig over om der findes nogen applikationer af incidensgeometri.

Bemærk at mange tekster (blandt andet [6], [2]) bruger nogle lidt andre aksiomer for projektive og affine planer, end vi gør – vores er mere i stil med [1]. Særligt er PP_3 og AP_3 forskellige fra vores. Vi vil kalde deres for PP_3' og AP_3' (I [6] og [2] er PP_3 "Der findes fire punkter, så der ikke findes 3 af dem, som ligger på samme linje". AP_3 [6] er det samme som PP_3, og i [2] er AP_3 "der findes 3 punkter, som ikke ligger på linje"). Vi påstår dog, at dette ikke gør nogen reel forskel i det projektive tilfælde; at vores PP_3 og deres PP_3 er ækvivalente.[13,14] Her er en idé til hvordan man viser, at $(PP_1, PP_2, PP_3) \Rightarrow PP_3'$:

Bevisidé. Lad **P** være en geometri, som opfylder PP_1, PP_2 og PP_3. Så er der mindst 2 linjer, pr. PP_3 er der mindst 3 punkter på hver linje og pr. PP_2 *skal* de i hvert fald mødes i ét punkt, så vi har som minimum følgende punkter og linjer:

[13]Den arbejdsomme læser ved selvfølgelig allerede dette, da vedkommende klart har lavet alle opgaverne.

[14]Samtidig medfører vores (AP_3 deres AP_3, og givet at deres affine og projektive planer er lineære rum, kan man også vise den anden vej.



Bemærk at dette kun er et udsnit af et billede af et projektivt plan. Der skal for eksempel være linjer gennem alle par af punkter, og der kan sagtens være flere punkter – i tilfældet af at hver linje rent faktisk kun har tre punkter i planet, ved vi jo faktisk præcis hvilket plan denne struktur ville resultere i (Sætning 6.27).

Uanset hvad giver PP_3 og PP_2 os i hvert fald nogle punkter og linjer at arbejde med som minimum. Overvej om du kan finde 4 punkter på figuren, hvoraf der ikke findes 3, som er på linje, uanset hvilket projektivt plan strukturen på billedet er en del af. ■

Se også Opgave 8.27. De andre implikationer opfordres læseren til selv at vise som en hyggelig aktivitet i det tilfælde at læseren får fingrende i et eksempel af [2] og [6].

På samme måde er konceptet om en geometri fremfor en prægeometri ikke til stede i al litteratur (se igen [2], [6]). Bemærk dog at lineære rum er geometrier på baggrund af deres aksiomer alene – så vi behøvede faktisk ikke at kræve, at de var geometrier i første omgang. Overvej hvorfor.

8 Opgaver

- **Opgave 8.1:**
  Diskuter følgende begreber med en makker.

  1) Prægeometri.

  2) Incidens.

  3) Flag.

  4) Geometri.

- **Opgave 8.2:**
  Er det muligt at lave en geometri over $\{punkt, linje\}$ med 4 punkter og 2 linjer? Hvis ja, tegn én.

- **Opgave 8.3:**
  Tegn en prægeometri over typemængden $\{punkt, linje\}$ med 6 punkter, 7 linjer og hvor ingen linjer har præcis 4 punkter.

- **Opgave 8.4:**
  Lad Γ være en prægeometri over typemængden I, og lad F være et flag i Γ. Vis følgende:

  1) Mængderne F og $\mathbf{type}(F)$ har samme kardinalitet, altså $|F| = |\mathbf{type}(F)|$.

  2) Hvis F er et maksimalt flag, så gælder $|F| = |I|$.

- **Opgave 8.5:**
  Bevis følgende udsagn.

  Lad Γ være en geometri med endelig rang n over typemængden I. Et flag i Γ er co-maksimalt hvis og kun hvis $|\mathbf{type}(F)| = n - 1$.

- **Opgave 8.6:**
  Diskuter følgende begreber med en makker.

  1) Nær-lineært rum.

  2) Lineært rum.

•• **Opgave 8.7:**
Lad ℓ_1 og ℓ_2 være linjer i et nær-lineært rum. Vis, at antallet af linjer der skærer både ℓ_1 og ℓ_2 er mindre eller lig med $v_1 \cdot v_2$.

• **Opgave 8.8:**
Vis, at to forskellige linjer i et nær-lineært rum højst kan skære i ét punkt.

• **Opgave 8.9:**
Lad $\mathcal{L}$ være et lineært rum. Vi betragter linjerne i $\mathcal{L}$ som mængderne af de punkter, de inciderer med. Vis, at hvis ℓ_1 og ℓ_2 er linjer, så $\ell_1 \subseteq \ell_2$, så har vi at $\ell_1 = \ell_2$.

•• **Opgave 8.10:**
Tænk tilbage på vores definition af lineære rum. Er det nødvendigt at antage, at $\mathcal{L}$ er en geometri?

• **Opgave 8.11:**
Lad $\mathcal{L}$ være et lineært rum. Vis, at følgende er underrum af $\mathcal{L}$:

1) Den tomme mængde.

2) Et enkelt punkt.

3) En enkel linje.

4) Mængden af alle punkter i $\mathcal{L}$.

• **Opgave 8.12:**
Er det muligt at konstruere et lineært rum med et uendeligt antal punkter, men et endeligt antal linjer? Begrund dit svar.

•• **Opgave 8.13:**
Vis, at spannet i et lineært rum opfylder følgende egenskaber:

1) $X \subseteq \langle X \rangle$.

2) $\langle X \rangle = \langle \langle X \rangle \rangle$.

3) $X \subseteq Y \Rightarrow \langle X \rangle \trianglelefteq \langle Y \rangle$.

- **Opgave 8.14:**
  Diskutér følgende begreber med en makker:

  1) Underrum.

  2) Span.

  3) Hyperplan.

- **Opgave 8.15:**
  Lad $\mathcal{L}$ være et lineært rum med 5 punkter. Hvor mange linjer kan der højst være i $\mathcal{L}$?

- **Opgave 8.16:**
  Giv et eksempel på et nær-lineært rum, som har mindst en linje med uendeligt mange punkter, og mindst en linje med endeligt mange punkter.

- **Opgave 8.17:**

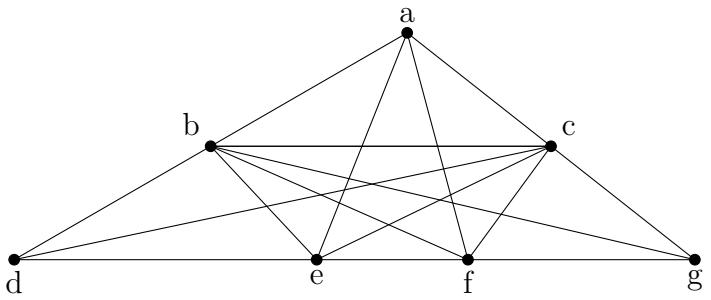
  1) Find alle nær-lineære rum med 4 punkter.

  2) Find alle lineære rum med 4 punkter.

•• **Opgave 8.18:**
  Lad $\mathcal{N}$ være et nær-lineært rum med v punkter, og hvor hver linje inciderer med præcis 3 punkter. Hvad er det maksimale antal linjer, som et givet punkt kan ligge på?

•• **Opgave 8.19:**
  Hvis et nær-lineært rum har v punkter, hvad er det højest mulige antal linjer, rummet kan have? Forklar din tankegang.

•• **Opgave 8.20:**
Find hyperplanerne i det lineære rum.



Med linjerne $\{a,b,d\}$, $\{a,c,g\}$, $\{a,e\}$, $\{a,f\}$, $\{b.c\}$, $\{b,e\}$, $\{b,f\}$, $\{c,d\}$, $\{c,e\}$, $\{c,f\}$, $\{d,e,f,g\}$.

• **Opgave 8.21:**
Diskuter følgende begreber med en makker.

1) Projektivt plan.

2) Dualudsagn.

3) Projektiv lukning.

4) Ækvivalensrelation.

5) Parallelklasse.

• **Opgave 8.22:**
Diskuter med en makker:

1) Hvordan er projektive planer forskellige fra almene lineære rum?

2) Hvordan er affine planer forskellige fra almene lineære rum?

•• **Opgave 8.23:**
Ligesom vi definerede dualudsagnet af et udsagn (Definition 4.4), kan vi også definere dualen af et nær-lineært rum på følgende måde.

Lad $\mathcal{L}$ være et nær-lineært rum med punkter P og linjer L. Det duale nær-lineære rum er rummet $\mathcal{L}'$, som har linjerne i $\mathcal{L}$ som punkter. Linjerne i $\mathcal{L}'$ er defineret på følgende måde:

Lad $p \in P$ være et punkt i $\mathcal{L}$. Hvis $b(p) \geq 2$, så er mængden af alle linjer gennem p en linje i $\mathcal{L}'$. Linjer opnået på denne måde er de eneste linjer i $\mathcal{L}'$. Da punkterne i $\mathcal{L}'$ er linjerne i $\mathcal{L}$, er linjerne i $\mathcal{L}'$ altså mængder af punkter i $\mathcal{L}'$, og vi kan derfor definere incidens på den naturlige måde.

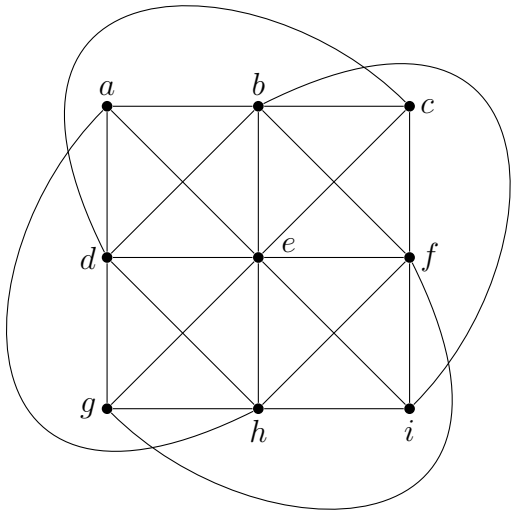
1) Vis, at dualen af et nær-lineært rum er lineært.

2) Vis, at dualen af et lineært rum ikke nødvendigvis er lineært.

3) Afgør hvad der skal til for at dualen af et givet lineært rum er lineært.

•• **Opgave 8.24:**
Find et nær-lineært rum, som er sin egen dual.

•• **Opgave 8.25:**
Betragt geometrien **A**



Med linjerne $\{a, b, c\}$, $\{d, e, f\}$, $\{g, h, i\}$, $\{a, d, g\}$, $\{b, e, h\}$, $\{c, f, i\}$, $\{h, d, c\}$, $\{a, e, i\}$, $\{b, f, g\}$, $\{d, b, i\}$, $\{g, e, c\}$, $\{a, h, f\}$.

1) Er **A** et affint plan?

2) Hvilke parallelklasser har $\mathbf{A}$?

3) Konstruer $\mathbf{P}(\mathbf{A})$.

4) Hvor mange punkter har $\mathbf{P}(\mathbf{A})$?

••• **Opgave 8.26:**
Vis, at enhver linje i et affint plan må incidere det samme antal punkter – mere specifikt er der for ethvert par af linjer ℓ og m en bijektion mellem deres punkter.

••• **Opgave 8.27:**
Vis, at en geometri $\mathbf{P}$ over typemængden $\{punkt, linje\}$ er et projektivt plan hvis og kun hvis den opfylder PP_1, PP_2 og det modificerede aksiom PP_3':

PP_3' Der findes fire punkter, hvoraf der ikke er tre, som ligger på samme linje.

•• **Opgave 8.28: Möbius-planerne**
Et *Möbius-plan* er en prægeometri μ over $\{punkt, cirkel\}$, som opfylder følgende aksiomer.

(IP_1) Hvis x,y,z er tre punkter i μ, så findes der præcis en cirkel C, så alle tre punkter inciderer med C.

(IP_2) Hvis x,y er punkter og C er en cirkel, som indeholder x men ikke y, så findes der præcis én cirkel C', som indeholder y, således at C og C' har x og kun x til fælles.

(IP_3) Hver cirkel inciderer med mindst 3 punkter. Der findes 4 punkter, som ikke alle ligger på samme cirkel.

Vi betragter en cirkel som værende mængden af de punkter, som den inciderer med.

1) Lad μ være et Möbius-plan, hvor mængden af punkter kaldes $\mathcal{P}$ og mængden af cirkler kaldes $\mathcal{C}$. Lad desuden $x \in P$ Definér pregeometrien A over $\{punkt, linje\}$ således:

- Punkterne i A er $P \setminus \{x\}$.
- Linjerne er som følger: Hvis C er en cirkel i μ som indeholder x, så er $C \setminus \{x\}$ en linje i A.

2) Vis, at A er et affint plan.

3) Konstruer et Möbius-plan med så få punkter som muligt.

4) Konstruer et affint plan ud fra Möbius-planet konstrueret ovenfor.

- **Opgave 8.29:**
  Konstruer et projektivt plan med præcis 13 punkter. (Hint: hvor mange punkter skal være på hver linje?)

••• **Opgave 8.30: Moulton-planet**
Vi definerer for alle reele tal m,x følgende operation:

$$m \odot x := \begin{cases} m \cdot x, & \text{hvis } m \leq 0 \text{ eller } x \leq 0, \\ 2 \cdot m \cdot x, & \text{hvis } m > 0 \text{ og } x > 0. \end{cases}$$

Lad $P = \mathbb{R}^2$. Lad $\Lambda = \{\{x\} \times \mathbb{R} \mid x \in \mathbb{R}\}$ være de lodrette linjer i $\mathbb{R}^2$, som i Eksempel 3.6. Husk, at hvis $f : \mathbb{R} \to \mathbb{R}$ er en funktion, så kaldes mængden $G_f = \{(x,f(x)) \mid x \in \mathbb{R}\} \subseteq \mathbb{R}^2$ *grafen* for f. Vi kalder mængden af graferne af alle funktioner $f : \mathbb{R} \to \mathbb{R}$ hvor $f(x) = m \odot x + b$ med m,$b \in \mathbb{R}$ for $\mathcal{G}$. Lad slutteligt $L = \mathcal{G} \cup \Lambda$, og $X = P \cup L$.

Moulton-planet er prægeometrien $M = (X,\mathbf{type},*)$ over $I = \{punkt, linje\}$, hvor typefunktionen er givet ved

$$\mathbf{type}(x) = \begin{cases} punkt, & x \in P, \\ linje, & x \in L, \end{cases}$$

og hvor incidens er givet ved at hvis $x \in P$,$\ell \in L$, så har vi $x * \ell$ hvis og kun hvis $x \in \ell$. To elementer af samme type inciderer hvis og kun hvis de er det samme element.

1) Er Moulton-planet et affint plan?

• **Opgave 8.31:**
Diskuter følgende begreber med en makker.

1) Lineær funktion.

2) Isomorfi.

3) Kollination.

•• **Opgave 8.32:**
Betragt Fano planet fra Eksempel 4.3.

Sætning 4.13 giver os en måde at lave affine planer ud af Fano planet, ved at fjerne en linje. Er det ligegyldigt hvilken linje, vi fjerner? Altså, er de forskellige affine planer, vi opnår ved at fjerne en enkelt linje isomorfe med hinanden?

••• **Opgave 8.33:**
Lad $\mathcal{L}$ og $\mathcal{M}$ være nær-lineære rum, $U \trianglelefteq \mathcal{L}$ være et underrum, og λ være en lineær funktion fra $\mathcal{L}$ til $\mathcal{M}$. Vis, at $\lambda(U)$ ikke nødvendigvis er et underrum i $\mathcal{M}$.

Vis, at hvis $\mathcal{L}$ og $\mathcal{M}$ begge er lineære rum, så skal $\lambda(U)$ være et underrum i $\mathcal{M}$.

•• **Opgave 8.34:**
Lad V være et underrum af et lineært rum $\mathcal{M}$ og lad $\mathcal{L}$ være et andet lineært rum. Lad λ være en injektiv lineær funktion fra $\mathcal{L}$ til $\mathcal{M}$. Vis, at $\lambda^{-1}(V)$ er et underrum af $\mathcal{L}$.

••• **Opgave 8.35:**
Lad λ være en injektiv lineær funktion fra det lineære rum $\mathcal{L}$ til det lineære rum $\mathcal{M}$.

1) Vis, at hvis U er et underrum af $\mathcal{L}$, så er $\lambda^{-1}(\lambda(U)) = U$.

2) Vis, at hvis V er et underrum af $\mathcal{M}$, så er $\lambda(\lambda^{-1}(V)) \subseteq V$.

3) Vis, at hvis λ ydermere er bijektiv, så er $\lambda(\lambda^{-1}(V)) = V$.

Bemærk, at de sidste par opgaver har givet os følgende lemma.

Lemma 8.1 ([2, Lemma 2.6.7])**.** Lad $\mathcal{L}$,$\mathcal{M}$ være lineære rum, og X være en punktmængde i $\mathcal{L}$. Lad desuden λ være en injektiv lineær funktion fra $\mathcal{L}$ til $\mathcal{M}$.

a) For ethvert underrum $U \trianglelefteq \mathcal{L}$, som inderholder X, så er $\lambda(X) \subseteq \lambda(U) \trianglelefteq \mathcal{M}$. Desuden, hvis $V \trianglelefteq \mathcal{M}$, og V indeholder $\lambda(X)$, så er $X \subseteq \lambda^{-1}(U) \trianglelefteq \mathcal{L}$.

b) Ydermere, hvis λ er bijektiv, så findes der en bijektion mellem underrummene af $\mathcal{L}$, som indeholder X og underrummene af $\mathcal{M}$, som indeholder $\lambda(X)$.

•• **Opgave 8.36:**
Brug ovenstående lemma til at vise, at hvis $\mathcal{L}$,$\mathcal{M}$ er lineære rum, og X er en punktmængde i $\mathcal{L}$, mens λ er en lineær funktion fra $\mathcal{L}$ til $\mathcal{M}$, så er $\langle\lambda(X)\rangle = \lambda(\langle X\rangle)$.

• **Opgave 8.37:**
Lad $P = \{p_1,p_2,p_3,p_4,p_5\}$ og

$$L = \{\{p_1,p_2\},\ \{p_1,p_4\},\ \{p_4,p_5\},\ \{p_2,p_5\},\ \{p_2,p_3,p_4\},\ \{p_1,p_3,p_5\}\}.$$

Lad desuden $Q = \{a,b,c,d\}$ og $M = \{\{a,b\},\{a,c\},\{a,d\},\{b,c,d\}\}$. Lad $\mathcal{L}$ være det nær-lineære rum med punkter P og linjer L, mens M er det nær-lineære rum med punkter Q og linjer M. Her defineres incidens i begge tilfælde så et punkt inciderer med en linje hvis og kun hvis punktet er et element i linjen. Definer funktionen $f : P \to Q$ ved $f(p_1) = a$, $f(p_2) = b$, $f(p_3) = c$, $f(p_4) = d$, $f(p_5) = a$.

Og nu til den faktiske opgave: Er f lineær?

•• **Opgave 8.38:**
Lad S være et lineært rum, lad S' være et nær-lineært rum, og lad $f : S \to S'$ være en lineær funktion. Vis, at S' er et lineært rum.

•• **Opgave 8.39:**
Lad $\mathcal{L}$ være et nær-lineært rum. Antag at hvert punkt inciderer med r linjer, mens enhver linje inciderer med k punkter. Vis, at

$$v \cdot r = b \cdot k.$$

• **Opgave 8.40:**
Lad $\mathcal{L}$ være et lineært rum som opfylder $(b - v)^2 \leq v$.
1) Find alle lineære rum af denne slags med $v = 6$.

•• **Opgave 8.41:**
Vis, at Sætning 6.20 ikke nødvendigvis er sandt, hvis vi i stedet for at arbejde med et lineært rum, arbejder med et nær-lineært rum.

•• **Opgave 8.42:**
Find et eksempel på nær-lineære rum $\mathcal{L}$ og $\mathcal{M}$ og en injektiv lineær funktion λ fra $\mathcal{L}$ til $\mathcal{M}$, hvor der findes et punkt p i $\mathcal{L}$ sådan at $b(p) \neq b(\lambda(p))$. Hvorfor modstrider dette ikke Proposition 6.15?

•• **Opgave 8.43:**
Vis, at et nær-lineært rum med $v \geq b \geq 1$ ikke nødvendigvis er et lineært rum.

• **Opgave 8.44:**
Betragt følgende incidensmatrix.

$$\begin{array}{c} \\ p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{array} \begin{array}{c} \begin{array}{cccc} \ell_1 & \ell_2 & \ell_3 & \ell_4 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

Er dette et lineært rum? Konstruer dualen.

•• **Opgave 8.45:**
Betragt nedenstående incidensmatricer.

$$\begin{array}{c} \\ p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{array} \begin{array}{c} \begin{array}{cccccccc} \ell_1 & \ell_2 & \ell_3 & \ell_4 & \ell_5 & \ell_6 & \ell_7 & \ell_8 \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

$$\begin{array}{c} \\ p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \\ p_7 \end{array} \begin{array}{c} \begin{array}{ccccccc} \ell_1 & \ell_2 & \ell_3 & \ell_4 & \ell_5 & \ell_6 & \ell_7 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

Hvilke slags strukturer repræsenterer matricerne?

Det kunne for eksempel være affine/projektive planer, (nær)-lineære rum, etc.

• **Opgave 8.46:**
Findes der et projektivt plan med 15 punkter?

••• **Opgave 8.47:**
Er det muligt at konstruere et affint plan med et uendeligt antal punkter og et endeligt antal linjer? Begrund dit svar.

••• **Opgave 8.48:**
Lad $\mathbf{A}$ være et affint plan, hvor hver linje har n punkter for $n \in \mathbb{N}$. Vis følgende sætning.

Sætning 8.2 ([6, Sætning 2.4])**.** Hvert punkt i **A** ligger på præcis $n+1$ linjer, som representerer hver sin parallelklasse. Desuden består hver parallelklasse af n linjer, som partitionerer punkterne i **A**.[15] Særligt har **A** præcis n^2 punkter og $n^2 + n$ linjer.

Bevisidé. Dette bevis er lidt af en mundfuld, så vi giver en guide, punkt-for punkt.

1) Vi starter med at vise at hver parallelklasse består af n linjer. Lad ℓ være en linje. Da kan vi finde en linje, som ikke er parallel til ℓ (Hvorfor?). Kald denne linje m. m har n punkter. Hvad kan vi bruge de punkter til?

2) Nu vi har argumenteret for at der er n linjer i hver parallelklasse, hvorfor udgør $\pi(\ell)$ en partitionering af punkterne i **A**?. Hvis vi har et punkt x i **A**, kan vi så finde en linje parallel til ℓ gennem x?

3) Vi er nu nået til at vise at der findes n^2 punkter. Dette er mere ligetil end det første skridt. Argumentér for, at der skal findes mindst én linje ℓ i **A**. $\pi(\ell)$ partiotionerer **A**. Hvad kan vi bruge det til?

4) Vi er nået til at vise at der går $n + 1$ linjer gennem hvert punkt. Lad x være et punkt i **A**. Vi ved at der er $n^2 - 1$ punkter i **A**, og at hvert punkt $y \neq x$ giver os en unik linje xy med n punkter. Vi kan nu finde $b(x)$.

5) Vi mangler til sidst at vise at **A** har $n^2 + n$ linjer. Bemærk at ethvert punkt *skal* have en linje fra hver parallelklasse igennem sig. ■

[15] Altså hvis ℓ er en linje, så er alle punkter i **A** indeholdt i en af linjerne fra $\pi(\ell)$. Desuden er der selvfølgelig intet punkt i **A**, som ligger på to linjer i $\pi(\ell)$

9 Projekt: Kollineardistance

Vi har i løbet af dette kapitel kastet nørklede og unødvendige begreber, såsom vinkler, distance, og position ud af vinduet. Som vi har vist, er det trods alt muligt at lave rigtig meget geometri, selv uden overhovedet at gøre brug af disse koncepter...

Men hvad nu hvis vi havde lyst til at genindføre distancebegrebet, bare for hyggens skyld? Dette projekt kommer til at handle om hvad man kan gøre for at indføre et begreb om distance i et vilkårligt nær-lineært rum, selv hvis den struktur man arbejder med ikke har noget indbygget distancebegreb.

Første ting på dagsordenen er at give en præcis definition af hvad vi mener med *distance*.

Definition 9.1 (Metrik)**.** Lad X være en vilkårlig mængde. En *metrik* på X er en funktion $d : X \times X \to [0,\infty)$, som opfylder følgende aksiomer.

M_1 $d(x,y) = 0$ hvis og kun hvis $x = y$,

M_2 $d(x,y) = d(y,x)$ for alle $x,y \in X$,

M_3 $d(x,z) \leq d(x,y) + d(y,z)$ for alle $x,y,z \in X$.

Det sidste aksiom kendes som trekantsuligheden. Vi kalder (X,d) for et *metrisk rum*.

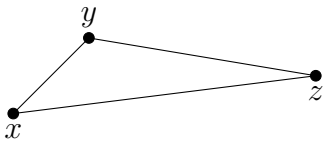
Bemærkning 9.2. En metrik er essentielt set en afstandsfunktion. Den giver en måde at måle afstanden mellem punkter på. Derfor er definitionen af en metrik også langt fra tilfældig. Hvert af aksiomerne er med til at sørge for at en metrik altid vil opfylde de krav, vi intuitivt set ville have for en afstandsfunktion. Vi gennemgår aksiomerne, og hvad de betyder.

M_1 Intuitivt set vil vi gerne have, at ethvert punkt har afstand 0 til sig selv. Det giver til gengæld ingen mening at to punkter, som

er forskellige skulle have afstand 0. M_1 sikrer os at metrikken opfylder disse forventninger.

M_2 M_2 fortæller os, at der skal være lige så langt fra x til y, som der er fra y til x. Det ville være svært at kalde $d(x,y)$ en afstandsfunktion, hvis afstanden mellem to punkter afhænger af hvilket punkt man starter med at måle ved.

M_3 Trekantsuligheden kommer til at virke meget naturlig når man overvejer følgende diagram.



M_3 fortæller os bare at afstanden fra x til z burde være kortere end eller lige så lang som afstanden fra x til y lagt sammen med afstanden fra y til z. Det giver god mening, da turen fra x til y til z er lidt af en omvej.

Eksempel 9.3 (Den Euklidiske metrik d_2)**.** Lad $X = \mathbb{R}^2$. Definér $d_2 : \mathbb{R}^2 \times \mathbb{R}^2 \to [0,\infty)$ således.

- Hvis $(a,b),(c,d) \in \mathbb{R}^2$, har vi

$$d_2((a,b),(c,d)) := \sqrt{(a-c)^2 + (b-d)^2}.$$

Vi kalder denne metrik den *Euklidiske metrik*, og det er nok den læseren er vant til at bruge til at måle afstande i $\mathbb{R}^2$. ∘

Eksempel 9.4. Lad $X = \mathbb{R}$. Funktionen $d_s : \mathbb{R} \times \mathbb{R} \to [0,\infty)$ givet ved

$$d_s(a,b) = |a-b|$$

er en metrik. ∘

Opgave 9.1:
Verificer at d_s er en metrik på $\mathbb{R}$.

Nu ved vi hvad en metrik er, så nu mangler vi at finde på en måde at give et vilkårligt nær-lineært rum en metrik på.

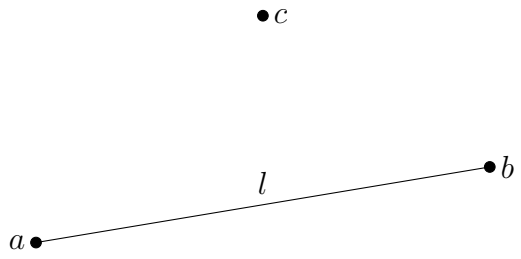
Definition 9.5 (Kollineardistance)**.** Lad $\mathcal{L} = (X,\mathbf{type},*)$ være et nær-lineært rum.

Vi definerer en metrik d_c på $\mathbf{type}^{-1}(\mathit{punkt})$ induktivt, i stor grad på samme måde som Erdős-tallet.

- Hvis $x = y$, så er $d_c(x,y) = 0$.
- Hvis $x \neq y$, men x og y ligger på samme linje, er $d_c(x,y) = 1$.
- Hvis $d(x,y) \neq k$ for alle $k \in \{0,\ldots,n\}$, hvor $n \in \mathbb{N}$, men y ligger på linje med et punkt z med $d(x,z) = n$, så er $d(x,y) = n + 1$.

Vi kalder d_c for *kollinearmetrikken.*

Bemærkning 9.6. $d_c(x,y)$ giver i store træk et mål for hvor langt x er fra at ligge på linje med y. Bemærk at vi støder på et problem. Hvad nu hvis vi har et nær-lineært rum, hvor ikke alle punkter er forbundne med hinanden gennem en række af linjer? Tag for eksempel det nær-lineære rum fra Eksempel 2.3.



Hvad skal $d_c(a,c)$ være? Der er et par måder at tilgå problemet på. Som med Erdős-tallet kunne man vælge at sige, at $d_c(a,c) = \infty$, men så ville d_c ikke være en metrik med mindre vi udvider vores definition

(hvorfor?). En anden tilgang er kun at tage højde for de *forbundne* nær-lineære rum. Det er de nær-lineære rum, hvor $d_c(x,y) < \infty$ for alle punkter x,y – altså alle de nær-lineære rum hvor vi kan finde en "sti" af linjer, som forbinder x til y, uanset hvilket x og y vi vælger. Det er den tilgang, vi vælger, så vi rent faktisk har en vaskeægte metrik at arbejde med.

Opgave 9.2:
Vis, at kollinearmetrikken, med forbeholdene fra bemærkning 9.6 er en metrik.

Opgave 9.3:
Lad $\mathcal{L}$ være et lineært rum. Bestem $d_c(x,y)$ for alle punkter x og y.

Tillykke! Nu kan vi måle afstand mellem punkter i et nær-lineært rum! Alle vores idealer om en distanceløs verden er blevet revet fra os, og der er ikke noget vi kan gøre ved det. Vores metrik på nær-lineære rum er dog lidt spøjs. Den minder for eksempel ikke rigtig på nogen måde om d_2 fra Eksempel 9.3. Med d_2 kan afstande mellem punkter være alle mulige spøjse reelle tal, eksempelvis er $d_2((0,0),(0,\pi)) = \pi$, men d_c kan kun antage heltalsværdier! Hvis vi betragter $\mathbb{R}^2$ som et metrisk rum, både med hensyn til d_c og d_2, så *må* vi altså næsten kunne se en meningsfuld forskel af en art.[16] Vi skal nu forsøge at prikke lidt til de to metrikker for at se, om vi kan finde sådan en forskel (ud over de åbenlyse).

Definition 9.7 (Følge)**.** En *følge* $(x_n)_{n\in\mathbb{N}}$ af elementer i en mængde X er en funktion $\mathbb{N} \to X$. Altså er det noget, som for hvert naturligt tal n giver os et element $x_n \in X$. x_n er altså det n'te element af følgen $(x_n)_{n\in\mathbb{N}}$.

Eksempel 9.8. Den uendelige talfølge

$$1,\frac{1}{2},\frac{1}{3},\ldots = \left(\frac{1}{n}\right)_{n\in\mathbb{N}}$$

[16]For at kunne bruge d_c skal vi bruge et nær-lineært rum. Det nær-lineære rum, vi bruger til at definere d_c på $\mathbb{R}^2$ er det euklidiske plan E fra Eksempel 3.6.

er en følge i $\mathbb{R}$.

$$(0,1),(0,2),(0,3),\ldots = ((0,n))_{n\in\mathbb{N}}$$

er en følge i $\mathbb{R}^2$. ∘

Definition 9.9 (Konvergens)**.** Lad (X,d) være et metrisk rum. En følge $(x_n)_{n\in\mathbb{N}}$ siges at *konvergere* til et punkt $x \in X$ med hensyn til d hvis og kun hvis følgende gælder.

> For alle $\varepsilon > 0$ findes der et $N \in \mathbb{N}$ så $d(x_n,x) < \varepsilon$ så længe $n \geq N$.

Vi skriver $\lim_{n\to\infty} x_n = x$.

Bemærkning 9.10. Definitionen ovenfor siger at hvis vi har at $\lim_{n\to\infty} x_n = x$, så betyder det at uanset hvor lille en fejlmargen vi får, så kan vi finde et $N \in \mathbb{N}$, som sørger for at alle vores følgeelementer efter punktet x_N ligger inden for denne fejlmargen af x. Bemærk at hvorvidt en følge konvergerer lige så meget er et spørgsmål om den metrik man bruger, som følgen man har med at gøre.

Eksempel 9.11. Den konstante følge $(0,0),(0,0),\ldots = ((0,0))_{n\in\mathbb{N}}$ konvergerer i $\mathbb{R}^2$ med hensyn til en hvilken som helst metrik. ∘

Eksempel 9.12. Følgen $-1,1,-1,\ldots = (-1^n)_{n\in\mathbb{N}}$ konvergerer *ikke* i $\mathbb{R}$ med hensyn til d_s fra Eksempel 9.4: Lad $\varepsilon = \frac{1}{2}$. Hvis $d_s(x,-1) < \varepsilon$, så er $d_s(1,-1) \leq d_s(1,x) + d_s(-1,x)$ pr. trekantsuligheden, så altså $2 \leq d_s(1,x) + d_s(-1,x)$, sådan at $d_s(1,x) \geq 2 - d_s(-1,x) > \frac{3}{2}$, da $d_s(-1,x) < \varepsilon$. Vice versa gælder hvis $d_s(x,1) < \varepsilon$. Da både -1 og 1 optræder uanset hvor langt ud i følgen vi går, kan vi ikke finde et $N \in \mathbb{N}$, så alle punkter efter x_n er tættere end ε på x, uanset hvilket punkt $x \in \mathbb{R}$ er. ∘

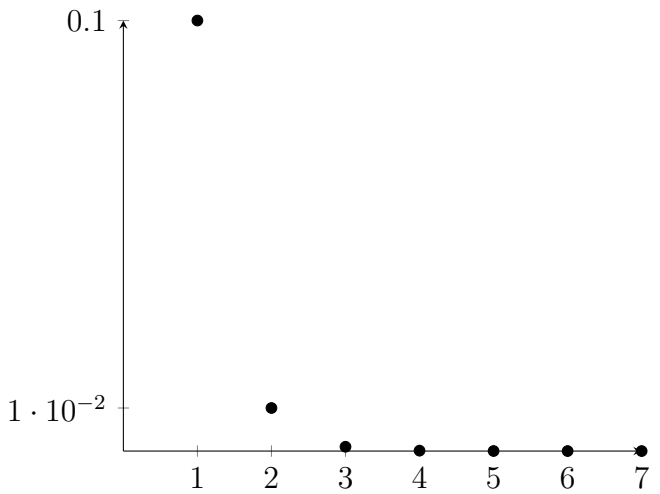
Sætning 9.13. En følge $((x_n,y_n))_{n\in\mathbb{N}}$ i $\mathbb{R}^2$ konvergerer til punktet (x,y) med hensyn til d_2 hvis og kun hvis følgen $(x_n)_{n\in\mathbb{N}}$ konvergerer til $x \in \mathbb{R}$ med hensyn til d_s, og $(y_n)_{n\in\mathbb{N}}$ konvergerer til $y \in \mathbb{R}$ med hensyn til d_s.

Opgave 9.4:
Bevis Sætning 9.13.

Nu skal vi se en af de store forskelle mellem d_c og d_2, netop hvilke følger konvergerer med hensyn til de to forskellige metrikker.

Opgave 9.5:
Vis, at en følge $((x_n,y_n))_{n\in\mathbb{N}}$ i $\mathbb{R}^2$ konvergerer med hensyn til d_c hvis og kun hvis den i sidste ende er konstant, altså hvis der findes et $N \in \mathbb{N}$ sådan at (x_n,y_n) konstant er lig (a,b) for et eller andet talpar $(a,b) \in \mathbb{R}^2$ og for alle naturlige tal $n \geq N$.

Opgave 9.6:
Vis, at følgen $(\frac{1}{10^n},\frac{1}{10^n})_{n\in\mathbb{N}}$ konvergerer til (0,0) i $\mathbb{R}^2$ med hensyn til d_2, men ikke konvergerer med hensyn til d_c.
(Hint: Vis, at $\lim_{n\to\infty} 1/(10^n) = 0$ i $\mathbb{R}$ med hensyn til d_s. Se desuden Figur 1.9).



Figur 1.9: Her ses de første 7 elementer af følgen $(1/(10^n))_{n\in\mathbb{N}}$ med n på x-aksen og $1/(10^n)$ på y-aksen.

Vi kan fra ovenstående opgaver konkludere, at d_c er ret streng med hensyn til konvergens. Hvis en følge i en intuitiv forstand "lig-

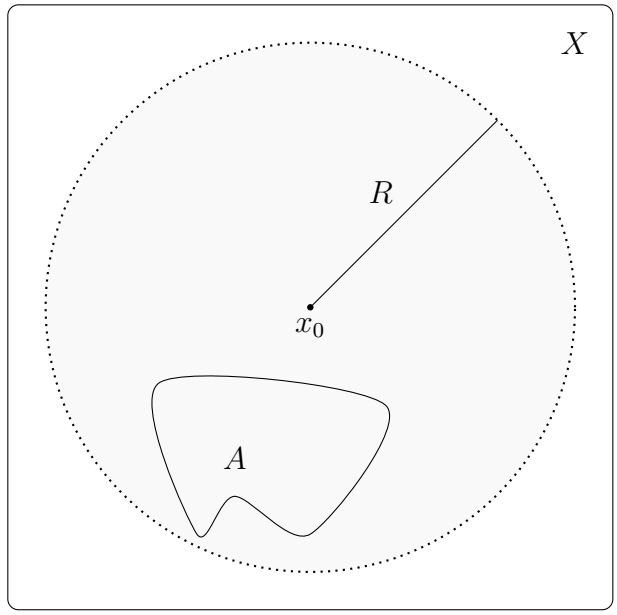
ner", at den nærmer sig et punkt $(x,y) \in \mathbb{R}^2$, så er der god sandsynlighed for at den konvergerer med hensyn til d_2. d_2 fanger altså rimelig godt det, vi intuitivt vil tænke på konvergens som. d_c er *meget* strengere og tillader kun konvergens hvis følgen i sidste ende bare bliver på det samme punkt.

Vi skal nu se en anden måde d_c er lidt mærkelig på.

Definition 9.14. Lad (X,d) være et metrisk rum. Vi siger at en delmængde $A \subseteq X$ er *begrænset* med hensyn til d hvis og kun hvis der findes et $x_0 \in X$ og et $R > 0$ så $d(x_0,a) < R$ for alle $a \in A$. Man kan tænke på det som at alle punkter i A ligger i en *åben kugle* af radius R med centrum i x_0, altså mængden

$$B_d(x_0,R) = \{x \in X \mid d(x,x_0) < R\}.$$

Se Figur 1.10. Vi siger, at (X,d) er begrænset hvis X er begrænset som delmængde af sig selv.



Figur 1.10: A ligger i $B_d(x_0,R)$. Bemærk at kugler ikke behøver være runde.

Bemærkning 9.15. Grunden til at vi kalder $B_d(x_0,R)$ en kugle er, at hvis man betragter $\mathbb{R}^2$ med hensyn til d_2, så er kuglen rund! Man kan desuden udvide d_2 til $\mathbb{R}^3$, hvor den åbne kugle rent faktisk er en kugle i 3D-rum.

Intuitivt set ville det give mening hvis $\mathbb{R}^2$ er ubegrænset – det er trods alt et uendeligt plan, så man skulle tro, at det var uendeligt stort. Det påstår vi også at det er, når vi bruger d_2. Det kan vi dog ikke tage for givet, så det bliver læserens opgave at verificere dette.

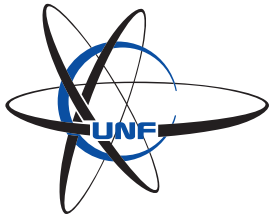
Opgave 9.7:
Vis, at $(\mathbb{R}^2,d_2)$ ikke er begrænset.

Hvis vi bruger d_c får vi til gengæld helt andre resultater...

Opgave 9.8:
Vis, at $(\mathbb{R}^2,d_c)$ er begrænset.

Der er mange andre sammenligninger at drage mellem d_c og d_2, som yderligere illustrer hvor fjollet d_c kan virke, og læseren er velkommen til at kontakte rhh eller syhe fysisk eller over mail hvis de vil høre mere, men dette er enden på vores projekt.

Vi har nu set en måde at genindføre afstandsbegrebet på, givet et sammenhængende nær-lineært rum, og i mellemtiden vist at der faktisk findes andre måder at måle afstand på i $\mathbb{R}^2$ end d_2, som vi ellers normalt bruger – og ikke nok med at der findes andre måder at gøre det på, så er der faktisk en meningsfuld forskel på metrikkerne, så man i virkeligheden får to forskellige perspektiver på eksempelvis konvergens.



2 Talteori

1 Introduktion

Generelt er talteori, som navnet indikerer, teorien om tal - i særlig grad, de hele tal. Hvilke seje egenskaber har de forskellige tal? Hvad kan vi sige om disse egenskaber? Og hvor mange tal har de samme egenskaber? Dette er nogle af de spørgsmål der ligger til grund for talteori og har fascineret matematikere i mange tusind år.

Du kender nok allerede nogle specielle typer af tal; lige tal, primtal eller måske kvadrattal. Der er mange forskellige typer af tal og de har allesammen en egenskab, der gør dem interessante.

Denne gren af matematikken har optaget flere af tidens største matematikere. Du har måske hørt om Euklid, Fermat og Euler, og der er mange flere.

"Matematik er dronningen af videnskab, og talteori er dronningen af matematik." – Carl Friedrich Gauss

2 Delelighed

Definition 2.1 (Delelighed)**.** For hele tal d og n, siger vi at d er en *divisor* i n hvis der findes et heltal k således at

$$n = d \cdot k. \tag{2.1}$$

Vi skriver $d \mid n$ når vi noterer at d deler n.

Sætning 2.2. Lad a, b og c være hele tal. Da gælder følgende regler for divisorer:

1) Hvis $a \mid b$ og $b \mid c$, så vil $a \mid c$.

2) Hvis $a \mid b$, så vil $a \mid bc$.

3) Hvis $a \mid b$ og $a \mid c$, så vil $a \mid mb + nc$, for vilkårlige hele tal, m og n.

Bevis. Vi giver et bevis for regel 1). Resten vises i Opgave 6.4. Når $a \mid b$ kan vi finde d så $b = da$. Ligeledes, når $b \mid c$ kan vi finde d' så $c = d'b$. Sammen giver det

$$c = d'b = d'da \tag{2.2}$$

og ergo er a divisor i c. ■

Definition 2.3 (Primtal)**.** Et helt tal $p > 1$ kaldes for et primtal hvis det ikke har andre positive divisorer end 1 og p. Et helt tal $n > 1$ som ikke er et primtal kaldes for et sammensat tal.

Sætning 2.4. Ethvert naturligt tal kan skrives som et produkt af primtal.

Bevis. Lad n være et naturligt tal større end 1. Lad p_1 være den mindste divisor større end 1.

Antag for modstrid at p_1 er et sammensat tal. Da har p_1 en divisor d hvor $1 < d < p_1$. d må nødvendigvis også være divisor i n, hvilket er i modstrid med at p_1 var mindste divisor.

Derfor må p være et primtal og vi kan skrive $n = p_1 q_1$. Hvis $q_1 = 1$ kan vi stoppe. Hvis ikke kan vi bruge samme argument til at finde primtallet p_2 så $n = p_1 p_2 q_2$. Eftersom $q_1, q_2, \ldots$ er en aftagende følge af hele tal, vil vi før eller senere ramme 1. Så n er skrevet som et produkt, udelukkende af primtal. ■

Sætning 2.5 (Euklids Lemma)**.** Hvis p går op i produktet ab, så går p enten op i a eller b.

Denne sætning bevises til sidst i dette afsnit, men for nu antager vi at det er sandt.

Sætning 2.6 (Aritmetikkens Fundamentalsætning)**.** Ethvert positivt heltal n har en entydig primtalsopløsning (op til rækkefølgen)

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}. \tag{2.3}$$

Bevis. Vi har allerede set bevis for eksistensen af primopløsning for n. Vi vil nu vise entydigheden.

Antag for modstrid at der findes tal hvis primopløsning ikke er entydig, og lad n være det mindste af disse tal.

$$n = p_1 p_2 p_3 \cdots p_r = q_1 q_2 q_3 \cdots q_s \tag{2.4}$$

Læg mærke til at vi ikke kan have et primtal der går igen i hver opløsning, da hvis $p_i = q_j$ kunne vi forkorte n med dette og vi ville få et tal mindre en n med ikke-entydig primopløsning, hvilket er i modstrid med vores antagelse. Dette sikrer os at $q_1 - p_1 \neq 0$. Antag uden tab af generalitet at q_1 er større end p_1. Dermed er $(q_1 - p_1)$ et naturligt tal. Vi betragter nu

$$m = (q_1 - p_1) q_2 q_3 \cdots q_s \tag{2.5}$$

og ser at m har en primopløsning, hvor alle q_2 til q_s indgår. p_1 kan ikke indgå i denne, da det sidste vi mangler i primopløsningen er $q_1 - p_1$ som p_1 ikke er divisor i.

På den anden side ser vi

$$\begin{aligned} m &= (q_1 - p_1) q_2 q_3 \cdots q_s \\ &= q_1 q_2 q_3 \cdots q_s - p_1 q_2 q_3 \cdots q_s \\ &= n - p_1 q_2 q_3 \cdots q_s \\ &= p_1 (p_2 p_3 \cdots p_r - q_2 q_3 \cdots q_s) \end{aligned} \tag{2.6}$$

Hvilket fortæller os at $p_1 \mid m$ så m har en primopløsning hvor p_1 indgår. Primopløsningen for m er altså ikke entydig, men da m er mindre end n er dette i modstrid med at n var mindst.

Derfor kan vi konkludere at den type tal ikke eksisterer. ■

Sætning 2.7 (Euklids Sætning)**.** Der findes uendeligt mange primtal.

Bevis. Antag for modstrid at der findes endeligt mange primtal p_1, $p_2, \ldots, p_k$. Betragt nu tallet $n = p_1 p_2 \cdots p_k + 1$. Da n har en primopløsning kan vi specielt finde et primtal p der deler n. Eftersom p er et primtal må den indgå i vores liste af k primtal og derfor også dele produktet $p_1 p_2 \cdots p_k$.

$p \mid n$ og $p \mid p_1 p_2 \cdots p_k$ medfører at $p \mid n - p_1 p_2 \cdots p_k$, hvilket giver 1.

Et primtal kan ikke være divisor i 1, så vi har opnået modstrid og der må derfor findes uendeligt mange primtal. ■

Definition 2.8 (Største fælles divisor)**.** For tallene n og m siger vi at $d \geq 0$ er deres største fælles divisor (*greatest common divisor*), hvis

- $d \mid n$ og $d \mid m$
- hvis vi har d' således at $d' \mid n$ og $d' \mid m$ så vil $d' \mid d$.

Vi skriver

$$d = \gcd(n,m) \quad \text{eller blot} \quad d = (n,m). \tag{2.7}$$

Hvis $\gcd(n,m) = 1$ siger vi at n og m er indbyrdes primiske.

Eksempel 2.9. Se som eksempel på tallene 12 og 16. De tal (større end eller lig 0) der opfylder at være divisor i både 12 og 16 er netop 1, 2 og 4. Vi ser at gcd(12,16) = 4 da 1 | 4 og 2 | 4.

Et andet eksempel er tallene 3 og 10. Disse tal har kun 1 som fælles divisor og derfor er 3 og 10 indbyrdes primiske. ∘

Sætning 2.10 (Regler for største fælles divisor)**.** Lad n, m og k være hele tal.

1) $\gcd(n,m) = \gcd(m,n)$

2) $\gcd(n, \gcd(m,k)) = \gcd(\gcd(n,m), k)$

3) $\gcd(kn, km) = |k| \gcd(n,m)$

4) $\gcd(n,m) = \gcd(m, n-km)$

Disse regneregler bevises i Opgave 6.11.

Sætning 2.11 (Division med rest)**.** Lad n være et helt tal og m et naturligt tal. Da findes to entydigt bestemte tal q og r (kvotient og rest) således at

$$n = qm + r, \qquad 0 \leq r < m. \tag{2.8}$$

Bevis. Betragt følgende

$$\cdots < -3m < -2m < -m < 0 < m < 2m < 3m < \cdots \tag{2.9}$$

Tallet n vil netop befinde sig i ét af intervallerne, dvs. at der findes et $q \in \mathbb{Z}$ så

$$qm \leq n < (q+1)m$$

hvilket giver $0 \leq n - qm < m$, så sæt $r = n - qm$ og vi har dermed vist eksistensen af q og r.

For at vise entydighed antager vi at vi har både q,q',r,r' således at

$$qm + r = n = q'm + r' \tag{2.10}$$

og $0 \leq r,\ r' < m$. Vi har altså $qm + r = q'm + r'$ og dermed $m(q - q') = r' - r$. Bemærk nu at $|r' - r|$ må være mindre end m da $0 \leq r$ og $r' < m$, hvilket giver os $|m(q - q')| < m$. Derfor må $q = q'$. Dette gør nu at ligningen $qm + r = q'm + r'$ er ækvivalent med $r = r'$. Vi har dermed vist at q og r entydigt bestemt. ■

Metode 2.12 (Euklids algoritme)**.** Euklids algoritme er en metode til at finde største fælles divisor for n og m, hvor vi laver division med rest gentagne gange indtil vores rest er 0.

$$\begin{aligned} n &= q_1 m + r_1 && 0 \leq r_1 < m \\ m &= q_2 r_1 + r_2 && 0 \leq r_2 < r_1 \\ r_1 &= q_3 r_2 + r_3 && 0 \leq r_3 < r_2 \\ &\vdots \\ r_{k-1} &= q_{k+1} r_k + 0 \end{aligned}$$

Og så er den største fælles divisor for n og m givet ved r_k.

Bevis. Ifølge Sætning 2.10 er $\gcd(n,m) = \gcd(m, n - q_1 m)$ og fra første lighed i algoritmen har vi at $n - q_1 m = r_1$. Denne argumentation gentages og vi opnår at

$$\gcd(n,m) = \gcd(m, r_1) = \gcd(r_1, r_2) = \cdots = \gcd(r_k, 0) = r_k.$$

■

Eksempel 2.13. Lad os bruge Euklid's algoritme til at finde den største fælles divisor af 33 og 12.

$$\begin{aligned} 33 &= 2 \cdot 12 + 9 \\ 12 &= 1 \cdot 9 + 3 \\ 9 &= 3 \cdot 3 + 0 \end{aligned}$$

Så $\gcd(33, 12) = 3$ da 3 er den sidste rest som ikke er nul. ○

Eksempel 2.14. Lad os nu finde gcd(24,42).

$$\begin{aligned} 42 &= 1 \cdot 24 + 18 \\ 24 &= 1 \cdot 18 + 6 \\ 18 &= 3 \cdot 6 + 0 \end{aligned}$$

Så $\gcd(24, 42) = 6$. ∘

Eksempel 2.15. Lad os nu finde gcd(175, 78).

$$\begin{aligned} 175 &= 2 \cdot 78 + 19 \\ 78 &= 4 \cdot 19 + 2 \\ 19 &= 9 \cdot 2 + 1 \end{aligned}$$

Så $\gcd(175, 78) = 1$, så de er altså indbyrdes primiske. ∘

Sætning 2.16 (Bezouts identitet)**.** For to hele tal n og m, kan vi finde to andre hele tal s og t, således at

$$\gcd(n,m) = sn + tm \tag{2.11}$$

Bevis. Til dette bevis benytter vi os af Euklids algoritme og laver et induktionsbevis over resterne r_i for $i = 0,...,k$ (hvor $r_0 = m$). Vores induktion kræver to startbetingelser, så vi kigger på r_0 og r_1 og ser at vi har linearkombinationerne $r_0 = 0n + 1m$ og $r_1 = 1n - q_1 m$.

I induktionsskridtet antager vi at vi kan skrive $r_k = s_k n + t_k m$ samt $r_{k-1} = s_{k-1} n + t_{k-1} m$ og vil nu vise at r_{k+1} også kan skrives som heltallig linearkombination af n og m.

$$\begin{aligned} r_{k+1} &= r_{k-1} - q_{k+1} r_k \\ &= (s_{k-1} n + t_{k-1} m) - q_{k+1}(s_k n + t_k m) \\ &= (s_{k-1} - q_{k+1} s_k) n + (t_{k-1} - q_{k+1} t_k) m \end{aligned} \tag{2.12}$$

Dermed kan alle resterne i Euklids algoritme skrives som linearkombinationer af n og m og specielt gælder det derfor også for gcd(n,m). ■

Nu har vi alle værktøjerne til at vise Sætning 2.5, men beviset her er faktisk konstruktivt i forhold til at finde s og t så lad os først se på et eksempel.

Eksempel 2.17. Lad os finde s og t for $n = 33$ og $m = 12$. Vi har i Eksempel 2.13 fundet gcd(33,12) = 3 med $r_1 = 9$, $r_2 = 3$ og $r_3 = 0$. Dermed får vi hvis vi indsætter udtrykkene fra vores udregning i eksemplet at

$$\begin{aligned} 3 &= 12 - 9 \\ &= 12 - (33 - 2 \cdot 12) \\ &= 12 - 33 + 2 \cdot 12 \\ &= 3 \cdot 12 + (-1) \cdot 33 \end{aligned}$$

så $s = -1$ og $t = 3$, da vi nu har $3 = 3 \cdot 12 + (-1) \cdot 33$.

Tilsvarendetfor Eksempel 2.15 har vi at $1 = 9 + (-4) \cdot 2$. ∘

Nu er vi klar til beviset for Sætning 2.5.

Bevis for Euklids Lemma (Sætning 2.5). Lad p være et primtal og lad $p \mid ab$. Hvis $p \mid a$ er vi færdige, så vi antager at $p \nmid a$. Derfor må p og a være indbyrdes primiske. Vi bruger derfor Bezouts identitet til at sige at der findes et s og et t således at

$$1 = sa + tp. \tag{2.13}$$

Vi ganger igennem med b og får

$$b = sab + tpb. \tag{2.14}$$

Vi ser at $p \mid sab$ da $p \mid ab$ per antagelse, samt $p \mid tpb$ og dermed må $p \mid sab + tpb$. Med andre ord har vi altså vist at $p \mid b$. ■

3 Kongruenser

Definition 3.1 (Kongruens)**.** Lad n være et givet heltal. Vi siger at to tal, a og b er *kongruente modulo* n, hvis

$$n \mid a - b. \tag{2.15}$$

Vi skriver $a \equiv b \pmod{n}$.

Det betyder altså at a og b har samme rest ved division med n. Ækvivalent med definitionen kan vi beskrive kongruens mellem a og b som at vi kan finde et tal k således at $a = b + kn$.

Eksempel 3.2. Som eksempel er det nemt at se at to lige tal altid er kongruente modulo 2. Det samme gælder for to ulige tal. ∘

Eksempel 3.3. Et andet eksempel på kongruenser, som er ligetil at vurdere, er kongruens modulo 10. To positive tal er kongruente modulo 10 hvis og kun hvis de ender på samme ciffer. Overvej hvordan vi kan formulere et kriterium, så vi også får indraget de negative tal. ∘

Definition 3.4 (Restklasser)**.** Vi definerer restklassen for a modulo n som mængden af alle de tal, der er kongruente med a modulo n. Vi noterer ofte restklassen for a med

$$[a]_n = \{a + kn \mid k \in \mathbb{Z}\}, \tag{2.16}$$

sommetider blot $[a]$ hvis n er klar fra konteksten.

Det er altså en måde at gruppere de tal, der er kongruente modulo n. Og hvis $[a] = [b]$, så kan vi finde på at skrive $a = b$, selvom det kun betyder at $a \equiv b \pmod{n}$. Når det så er sagt, opfører det sig pænt. Faktisk så pænt, at vores regneoperationer + (addition) og · (multiplikation) kan overføres direkte til restklasser.

Eksempel 3.5. Vi har for eksempel at $[3]_5 = \{..., -2,3,8,13,...\}$ og $[7]_7 = \{..., -14, -7,0,7,14,...\}$ så vi har for eksempel at $[3]_5 = [13]_5$ og $[-14]_7 = [7]_7$. ∘

Sætning 3.6 (Regler for restklasser)**.** Følgende gælder for restklasser.

1) $[a] + [b] = [a + b]$

2) $[a][b] = [ab]$

Bevis. Vi vil gerne vise at $[a] + [b] = [a + b]$ hvilket er ækvivalent til at

$$n \mid \underbrace{(a + k_1 n) + (b + k_2 n) - (a + b + kn)}_{=n(k_1+k_2-k)}. \tag{2.17}$$

Siden $n(k_1 + k_2 - k)$ altid er delelig med n, vil $[a] + [b]$ altid være lig med $[a + b]$. Tilsvarende har vi at $[a][b] = [ab]$ er ækvivalent med

$$n \mid \underbrace{(a + k_1 n)(b + k_2 n) - (ab + kn)}_{=n(ak_2+bk_1+k_1k_2n-k)}, \tag{2.18}$$

og da vi ender ud med et produkt af n, gælder samme argument som før. ■

Definition 3.7. Vi definerer $\mathbb{Z}/n\mathbb{Z}$ til at være mængden af restklasser modulo n. Ofte vil vi blot notere en restklasse $[a] \in \mathbb{Z}/n\mathbb{Z}$ med dens repræsentant a. Det vil sige $\mathbb{Z}/n\mathbb{Z}$ kan repræsenteres som $\{0, 1, 2, \ldots, n - 1\}$.

Sætning 3.8. Hvis a og n er indbyrdes primiske, så eksisterer der et tal b således at

$$ab \equiv 1 \pmod{n}. \tag{2.19}$$

Vi kalder b for den inverse til a modulo n og skriver $b = a^{-1}$.

Bevis. Lad $\gcd(a,n) = 1$. Bezouts identitet fortæller at der eksisterer tallene s og t således at

$$sa + tn = \gcd(a,n) = 1.$$

Ved at omskrive dette til $tn = 1 - sa$ ser vi at n deler $1 - sa$, så pr. definition er $sa \equiv 1 \pmod{n}$ og s er dermed invers til a. ■

Vi bemærker at den inverse til a er entydig modulo n, da hvis både b og c opfylder

$$\begin{aligned} ab &\equiv 1 \pmod{n} \\ ac &\equiv 1 \pmod{n} \end{aligned} \tag{2.20}$$

får vi at

$$b \equiv b(ac) \equiv (ba)c \equiv c \pmod{n}. \tag{2.21}$$

Det at vi kan finde en invers til a betyder at vi kan løse ligninger på formen

$$ax = m \pmod{n} \tag{2.22}$$

for x. Vi kan finde a's invers b og få

$$bax = bm \pmod{n}, \tag{2.23}$$

hvor $ba \equiv 1 \pmod{n}$. Så bm er en løsning til ligningen.

Eksempel 3.9. Fra Eksempel 2.17 ved vi at hvis vi lader $a = 2$ og $n = 9$, som jo er indbyrdes primiske med $1 = 9 + (-4) \cdot 2$ så har vi jo at

$$1 \equiv 9 - 4 \cdot 2 \equiv 2 \cdot (-4) \pmod{9}.$$

Dermed har vi en løsning $b = -4$. ∘

Sætning 3.10 (Wilsons Sætning)**.** Lad $p > 1$ være et heltal. p er et primtal hvis og kun hvis

$$(p-1)! \equiv -1 \pmod{p}. \tag{2.24}$$

Bevis. Lad p være et primtal. $(p-1)!$ er et produkt af tallene fra 1 til $p-1$, som alle er indbyrdes primiske med p. Derfor har alle disse en invers modulo p pr. Sætning 3.8, og disse inverser har en repræsentant mellem 1 og $p-1$. Vi vil derfor se en masse par gå ud med hinanden, på nær de af tallene, der har sig selv som invers. Ligningen $a^2 \equiv 1 \pmod{p}$ kan højest have to løsninger modulo p og vi ved at 1 og -1 er løsninger, så der findes altså ikke flere. Husk at $-1 \equiv p-1 \pmod{p}$. Dermed er $(p-1)! = 1 \cdot 2 \cdots (p-1) \equiv p-1 \pmod{p}$.

Vi skal også bevise påstanden den modsatte retning. Så start med et p hvor der gælder at $(p-1)! = -1 \pmod{p}$. Antag for modstrid at p ikke er et primtal. Altså eksisterer der et $d \neq 1$ hvor $d \mid p$.

Fordi $d \mid p$ og $p \mid (p-1)! - (-1)$ får vi at $d \mid (p-1)! + 1$. Men eftersom at d indgår i produktet $(p-1)!$ har vi samtidigt at $d \mid (p-1)!$. Det giver tilsammen at $d \mid 1$. Men fordi $d > 1$, er dette ikke muligt.

Vi har dermed opnået modstrid og vi konkluderer derfor at p må være et primtal. ■

Definition 3.11 (Eulers φ-funktion)**.** For et naturligt tal n, giver Eulers φ-funktion antallet af tal, fra 1 til n, som er indbyrdes primiske med n.

Eksempel 3.12. Lad os se på tallet 10. Listen af tal op til 10, der er indbyrdes primiske med 10 er netop 1, 3, 7, 9. Så $\varphi(10) = 4$. ∘

Bemærkning 3.13. Bemærk at for et primtal p er $\varphi(p) = p-1$, da $\gcd(p,a) = 1$ for alle a fra 1 til $p-1$.

Sætning 3.14 (Euler-Fermats Sætning)**.** Hvis $\gcd(a,n) = 1$, så er

$$a^{\varphi(n)} \equiv 1 \pmod{n}. \tag{2.25}$$

Bevis. Lad $X = \{x_1,x_2,\ldots,x_{\varphi(n)}\}$ være mængden af de $\varphi(n)$ tal mellem 0 og n, hvor $\gcd(x_i,n) = 1$ for alle $i = 1,\ldots,\varphi(n)$. For et tal a, der er indbyrdes primisk med n, vil der også gælde at $\gcd(ax_i,n) = 1$. Mængden $aX = \{ax_1,ax_2,\ldots,ax_{\varphi(n)}\}$ består altså af tal der, modulo n, er ækvivalente med et tal fra X.

Vi ser også, at hvis $ax \equiv ay \pmod{n}$, kan vi gange med en invers til a, og få $x \equiv y \pmod{n}$. Da ingen af elementerne i X er ækvivalente modulo n, da de er forskellige tal mellem 0 og n, betyder det derfor at ingen elementer i aX er ækvivalente modulo n. Vi har derfor at de to mængder er helt ens modulo n, og vi ser derfor at

$$x_1 \cdot x_2 \cdots x_{\varphi(n)} \equiv ax_1 \cdot ax_2 \cdots ax_{\varphi(n)} \pmod{n}. \tag{2.26}$$

Vi ved at vi kan finde inverser til alle elementer i X og kan dermed nå frem til

$$1 \equiv a^{\varphi(n)} \pmod{n} \tag{2.27}$$

som ønsket. ■

Hvis vi ser på tilfældet hvor p er et primtal, husker vi at $\varphi(n) = p - 1$. Ved at benytte dette i sammenhæng med Euler-Fermats sætning, får vi Fermats lille sætning.

Korollar 3.15 (Fermats lille sætning)**.** Lad p være et primtal. Hvis $\gcd(p,a) = 1$ så gælder

$$a^{p-1} \equiv 1 \pmod{p}. \tag{2.28}$$

Definition 3.16 (Orden)**.** Vi siger at k er ordnen af a modulo n, hvis k er det mindste naturlige tal der opfylder at

$$a^k \equiv 1 \pmod{n}. \tag{2.29}$$

Hvis et sådant k ikke findes, siger vi at ordnen af a modulo n er ∞.

Eksempel 3.17. Som eksempel har tallet 2 orden 3 modulo 7, da $2^3 = 8 \equiv 1 \pmod{7}$ og $k = 3$ er det mindste naturlige tal, der opfylder dette. ○

Eksempel 3.18. Som et andet eksempel har tallet 2 uendelig orden modulo 6, da 2^k kun vil tage værdierne 2 og 4 modulo 6, og kan derfor aldrig blive 1. ○

Definition 3.19 (Primitive rødder)**.** Et tal a kaldes en primitiv rod modulo n, hvis der for alle x hvor $\gcd(x,n) = 1$ findes et k således at

$$x \equiv a^k \pmod{n}. \tag{2.30}$$

Når vi ser på primitive rødder modulo et primtal p, betyder det, at hvis a er en primitiv rod modulo p, kan vi beskrive samtlige restklasser i $\mathbb{Z}/p\mathbb{Z}$, på nær $[0]_p$, med potenser af a. Primitive rødder er derfor vældigt nyttige, fordi vi kan bruge dem som byggeklodser for $\mathbb{Z}/p\mathbb{Z}$.

Sætning 3.20. Lad p være et primtal. Der vil altid kunne findes en primitiv rod i $\mathbb{Z}/p\mathbb{Z}$.

Bevis. Vi starter med at vise, at hvis p er et primtal og d er divisor i $p-1$, så vil polynomiet x^d-1 have *præcis* d rødder i $\mathbb{Z}/p\mathbb{Z}$.

Da $d \mid p-1$, vil $e=(p-1)/d$ være et heltal. Vi ved fra Fermats lille sætning at alle a fra 1 til $p-1$ er rødder til $x^{p-1}-1$ modulo p, så vi har præcis $p-1$ rødder i $\mathbb{Z}/p\mathbb{Z}$. Vi omskriver dette polynomium og får

$$\begin{aligned} x^{p-1}-1 &= \left(x^d\right)^{\frac{p-1}{d}}-1 \\ &= \left(x^d\right)^e-1 \\ &= \left(x^d-1\right)\left((x^d)^{e-1}+(x^d)^{e-2}+\cdots+1\right) \\ &= \left(x^d-1\right)g(x), \end{aligned} \tag{2.31}$$

hvor vi har defineret $g(x)$ som polynomiet i parantesen. $g(x)$ har grad $d(e-1)=de-d=p-1-d$. Det har derfor maksimalt $p-1-d$ rødder. Imens har x^d-1 grad d og derfor maksimalt d rødder. For at være en rod i $\left(x^d-1\right)g(x)$ skal man være en rod i en af de to faktorer, så udtrykket har derfor maksimalt $p-1-d+d=p-1$ rødder. Men som vi så tidligere, har polynomiet i alt præcist $p-1$ rødder. Så $g(x)$ har præcist $p-1-d$ rødder, imens x^d-1 har præcist d rødder.

Så langt, så godt. Nu ser vi på primfaktoriseringen af $p-1$

$$p-1=q_1^{n_1}q_2^{n_2}\cdots q_r^{n_r}. \tag{2.32}$$

Definerer vi $f_1(x)=x^{q_i^{n_i}}-1$ ved vi nu at polynomiet har præcis $q_i^{n_i}$ rødder i $\mathbb{Z}/p\mathbb{Z}$, og ligeledes at $f_2(x)=x^{q_i^{n_i-1}}-1$ har $q_i^{n_i-1}$ rødder i $\mathbb{Z}/p\mathbb{Z}$. Alle rødderne i f_2 er samtidigt rødder i f_1, idet en given rod, s, i f_2 opfylder at $s^{q_i^{n_i-1}}=1$ så $1=1^q=(s^{q_i^{n_i-1}})^q=s^{q_i^{n_i}}$.

Derfor har vi $q_i^{n_i}-q_i^{n_i-1}$ rødder i f_1, der *ikke* er rødder i f_2. Med andre ord; vi har $q_i^{n_i}-q_i^{n_i-1}$ tal der har orden $q_i^{n_i}$. For hvert $i=1,\ldots,r$ vælger vi et a_i med orden $q_i^{n_i}$ og ser nu på produktet af disse.

$$a=a_1a_2\cdots a_r \tag{2.33}$$

Tallet a vil have orden $q_1^{n_1} q_2^{n_2} \cdots q_r^{n_r} = p-1$ og derfor er a en primitiv rod modulo p. ■

4 Kvadratiske Rester

Matematikken i dette afsnit er i høj grad motiveret af spørgsmålet om, under hvilke betingelser ligningen

$$ax^2 + bx + c \equiv 0 \pmod{n} \tag{2.34}$$

har løsninger. Svaret til dette spørgsmål kan ofte omskrives til problemet om hvorvidt diskriminanten, d, har en kvadratrod modulo et primtal p. Vi lægger ud med en definition.

Definition 4.1 (Kvadratisk rest)**.** Lad $n \in \mathbb{N}$ og $a \in \mathbb{Z}$. Tallet a kaldes for en *kvadratisk rest* modulo n, hvis der findes et $x \in \mathbb{Z}$ så

$$x^2 \equiv a \pmod{n}. \tag{2.35}$$

Eksempel 4.2. Lad os kigge på $n = 7$. For at finde de kvadratiske rester modulo 7, beregner vi blot x^2 for alle x.

$$0^2 \equiv 0, \quad 1^2 \equiv 1, \quad 2^2 \equiv 4, \quad 3^2 \equiv 2,$$
$$4^2 \equiv 2, \quad 5^2 \equiv 4, \quad 6^2 \equiv 1.$$

Så de kvadratiske rester modulo 7 er tallene 0, 1, 2 og 4. ∘

Legendresymbolet

Definition 4.3 (Legendresymbol)**.** Lad p være et ulige primtal. For $a \in \mathbb{Z}$ defineres *legendresymbolet* ved

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{hvis } a \text{ er en kvadratisk rest modulo} \\ & p \text{ og } a \not\equiv 0 \pmod{p}, \\ -1 & \text{hvis } a \text{ ikke er en kvadratisk rest modulo } p, \\ 0 & \text{hvis } a \equiv 0 \pmod{p}. \end{cases} \tag{2.36}$$

Eksempel 4.4. Lad os kigge på tilfældet, $p = 7$. Fra Eksempel 4.2 kender vi samtlige kvadratiske rester modulo 7, så vi får følgende værdier.

$$\left(\frac{0}{7}\right) = 0, \quad \left(\frac{1}{7}\right) = 1, \quad \left(\frac{2}{7}\right) = 1, \quad \left(\frac{3}{7}\right) = -1,$$
$$\left(\frac{4}{7}\right) = 1, \quad \left(\frac{5}{7}\right) = -1, \quad \left(\frac{6}{7}\right) = -1$$

Deruover kan vi også se på Legendresymboler for andre tal. For eksempel er $\left(\frac{28}{7}\right) = 0$ da $28 \equiv 0 \pmod 7$ og $\left(\frac{65}{7}\right) = 1$ da $65 \equiv 2 \pmod 7$. ∘

Bemærkning 4.5. Denne notation for legendresymbol er meget udbredt, men skal ikke forveksles med brøker!

Lemma 4.6. Lad p være et ulige primtal. Hvis g er en primitiv rod modulo p, da er

$$\left(\frac{g^i}{p}\right) = (-1)^i, \tag{2.37}$$

hvor i kan være et hvilket som helst ikke-negativt heltal.

Bevis. Fra Sætning 3.19 ved vi at der eksisterer en primitiv rod g modulo p. Vi ser at de kvadratiske rester er

$$(g)^2, (g^2)^2, (g^3)^2, \ldots, (g^{p-1})^2.$$

De kvadratiske rester er derfor netop de lige potenser af g, men det er også der hvor højresiden er 1. ■

Sætning 4.7. Lad p være et ulige primtal. For alle $a, b \in \mathbb{Z}$ gælder

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right) \tag{2.38}$$

Bevis. Hvis $p \mid ab$ så vil $p \mid a$ eller $p \mid b$, og begge sider vil da være 0. Hvis $p \nmid ab$, lad $a = g^i$ og $b = g^j$, hvor g er en primitiv rod modulo p. Da giver $ab = g^{i+j}$ og af Lemma 4.6 ser vi at

$$\left(\frac{ab}{p}\right) = (-1)^{i+j} = (-1)^i(-1)^j = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right). \tag{2.39}$$

■

Sætning 4.8 (Eulers kriterium)**.** Lad p være et ulige primtal. For alle $a \in \mathbb{Z}$ gælder

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod{p}. \tag{2.40}$$

Bevis. Hvis $p \mid a$ får vi $\left(\frac{a}{p}\right) = 0$ samt at $a^{(p-1)/2} \equiv 0 \pmod{p}$, så sætningen gælder i dette tilfælde.

Antag at $p \nmid a$. Da p er et primtal, kan vi finde en primitiv rod g modulo p, og derfor kan vi skrive $a = g^i$.

Vi definerer nu $h = g^{(p-1)/2}$ og ser at $h^2 = g^{p-1} \equiv 1 \pmod{p}$ ifølge Fermats lille sætning. Af dette deducerer vi til $h = \pm 1$, men udelukker hurtigt at $h = 1$, da g har orden $p - 1$. Så vi kan derfor ikke have at $g^{(p-1)/2} = 1$. Ergo er $h = -1$, hvorefter med hjælp af Lemma 4.6 når frem til

$$a^{(p-1)/2} = (g^i)^{(p-1)/2} = (g^{(p-1)/2})^i = (-1)^i = \left(\frac{g^i}{p}\right) = \left(\frac{a}{p}\right) \tag{2.41}$$

som ønsket. ■

Korollar 4.9. Lad p være et ulige primtal. Da er

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}. \tag{2.42}$$

Bevis. Dette følger direkte af Sætning 4.8, ved at sætte $a = -1$. ■

Sætning 4.10 (Gauss' lemma)**.** Lad p være et ulige primtal og lad $a \in \mathbb{Z}$ være indbyrdes primisk med p. Betragt tallene

$$a, 2a, 3a, \ldots, \frac{p-1}{2}a$$

og deres mindste positive rester modulo p. Lad μ være antallet af disse rester som er større end $\frac{p}{2}$. Da er

$$\left(\frac{a}{p}\right) = (-1)^{\mu}. \tag{2.43}$$

Bevis. Vi vil gerne bemærke at hvis ka optræder, gør $-ka$ det ikke. Antag for modstrid at der findes et l, hvor $0 < l \leq \frac{p-1}{2}$ så $la \equiv -ka \pmod{p}$. Betragt nu $ka+la \equiv 0 \pmod{p}$. Da vil $p \mid k+l$ da a og p er indbyrdes primiske, men da $0 < k+l \leq p-1$, kan p ikke dele $k+l$, så sådan et l kan ikke findes, og tallene må bare være en permutation af tallene $1,2,\ldots,\frac{p-1}{2}$, hvor nogle har skiftet fortegn.

$$\begin{aligned} a^{(p-1)/2}\left(\frac{p-1}{2}\right)! &= a \cdot 2a \cdots \frac{p-1}{2}a \\ &\equiv (-1)^{\mu}\left(\frac{p-1}{2}\right)! \pmod{p} \end{aligned} \tag{2.44}$$

Fra Eulers kriterium, og hvis vi lader de to $\left(\frac{p-1}{2}\right)!$ gå ud, ser vi altså at

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \equiv (-1)^{\mu} \pmod{p}. \tag{2.45}$$

■

Korollar 4.11. Lad p være et ulige primtal. Da er

$$\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}. \tag{2.46}$$

Bevis. Ved at sætte $a = 2$ i Gauss' lemma, får vi tallene

$$2, 4, 6, \ldots, p-1.$$

Vi ønsker at vurdere, hvor mange tal på listen, der er større end $p/2$.

Vi ser først på tilfældet hvor $p \equiv 1 \pmod{4}$. De første $(p-1)/4$ elementer, $2, 4, \ldots, (p-1)/2$ vil være mindre end $p/2$, imens de resterende $(p-1)/4$ elementer $(p+3)/2, \ldots, p-1$ er større end $p/2$.

Altså får vi at $\mu = (p-1)/4$ som ifølge Gauss' lemma giver

$$\left(\frac{2}{p}\right) = (-1)^{(p-1)/4}. \tag{2.47}$$

Da tallet $(p+1)/2$ er ulige når $p \equiv 1 \pmod{4}$, kan vi lave følgende omskrivning:

$$(-1)^{(p-1)/4} = \left((-1)^{(p-1)/4}\right)^{(p+1)/2} = (-1)^{(p^2-1)/8}. \tag{2.48}$$

Tag nu tilfældet hvor $p \equiv 3 \pmod 4$. Da er de første $(p-3)/4$ elementer mindre end $p/2$, mens de resterende $(p+1)/4$ er større end $p/2$. Tilsvarende ser vi at

$$\left(\frac{2}{p}\right) = (-1)^{(p+1)/4} = \left((-1)^{(p+1)/4}\right)^{(p-1)/2} = (-1)^{(p^2-1)/8}. \quad (2.49)$$

■

Den Kvadratiske Reciprocitetssætning

Sætning 4.12 (Eisensteins lemma)**.** Lad p og q være to forskellige ulige primtal. Lad endvidere $\alpha(q,p)$ være summen

$$\alpha(q,p) = \left\lfloor \frac{q}{p} \right\rfloor + \left\lfloor \frac{2q}{p} \right\rfloor + \left\lfloor \frac{3q}{p} \right\rfloor + \cdots + \left\lfloor \frac{\frac{p-1}{2}q}{p} \right\rfloor, \quad (2.50)$$

hvor $\lfloor x \rfloor$ betegner *floor* funktionen (det største heltal mindre end eller lig med x). Da gælder

$$\left(\frac{q}{p}\right) = (-1)^{\alpha(q,\,p)}. \quad (2.51)$$

Bevis. Fra sætningen om heltalsdivision med rest ved vi at

$$kq = p\left\lfloor \frac{kq}{p} \right\rfloor + r_k, \text{ hvor } 0 \le r_k < p \quad (2.52)$$

Fra dette ser vi at

$$\begin{aligned} 1q + 2q + \cdots + p'q &= \left(p\left\lfloor \frac{1q}{p} \right\rfloor + p\left\lfloor \frac{2q}{p} \right\rfloor + \cdots + p\left\lfloor \frac{p'q}{p} \right\rfloor\right) \\ &\quad + (r_1 + r_2 + \cdots + r_{p'}) \\ q\,(1 + 2 + \cdots + p') &= p\,\alpha(q,p) + (r_1 + r_2 + \cdots + r_{p'}) \end{aligned}$$

Lad A være mængden af rester r_i som er mindre eller lig p' og B være mængden af rester som er større end p'. Sæt $\nu = |A|$ og $\mu = |B|$. Vi

har da at

$$\begin{aligned}(r_1 + \cdots + r_{p'}) &= (a_1 + \cdots + a_\nu) + (b_1 + \cdots + b_\mu) \\ &= (a_1 + \cdots + a_\nu) + (p - b_1 + \cdots + p - b_\mu) \\ &\quad - \mu\, p + 2(b_1 + \cdots + b_\mu) \\ &= (1 + 2 + \cdots + p') - \mu\, p + 2(b_1 + \cdots + b_\mu)\end{aligned}$$

Vi indsætter dette i ligningen fra før

$$\begin{aligned}q\,(1 + 2 + \cdots + p') &= p\,\alpha(q,p) + (1 + 2 + \cdots + p') \\ &\quad - \mu\, p + 2(b_1 + \cdots + b_\mu) \\ (q-1)\,(1 + 2 + \cdots + p') &= p\,\alpha(q,p) - \mu\, p + 2(b_1 + \cdots + b_\mu)\end{aligned}$$

Da både p og q er ulige primtal ser vi at

$$\alpha(q,p) \equiv \mu \pmod 2$$

Fra Gauss Lemma ser vi at

$$\left(\frac{q}{p}\right) = (-1)^\mu = (-1)^{\alpha(q,p)}$$

hvilket skulle vises. ■

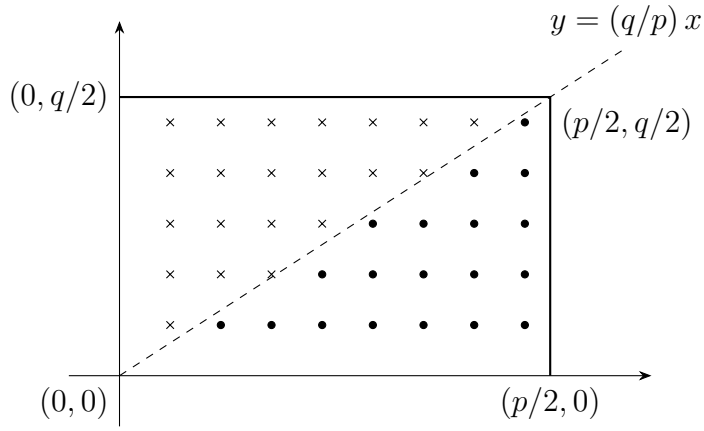
Sætning 4.13 (Den kvadratiske reciprocitetssætning)**.** Lad p og q være to forskellige ulige primtal. Da er

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}.$$

Bevis. Måden vi vil bevise sætningen på, er ved at tælle heltalspunkterne i et rektangel på to forskellige måder. Lad R være rektanglet med hjørner i punkterne $(0,\, 0)$, $(p/2,\, 0)$, $(p/2,\, q/2)$ og $(0,\, q/2)$, og betragt heltalspunkterne som ligger inde i R men uden at være på kanten (se Figur 2.1).

Da p og q begge er ulige, vil heltalspunkterne i R være punkterne $(x,\, y)$, hvor $1 \le x \le (p-1)/2$ og $1 \le y \le (q-1)/2$. Det ses heraf at antallet af heltalspunkter i R må være

$$\frac{p-1}{2} \cdot \frac{q-1}{2}.$$



Figur 2.1: Rektanglet R med diagonal og heltalspunkter.

Lad D være diagonalen i rektanglet fra $(0,\ 0)$ til $(p/2,\ q/2)$, denne linje har forskriften $y = (q/p)\,x$. Da p og q er forskellige primtal, vil ingen af heltalspunkterne i R ligge på D. Det skyldes at punkterne ville skulle opfylde ligningen $py = qx$ som medfører at $p \mid x$ og $q \mid y$, men da $0 < x < p/2$ og $0 < y < q/2$ kan dette ikke lade sig gøre. Heltalspunkterne i R kan altså adskilles alt efter om de ligger under eller over diagonalen D.

Lad os starte med at finde antallet af heltalspunkter i R som ligger under D ved at tælle punkterne søjle for søjle. Vi fokuserer på søjlen hvor $x = k$, altså de heltalspunkter som har formen $(k,\ y)$. Punkterne i denne søjle svarer til heltallene i intervallet $0 < y < kq/p$, og antallet af dem er netop $\lfloor kq/p \rfloor$. Ved at summere søjlerne hvor $k = 1, 2, \ldots, (p-1)/2$, får vi at antallet af heltalspunkter under D er

$$\left\lfloor \frac{1q}{p} \right\rfloor + \left\lfloor \frac{2q}{p} \right\rfloor + \cdots + \left\lfloor \frac{(\frac{p-1}{2})q}{p} \right\rfloor = \alpha(q,\ p).$$

Tilsvarende ved at tælle rækkevis har vi at antallet af heltalspunkter i R over D er $\alpha(p,\ q)$. Det samlede antal heltalspunkter i R er altså

$$\alpha(q,\ p) + \alpha(p,\ q) = \frac{p-1}{2} \cdot \frac{q-1}{2}.$$

Fra Eisensteins lemma (Sætning 4.12) følger det at

$$\begin{aligned}\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) &= (-1)^{\alpha(p,\,q)}(-1)^{\alpha(q,\,p)} \\ &= (-1)^{\alpha(p,\,q)+\alpha(q,\,p)} \\ &= (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.\end{aligned}$$

Dette fuldender beviset for kvadratisk reciprocitet. ■

Opsummering af regler

Til sidst opsummerer vi de mest fundamentale regneregler, der gælder for legendresymbolerne. Lad p og q være to forskellige, ulige primtal. Da gælder følgende.

(i) $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$

(ii) $\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)(-1)^{(p-1)(q-1)/4}$

(iii) $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}$

(iv) $\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}$

Disse regneregler er yderst brugbare, så lad os betragte nogle eksempler, hvor vi benytter dem i praksis.

Eksempel 4.14. Som første eksempel, vil vi se på om 2 er en kvadratisk rest modulo 3. Der er mange veje, der fører til Rom, og vi vil vise tre af disse veje.

- Da 3 er et lille tal, kan vi bare udregne alle kvadratiske rester;

  $$0^2 \equiv 0, \quad 1^2 \equiv 1, \quad 2^2 \equiv 1,$$

  og se at 2 ikke er på listen så $\left(\frac{2}{3}\right) = -1$.

- Vi kan bruge regneregel (iii), da $2 \equiv -1 \pmod 3$.

$$\left(\frac{-1}{3}\right) = (-1)^{(3-1)/2} = -1.$$

- Slutteligt, kan vi selvfølgelig benytte regneregel (iv).

$$\left(\frac{2}{3}\right) = (-1)^{(3^2-1)/8} = -1.$$

Heldigvis er alle metoderne enige om at 2 ikke er en kvadratisk rest modulo 3. ∘

Eksempel 4.15. Lad os se om 6 er en kvadratisk rest modulo 13.

$$\begin{aligned}\left(\frac{6}{13}\right) &\overset{(i)}{=} \left(\frac{2}{13}\right)\cdot\left(\frac{3}{13}\right)\\ &\overset{(ii)}{=} \left(\frac{2}{13}\right)\cdot\left(\frac{13}{3}\right)(-1)^{(3-1)(13-1)/4}\\ &\overset{(iii)}{=} (-1)^{(13^2-1)/8}\cdot\left(\frac{13}{3}\right)(-1)^{(3-1)(13-1)/4}\\ &= (-1)^{21}\cdot\left(\frac{13}{3}\right)(-1)^{6}\\ &= -\left(\frac{13}{3}\right) = -\left(\frac{1}{3}\right) = -1\end{aligned}$$

På sidste linje bruger vi at $13 \equiv 1 \pmod 3$ samt at 1 altid er en kvadratisk rest. Vi får sammenlagt at 6 ikke er en kvadratisk rest modulo 13. ∘

5 Summer af Kvadrater

Summer af to kvadrater

Definition 5.1. For et givet positivt heltal k, lad

$$S_k = \{n \mid n = x_1^2 + \cdots + x_k^2,\ x_i \in \mathbb{Z},\ i = 1,\ldots,k\} \tag{2.53}$$

betegne mængden af summer af k kvadrater.

Eksempel 5.2. $S_1 = \{0, 1, 4, 9, 16, 25, \dots\}$ er mængden af alle kvadrater. Altså $0^2, 1^2, 2^2, 3^2$ osv. ∘

Eksempel 5.3. S_2 er mængden der består af summer af to kvadrater. Altså tal, der kan skrives på formen $x^2 + y^2$. Tallene $2, 5$ og 8 ligger fx i S_2, mens $3, 6$ og 7 ikke kan skrives som summer af to kvadrater. ∘

Bemærkning 5.4. Bemærk at S_k er en delmængde af S_{k+1}, da vi altid kan lægge kvadratet 0 til vores element. Altså $S_k \subset S_{k+1}$.

Lemma 5.5. Mængden S_2 er lukket under multiplikation. Med lukket mener vi, at hvis x,$y \in S_2$, så er $xy \in S_2$.

Bevis. Dette følger af identiteten

$$\begin{aligned}(a_1^2 + b_1^2)(a_2^2 + b_2^2) =& a_1^2 b_2^2 + b_1^2 a_2^2 + (2a_1 a_2 b_1 b_2 - 2a_1 a_2 b_1 b_2) \\ &+ a_1^2 a_2^2 + b_1^2 b_2^2 \\ =& (a_1 b_2 + b_1 a_2)^2 + (a_1 a_2 - b_1 b_2)^2.\end{aligned} \tag{2.54}$$

■

Lemma 5.6. Hvis $p > 2$ er et primtal, hvor $p \equiv 1 \pmod 4$, så findes x således at

$$x^2 + 1 = mp, \tag{2.55}$$

hvor $0 < m < p$.

Bevis. Fra Korollar 4.9 ses det at -1 er en kvadratisk rest modulo p når $p \equiv 1 \pmod 4$. Der findes altså et x således at $x^2 \equiv -1 \pmod p$. Hvis vi vælger repræsentanten for x, der ligger mellem 0 og $p-1$ har vi

$$mp = x^2 + 1 \leq (p-1)^2 + 1 < p^2 \tag{2.56}$$

hvor den sidste ulighed følger af at $(p-1)^2 + 1 = p^2 - 2(p-1)$ som er mindre en p^2, når vi som her har at $p > 1$. ■

Sætning 5.7. Ethvert primtal $p \equiv 1 \pmod 4$ kan skrives som en sum af to kvadrater.

Bevis. Lad $A_p = \{n \in \mathbb{N} \mid np \in S_2\}$. Fra foregående sætning ser vi at denne mængde ikke er tom, da der findes $0 < n < p$ hvor $np = x^2 + 1^2$. Eftersom mængden ikke er tom og begrænset nedadtil, må den have et mindste element, som vi kan kalde m. Hvis m er 1 er $1 \cdot p \in S_2$ så er vi færdige, så antag for modstrid at $m > 1$. Vi skriver

$$mp = a_1^2 + b_1^2. \tag{2.57}$$

Lad nu $a_2 \equiv a_1 \pmod{m}$ og $b_2 \equiv b_1 \pmod{m}$ sammen med uligheden $|a_2|, |b_2| \leq m/2$. Af dette får vi

$$a_2^2 + b_2^2 \equiv a_1^2 + b_1^2 \equiv mp \equiv 0 \pmod{m}. \tag{2.58}$$

Altså findes der et tal s så $a_2^2 + b_2^2 = sm$. Bemærk at venstresiden ikke er negativ, så $s \geq 0$.

Eftersom $|a_2|, |b_2| \leq m/2$, har vi

$$sm = a_2^2 + b_2^2 \leq 2(m/2)^2 = m^2/2. \tag{2.59}$$

Ergo er $s \leq m/2$ og særligt $s < m$.

Antag for modstrid at $s = 0$. Dette giver os $a_2^2 + b_2^2 = 0$ og dermed $a_2 = b_2 = 0$. Heraf fås $a_1 \equiv b_1 \equiv 0 \pmod{m}$, altså er m divisor i både a_1 og b_1. Så m^2 vil dele $a_1^2 + b_1^2 = mp$, hvilket betyder at m må dele p.

Dette er i modstrid med at p er et primtal og derfor kun har trivielle divisorer, imens m hverken er 1 eller p. Derfor må s være positiv. Vi har nu $0 < s < m$.

Vi betragter nu produktet

$$(a_1^2 + b_1^2)(a_2^2 + b_2^2) = mp \cdot sm = m^2 sp \tag{2.60}$$

og benytter identiteten fra Ligning (2.54) til at få

$$(a_1 a_2 + b_1 b_2)^2 + (a_1 b_2 - b_1 a_2)^2 = (a_1^2 + b_1^2)(a_2^2 + b_2^2) = m^2 sp. \tag{2.61}$$

Vi lægger mærke til at

$$a_1a_2 + b_1b_2 \equiv a_1^2 + b_1^2 \equiv 0 \pmod{m} \tag{2.62}$$

samt

$$a_1b_2 - b_1a_2 \equiv a_1b_1 - b_1a_1 \equiv 0 \pmod{m}, \tag{2.63}$$

så disse er delelige med m. Dette giver os

$$\left(\frac{a_1a_2 + b_1b_2}{m}\right)^2 + \left(\frac{a_1b_2 - b_1a_2}{m}\right)^2 = sp, \tag{2.64}$$

som betyder at $sp \in S_2$ imens $0 < s < m$.

Igen er dette i modstrid med antagelsen om, at m var det mindste tal med denne egenskab. Ergo må $m = 1$ og $p \in S_2$. ■

Sætning 5.8. Hvis $n \in S_2$ og primtallet $q \equiv 3 \pmod 4$ deler n, da vil q have lige eksponent i n's primopløsning.

Bevis. Lad $q \equiv 3 \pmod 4$ være et primtal der deler $n \in S_2$. Hvis $n = a^2 + b^2$, betyder det at $a^2 + b^2 \equiv 0 \pmod q$.

Antag for modstrid at q ikke deler a. Da kan vi finde en invers til a og får derfor

$$a^2(a^{-1})^2 + b^2(a^{-1})^2 \equiv 0 \pmod{q}. \tag{2.65}$$

Dermed er

$$1 + (ba^{-1})^2 \equiv 0 \pmod{q}. \tag{2.66}$$

Men dette er i modstrid med at -1 ikke er en kvadratisk rest modulo q, når $q \equiv 3 \pmod 4$, per Korollar 4.9.

Dermed vil q dele a. Vi kan bruge samme argument til at nå frem til, at q må dele b. Ergo er q^2 divisor i $a^2 + b^2 = n$.

Resten følger nu af at vi kan faktorisere $a = cq$ og $b = dq$, og dermed $a^2 + b^2 = q^2(c^2 + d^2)$. Hvis q deler $c^2 + d^2$ bruges samme argument. Dette fortsætter indtil alle potenser af q er faktoriseret ud. ■

Sætning 5.9. Et positivt heltal n kan skrives som sum af to kvadrater hvis og kun hvis

$$n = 2^e p_1^{e_1} \cdots p_k^{e_k} q_1^{2f_1} \cdots q_l^{2f_l}, \tag{2.67}$$

hvor $p_i \equiv 1 \pmod 4$ og $q_i \equiv 3 \pmod 4$.

Bevis. Først viser vi, at hvis n har den givne form, kan den skrives som sum af to kvadrater. Dette følger af at $2 = 1^2 + 1^2 \in S_2$ og $q_i^2 = 0^2 + q_i^2 \in S_2$ samt at $p_i \in S_2$ per Sætning 5.7. Og ikke mindst, at S_2 er lukket under multiplikation.

Vi mangler nu blot at vise den anden del af sætningen; at hvis $n \in S_2$, må n nødvendigvis have den givne form. Dette gør vi ved kontraposition. Altså vi vil vise at hvis et tal *ikke* kan skrives på den givne form, vil det *ikke* være et element i S_2.

Hvis n ikke kan skrives på denne form, betyder det at vi har et q_i der deler n præcist et ulige antal gange, og så følger resten af Sætning 5.8. ■

Summer af fire kvadrater

Vi fortsætter vores gennemgang af summer af kvadrater ved at se på summer af fire kvadrater. I dette afsnit kommer vi til at vise firkvadratssætningen, som siger at alle positive heltal kan skrives som sum af fire kvadrater. Sætningen blev formuleret af Bachet i 1621 men blev først bevist i 1770 af Lagrange.

Vores strategi for at bevise sætningen følger meget vores tilgang til summer af to kvadrater. Vi starter med at vise nogle lemmaer.

Lemma 5.10. Mængden S_4 er lukket under multiplikation.

Bevis. Dette følger af identiteten

$$\begin{aligned}&(a_1^2+b_1^2+c_1^2+d_1^2)(a_2^2+b_2^2+c_2^2+d_2^2)\\ =\quad &(a_1a_2+b_1b_2+c_1c_2+d_1d_2)^2+(a_1b_2-b_1a_2-c_1d_2+d_1c_2)^2\\ &+(a_1c_2+b_1d_2-c_1a_2-d_1b_2)^2+(a_1d_2-b_1c_2+c_1b_2-d_1a_2)^2.\end{aligned} \tag{2.68}$$

■

Lemma 5.11. Hvis p er et primtal, så findes x og y således at

$$x^2+y^2+1=mp, \tag{2.69}$$

hvor $0<m<p$.

Bevis. Eftersom $a^2\equiv b^2 \pmod p$ medfører at $a\equiv\pm b \pmod p$, da gælder det at tallene x^2 for $0\le x\le (p-1)/2$ alle er forskellige modulo p. Det samme er tallene $-(1+y^2)$ for $0\le y\le (p-1)/2$. Men da vi nu sammenlagt har

$$\frac{p-1}{2}+1+\frac{p-1}{2}+1=p+1$$

tal x^2 og $-(1+y^2)$, men kun p restklasser modulo p, da må der findes en løsning til

$$x^2\equiv -(1+y^2) \pmod p, \tag{2.70}$$

hvor $0\le x\le (p-1)/2$ og $0\le y\le (p-1)/2$. Dette giver os en løsning til $x^2+y^2+1\equiv 0 \pmod p$, som svarer til at $x^2+y^2+1=mp$. Bemærk at $m>0$, da det er en sum af kvadrater, og der gælder desuden at

$$m\le \frac{1}{p}(\tfrac{1}{4}p^2+\tfrac{1}{4}p^2+1)<p.$$

■

Sætning 5.12. Alle positive heltal kan skrives som en sum af fire kvadrater.

Bevis. Vi starter med at bemærke at $1,2 \in S_4$ eftersom

$$1 = 1^2 + 0^2 + 0^2 + 0^2 \quad \text{og} \quad 2 = 1^2 + 1^2 + 0^2 + 0^2.$$

Det vil altså være tilstrækkeligt for os at vise, at alle ulige primtal p ligger i S_4, da det følger af Aritmetikkens fundementalsætning og Lemma 5, at alle positive heltal så må være der. Lad nu m være det mindste tal så at der findes en løsning til ligningen

$$mp = a_1^2 + b_1^2 + c_1^2 + d_1^2, \tag{2.71}$$

Fra Lemma 5.11 ved vi at et sådant m må eksistere, da der findes x og y så $mp = x^2 + y^2 + 1^2 + 0^2$. Hvis $m = 1$ er vi færdige, så antag for modstrid at $m > 1$.

Vi betragter først tilfældet hvor m er lige. Men da må 0, 2 eller 4 af kvadraterne også være lige. Ved at bytte rundt på rækkefølgen, kan vi få $a_1 \equiv b_1 \pmod 2$ og $c_1 \equiv d_1 \pmod 2$, hvilket medfører at

$$\tfrac{1}{2}mp = \left(\frac{a_1 + b_1}{2}\right)^2 + \left(\frac{a_1 - b_1}{2}\right)^2 + \left(\frac{c_1 + d_1}{2}\right)^2 + \left(\frac{c_1 - d_1}{2}\right)^2. \tag{2.72}$$

Dette er i modstrid med vores antagelse om, at m er det mindste tal som opfylder $mp \in S_4$. Ovenstående ligning siger nemlig, at $\frac{1}{2}mp \in S_4$

Så m må være ulige.

Hvis m er ulige og $m \neq 1$, da kan a_1, b_1, c_1, d_1 ikke alle være delelige med m. Det ville betyde at $m^2 \mid mp$, og dermed at $m \mid p$. Vi kan vælge a_2, sådan så $a_2 \equiv a_1 \pmod m$ og $|a_2| < m/2$. På tilsvarende vis vælges b_2, c_2 og d_2. Vi ser nu at

$$0 < a_2^2 + b_2^2 + c_2^2 + d_2^2 < 4(m/2)^2 = m^2 \tag{2.73}$$

og

$$a_2^2 + b_2^2 + c_2^2 + d_2^2 \equiv 0 \pmod m. \tag{2.74}$$

Det følger heraf at

$$a_2^2 + b_2^2 + c_2^2 + d_2^2 = ms, \tag{2.75}$$

hvor $0 < s < m$. Fra Lemma 5 ser vi at

$$(mp)(ms) = m^2ps = x^2 + y^2 + z^2 + w^2, \tag{2.76}$$

hvor

$$x = a_1a_2+b_1b_2+c_1c_2+d_1d_2 \equiv a_1^2+b_1^2+c_1^2+d_1^2 \equiv 0 \pmod{m}. \tag{2.77}$$

Tilsvarende kan det vises at y, z og w også er delelige med m. Det giver os følgende:

$$sp = \left(\frac{x}{m}\right)^2 + \left(\frac{y}{m}\right)^2 + \left(\frac{z}{m}\right)^2 + \left(\frac{w}{m}\right)^2, \tag{2.78}$$

hvilket er i modstrid med at m skulle være det mindste.

Det følger hermed at $m = 1$, hvilket færdiggør beviset.

■

6 Opgaver

Delelighed

- **Opgave 6.1:**
  Hvilke af følgende tal deler hinanden?

4	9	27	31
63	12	36	20

- **Opgave 6.2:**
  Find samtlige divisorer for tallene 13, 64, 360.

- **Opgave 6.3:**
  Lad d og n være hele tal. Hvilke af følgende udsagn er sande og hvilke er falske?

a) $1 \mid n$	d) $d \mid n \Rightarrow n \mid d$
b) $0 \mid n$	e) $d \mid n \Rightarrow d \leq n$
c) $d \mid 0$	f) $a \mid b \Rightarrow a \mid -b$

- **Opgave 6.4:**
  Bevis Sætning 2.2, regel 2) og 3).

- **Opgave 6.5:**
  Hvilke af følgende tal er primtal og hvilke er sammensatte tal?

17	14	27	31	3
9	123	7	60	98765

- **Opgave 6.6:**
  Bestem primtalsopløsningen for hver af de følgende tal.

a) 33
b) 63
c) 84
d) 350
e) 550
f) 2565

•• **Opgave 6.7:**
Brug Euklids algoritme til at bestemme følgende værdier.

a) $\gcd(550, 84)$

b) $\gcd(550, 350)$

c) $\gcd(350, 84)$

Sammenlign resultaterne med de relevante primtalsopløsninger fra forrige opgave.

•• **Opgave 6.8:**
Brug Euklids algoritme til at bestemme største fælles divisor for følgende talpar.

a) 195, 154

b) 1771, 952

c) 1482, 935

d) 1970, 1066

Er nogle af disse talpar indbyrdes primiske?

•• **Opgave 6.9:**
Vis, at for tallene p og a, hvor p er et primtal, gælder at $\gcd(p,a) = 1$, hvis og kun hvis p ikke deler a.

•• **Opgave 6.10:**
Ifølge Bezouts identitet kan vi finde to hele tal, s og t således at

$$\gcd(15, 19) = 15s + 19t.$$

Find et talpar (s,t) der opfylder ligningen.

●●● **Opgave 6.11:**
Bevis regnereglerne fra Sætning 2.10 om den største fælles divisor.

●●● **Opgave 6.12:**
Hvad er den største potens af 5 som deler 1000! og den største potens af 2 som deler 1000!

Kongruenser

- **Opgave 6.13:**
Gruppér tal der er kongruente modulo 3.

17	14	27	31	41
9	300	7	64	0

●●● **Opgave 6.14:**
Lad $n \in \mathbb{N}$. Bevis, at de følgende udsagn er sande for alle a, b, $c \in \mathbb{Z}$:

a) $a \equiv a \pmod{n}$

b) $a \equiv b \pmod{n} \quad \Rightarrow \quad b \equiv a \pmod{n}$

c) $a \equiv b$ og $b \equiv c \pmod{n} \Rightarrow a \equiv c \pmod{n}$

- **Opgave 6.15:**
Udregn følgende udtryk:

a) $[23]_5 + [101]_5$

b) $[58]_{60} \cdot [59]_{60}$

c) $[31]_8 \cdot [33]_8$

d) $[3578]_{11} \cdot ([36]_{11} + [8]_{11})$

●●● **Opgave 6.16:**
Bevis følgende udsagn eller find et modeksempel.

$$[a^b] = [a]^{[b]}$$

● **Opgave 6.17:**
Forestil dig et ur med den lille viser på 12. Brug din viden om restklasser til at bestemme hvor viseren står når der er gået

a) 100 timer.

b) $25 \cdot 100$ timer.

c) 25^4 timer.

Hint: Vi kan se på viserens position, som en restklasse modulo 12. For eksempel er $[10]_{12} = [22]_{12}$ ligesom at viseren står samme sted på uret klokken 10.00 og klokken 22.00.

●● **Opgave 6.18:**
Hvad er de to sidste cifre i tallet $11111 \cdot 10203$?

●● **Opgave 6.19:**
Vis, at 6 går op i $n(n+1)(2n+1)$ for alle heltal n.

Hint: Det er nok at tjekke om 2 og 3 går op i udtrykket på grund af 6's primtalsopløsning.

●●● **Opgave 6.20:**
Bevis, når vi regner modulo 3, at ethvert heltal kongruent med sin tværsum.
Gælder det også når vi regner modulo 9? Hvad med modulo 10?

●● **Opgave 6.21:**
Find inverser til følgende;

a) $2 \pmod 7$ c) $-1 \pmod{987}$

b) $5 \pmod{13}$ d) $7 \pmod{24}$

- **Opgave 6.22:**
  Beregn følgende værdier:

  a) $\varphi(15)$ d) $\varphi(7)$

  b) $\varphi(4)$ e) $\varphi(21)$

  c) $\varphi(18)$ f) $\varphi(25)$

●●● **Opgave 6.23:**
For $n > 1$ vis at summen af tallene mindre end n og indbyrdisk primiske med n er $n \cdot \varphi(n)/2$.

- **Opgave 6.24:**
  Find ordenen af 2, 3 og 5 modulo 19.

Kvadratiske rester

- **Opgave 6.25:**
  Find alle de kvadratiske rester for primtallene 3, 5, 7, 11, 13, 17, 19.

- **Opgave 6.26:**
  Udregn følgende Legendre symboler:

  a) $\left(\frac{5}{7}\right)$ d) $\left(\frac{91}{167}\right)$ g) $\left(\frac{31}{167}\right)$

  b) $\left(\frac{3}{11}\right)$ e) $\left(\frac{11}{37}\right)$ h) $\left(\frac{5}{160465489}\right)$

  c) $\left(\frac{6}{13}\right)$ f) $\left(\frac{19}{31}\right)$ i) $\left(\frac{3083}{3911}\right)$

- **Opgave 6.27:**
  Find alle løsninger i $\mathbb{Z}/7\mathbb{Z}$ til ligningerne:

  a) $x^2 - 3x + 4 \equiv 0 \pmod 7$

b) $x^2 + 4x - 5 \equiv 0 \pmod 7$

c) $x^2 - 2x - 1 \equiv 0 \pmod 7$

•• **Opgave 6.28:**
Find antallet af løsninger til ligningerne

a) $x^2 + 1 \equiv 0 \pmod{83}$

b) $x^2 + x + 1 \equiv 0 \pmod{83}$

c) $x^2 + 21x - 11 \equiv 0 \pmod{83}$

d) $x^2 + x + 21 \equiv 0 \pmod{83}$

Hint: Kig på diskriminanten.

•• **Opgave 6.29:**
Vi kalder r for kvadratroden af a modulo p, hvis $r^2 \equiv a \pmod p$. Lad $p \equiv 3 \pmod 4$ og lad a være en kvadratisk rest modulo p. Vis, at $a^{\frac{p+1}{4}}$ er en kvadratrod af a modulo p.

•• **Opgave 6.30:**
Vis at antallet af løsninger til $x^2 \equiv a \pmod p$ er givet ved $\left(\frac{a}{p}\right) + 1$.

•• **Opgave 6.31:**
Vis, at følgende holder

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & \text{hvis } p \equiv 1 \text{ eller } 7 \pmod 8, \\ -1 & \text{hvis } p \equiv 3 \text{ eller } 5 \pmod 8. \end{cases}$$

•••• **Opgave 6.32:**
Lav formler som i forrige opgave for $\left(\frac{3}{p}\right)$ og $\left(\frac{5}{p}\right)$.

•• **Opgave 6.33:**
Bevis, at

$$\left(\frac{1}{p}\right) + \left(\frac{2}{p}\right) + \cdots + \left(\frac{p-1}{p}\right) = 0.$$

•• **Opgave 6.34:**
Vis at en kvadratisk rest aldrig kan være en primitiv rod modulo p.
Hint: Kig på diskriminanten.

••• **Opgave 6.35:**
Vis at der findes a og k så $a^2 = -1 + kp$, hvis og kun hvis $p = 1 \pmod 4$.

•• **Opgave 6.36:**
Lad p være et primtal. Vis for alle a,$b \in \mathbb{Z}$ at mindst en af a, b eller $a \cdot b$ vil være kvadratisk rest modulo p.

•• **Opgave 6.37:**
I denne opgave betragter vi polynomiet $f(x) = (x^2-7)(x^2-13)(x^2-91)$

a) Argumenter for at polynomiet har rødder modulo p for vilkårlige primtal.

b) Argumenter for at polynomiet ikke har nogle rødder i heltallene.

Summer af kvadrater

•• **Opgave 6.38:**
Skriv følgende tal som summer af to kvadrater eller vis at dette er umuligt:

$$130, 260, 847, 980 \text{ og } 1073$$

• **Opgave 6.39:**
Find alle talpar $(x,y) \in \mathbb{Z}^2$ som opfylder

$$x^2 + y^2 = 50.$$

•• **Opgave 6.40:**
Vi ser at $13 = 2^2 + 3^2$ og $29 = 2^2 + 5^2$.
Hvordan kan vi skrive $377 = 13 \cdot 29$ som sum af to kvadrater?

• **Opgave 6.41:**
På hvor mange måder kan man skrive tallene 11 og 14 som sum af tre kvadrater?

•• **Opgave 6.42:**
Find et modeksempel som viser at S_3 ikke er lukket under multiplikation.

••• **Opgave 6.43:**
I denne opgave kigger vi på mængden S_3.

a) Vis at hvis $n \in S_3$ så er $n \not\equiv 7 \pmod 8$.

b) Vis at hvis $n \in S_3$ og $4 \mid n$, så er $\frac{n}{4}$ et element i S_3.

c) Udled at hvis $n = 4^m(8k + 7)$ så er $n \notin S_3$.

Fun fact: del c) gælder faktisk som hvis og kun hvis, altså n er kun i S_3 netop når n ikke er på formen $n = 4^m(8k + 7)$.

•• **Opgave 6.44:**
Skriv følgende tal som sum af fire kvadrater:

247, 308 og 465.

Hint: Brug at S_4 er multiplikativt lukket, og kig i beviset for at dette gælder.

• **Opgave 6.45:**
På hvor mange måder kan man skrive tallet 28 som en sum af fire kvadrater?

•• **Opgave 6.46:**
Vis at hvis $8 \mid (a_1^2 + a_2^2 + a_3^2 + a_4^2)$ så må tallene a_1, a_2, a_3 og a_4 alle være lige.

Hint: Betragt de kvadratiske rester modulo 8.

••• **Opgave 6.47:**
Vi ser at $2 \cdot 877 = 1754 = 27^2 + 7^2 + 20^2 + 24^2$. Find en repræsentation for 877 som sum af fire kvadrater.

Hint: Kig i beviset for summer af fire kvadrater.

7 Projekt: Perfekte Tal

I dette projekt betragter vi perfekte tal og viser en sammenhæng mellem lige *perfekte tal* og nogle specielle primtal kaldet *Mersenne primtal*.

Vi starter med at definere de begreber som projektet omhandler.

Definition 7.1 (Perfekt tal)**.** Et naturligt tal n kaldes for et *perfekt tal*, hvis tallet er lig summen af tallets divisorer foruden tallet selv.

Eksempel 7.2. Tallet 496 er et perfekt tal, eftersom

$$496 = 1 + 2 + 4 + 8 + 16 + 31 + 62 + 124 + 248.$$

∘

Opgave 7.1:
Find mindst et eksempel på et perfekt tal.

Hint: Der eksisterer 2 perfekte tal der er mindre end 100.

Definition 7.3 (Mersenne primtal)**.** Et primtal p som kan skrives på formen $2^n - 1$ kaldes for et *Mersenne primtal*.

Eksempel 7.4. Tallene 3, 7 og 31 er alle Mersenne primtal, da de er primtal og kan skrives som henholdvis $2^2 - 1$, $2^3 - 1$ og $2^5 - 1$. ∘

Mersenne primtal er opkaldt efter franskmanden Marin Mersenne, som betragtede dem i det 17. århundrede. Det er stadig et uløst matematisk spørgsmål, om der findes endelig eller uendelig mange Mersenne primtal. Der er dog en formodning om, at der eksisterer uendelig mange af dem. Men indtil nu kender man kun ca. 50.

Opgave 7.2:
Vis at hvis $2^n - 1$ er et primtal, da skal n også være et primtal.

Hint: Benyt at $a^k - 1 = (a - 1)(a^{k-1} + a^{k-2} + \cdots + a + 1)$.

Definition 7.5 (Divisorfunktionen)**.** Divisorfunktionen $\sigma(n)$ er defineret som summen af alle de positive divisorer for n, altså

$$\sigma(n) = d_1 + d_2 + \cdots + d_n, \tag{2.79}$$

hvor $d_i \mid n$. Bemærk at både 1 og n må optræde blandt d_i-erne.

Eksempel 7.6. Lad $n = 12$. Så har vi

$$\sigma(12) = 1 + 2 + 3 + 4 + 6 + 12 = 28. \tag{2.80}$$

∘

Vi har nu brug for at vise en vigtig egenskab ved divisorfunktionen, nemlig at det er en multiplikativ funktion.

Definition 7.7 (Multiplikativ funktion)**.** Vi kalder en funktion for *multiplikativ*, hvis den opfylder

$$f(ab) = f(a)f(b) \quad \text{når} \quad \gcd(a,b) = 1. \tag{2.81}$$

Eftersom σ er multiplikativ, så givet primtalsfaktoriseringen af n, vil vi nemt kunne udregne $\sigma(n)$, hvis blot vi kan finde udtryk for σ i alle primtalspotenser.

Opgave 7.3:
Lad p være et primtal. Find et udtryk for

(a) $\sigma(p)$

(b) $\sigma(p^k)$

Benyt derefter dine formler på nogle små p og k, og bekræft at de giver de rigtige resultater.

Hint: Hvis dit udtryk bliver meget langt, så prøv at gange igennem med $p - 1$.

Det vil være rart, hvis vi kan finde en betingelse for perfekte tal, som er lettere at arbejde med end definitionen af dem.

Opgave 7.4:
Vis at et tal n er perfekt hvis og kun hvis at

$$\sigma(n) = 2n. \tag{2.82}$$

Hvis vi betragter 496 og de andre perfekte tal der blev fundet i Opgave 7.1, kan vi se at de alle kan skrives på formen

$$n = 2^{k-1}(2^k - 1). \tag{2.83}$$

Lad os undersøge om der eksisterer flere perfekte tal på denne form.
Opgave 7.5:
Lav en formodning om for hvilke k det gælder at $2^{k-1}(2^k - 1)$ er et perfekt tal.

Hint: Prøv at indsætte i formlen fra Opgave 7.4.

Vi vil nu arbejde os igennem beviset for sætningen at det også gælder den modsatte vej.

Sætning 7.8. Et tal n er et lige, perfekt tal, hvis og kun hvis det er på formen $2^{p-1}(2^p - 1)$, hvor $(2^p - 1)$ er et Mersenne primtal.

Bevis. Da n er et lige tal, kan vi dele n op som

$$n = 2^k \cdot u, \tag{2.84}$$

hvor $k \geq 1$ og u er ulige.
Opgave 7.6:
Start med at vise at

$$(2^{k+1} - 1)\,\sigma(u) = 2^{k+1}u. \tag{2.85}$$

Opgave 7.7:
Argumenter for at u må kunne skrives som

$$u = (2^{k+1} - 1)q, \tag{2.86}$$

hvor q er et heltal.

Opgave 7.8:
Vis nu at

$$\sigma(u) = 2^{k+1}q. \tag{2.87}$$

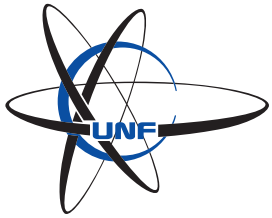
Opgave 7.9:
Vis derefter at $u + q = \sigma(u)$ og benyt dette til at vise at $q = 1$.

Hint: Læg mærke til at $q \mid u$.

Opgave 7.10:
Færdiggør beviset.

■

Det vides ikke om der eksisterer ulige perfekte tal. Man har dog vist, at hvis de findes, skal de være større end 10^{200} og have mere end 8 forskellige primfaktorer.



3 Ringteori

1 Introduktion

Grundlæggende egenskaber for ringe

I dette forløb skal vi studere ringe. Ringteori er en del af algebra, som i meget korte træk går ud på at studere mængder med struktur. Strukturen består i nogle såkaldte *kompositionsregler*, der i daglig tale kan kaldes regneregler. Vi lægger ud med at definere en ring helt formelt. Definitionen er lang og måske skræmmende, men man har heldigvis ikke brug for at huske den udenad. Det er som regel tilstrækkeligt at tænke på en ring som en mængde, hvor vi må lægge ting til hinanden og gange dem sammen, og de to regneregler spiller godt sammen, som vi kender det fra tal. For at kunne definere en ring, skal vi definere en kompositionsregel.

Definition 1.1. Lad A være en mængde. En *kompositionsregel* $*$ er en afbildning $* : A \times A \to A$. I stedet for $*(a_1, a_2)$ skriver vi $a_1 * a_2$.

Vi er allerede bekendte med kompositionsregler fra dagligdagen. En vigtig ting at bemærke er, at kompositionsreglen på to elementer igen skal være et element i mængden. Sagt lidt løst: Kompositionsreglen må ikke "tage os ud af mængden".

Eksempel 1.2. Betragt de reelle tal $\mathbb{R}$. $+$ (addition), $\cdot$ (multiplikation) og $-$ (subtraktion) er alle kompositionsregler. Dog er $/$ (di-

vision) ikke en kompositionsregel, idet $a/0$ ikke er et element i $\mathbb{R}$ (division med nul er ikke defineret). ∘

Eksempel 1.3. Betragt de naturlige tal $\mathbb{N}$. $+$ og $\cdot$ er igen kompositionsregler, thi summen og produktet af to naturlige tal igen er et naturligt tal. Dog er $-$ ikke en kompositionsregel. F.eks. er $2 - 3 = -1$, og -1 er ikke et element i $\mathbb{N}$. ∘

Kompositionsregler kan have bestemte egenskaber. Nogle af disse er skrevet i definitionen herunder.

Definition 1.4. Lad $*$ være en kompositionsregel på mængden A.

- $*$ kaldes *associativ* , hvis $a*(b*c) = (a*b)*c$ for alle $a, b, c \in A$.
- $*$ kaldes *kommutativ* , hvis $a * b = b * a$ for alle $a, b \in A$.

Definition 1.5. En *ring* R er en mængde R med to kompositionsregler betegnet $+$ og $\cdot$, som opfylder følgende regler:

1. $+$ er associativ.
2. $+$ er kommutativ.
3. Der findes et element $0 \in R$, så $a + 0 = a$ for alle $a \in R$.
4. For alle $a \in R$ findes et element $b \in R$, så $a + b = 0$. b betegnes som regel $-a$.
5. $\cdot$ er associativ.
6. For alle $a, b, c \in R$ gælder

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{og} \quad (b + c) \cdot a = b \cdot a + c \cdot a.$$

Bemærkning 1.6. Idet matematikere er dovne, vil vi fra nu af undlade at skrive $\cdot$ så ofte, vi kan slippe afsted med det. Med andre ord, hvis $a, b \in R$, da vil vi blot skrive ab i stedet for $a \cdot b$.

Der er en masse terminologi associeret med reglerne i definitionen ovenover. 0 i punkt 3 kaldes nogle gange *neutralelementet for addition*. b i punkt 4 kaldes den *additive invers* til a. De to regneregler i punkt 6 kaldes de *distributive love*. Det er ikke essentielt, at I husker alle disse navne. Ud fra reglerne for en ring, kan vi sige følgende.

Proposition 1.7. Lad R være en ring. Da vil $0 \cdot a = a \cdot 0 = 0$ for alle $a \in R$.

Bevis. Vi bemærker, at $0 = 0 + 0$ per regel 3. Vi kan nu benytte regel 6 og få

$$a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0.$$

Regel nummer 4 siger, at der findes et element $-a \cdot 0$, så $a \cdot 0 + (-a \cdot 0) = 0$. Lægges $-a \cdot 0$ til på begge sider, fås $0 = a \cdot 0$ som ønsket. Ligheden $0 \cdot a = 0$ fås ved at bruge den anden del af regel 6 og gentage udregningerne.

■

Lad os se nogle simple eksempler (og ikke-eksempler) på ringe.

Eksempel 1.8. Den simpleste ring, man kan konstruere, er *nulringen* $R = \{0\}$. Her er kompositionsreglerne defineret ved $0 + 0 = 0$ og $0 \cdot 0 = 0$. ○

Eksempel 1.9. $\mathbb{Z}$ udgør klart en ring med de sædvanlige regneoperationer $+$ og $\cdot$. Det samme gælder $\mathbb{R}$. ○

Eksempel 1.10. Betragt de rationale tal $\mathbb{Q}$. Idet

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{cd} \quad \text{og} \quad \frac{a}{b}\frac{c}{d} = \frac{ac}{bd},$$

ses det, at summen og produktet af to brøker igen er en brøk. Dermed er sædvanlig addition og multiplikation kompositionsregler, og det ses let, at disse opfylder alle reglerne i definitionen af en ring. Dermed er $\mathbb{Q}$ en ring. ○

Eksempel 1.11. Betragt $\mathbb{N}$ med operationerne $+$ og $\cdot$. Dette er ikke en ring. F.eks. findes der i $\mathbb{N}$ ikke et element 0, så $n + 0 = n$ for alle n, idet det mindste tal i $\mathbb{N}$ er 1. Selv hvis man medtager 0 i $\mathbb{N}$, går det galt. F.eks. findes intet naturligt tal n, så $1 + n = 0$. ∘

Eksempel 1.12. Lad os betragte mængden $2\mathbb{Z} = \{2n \mid n \in \mathbb{Z}\} = \{...-4, -2, 0, 2, 4, ...\}$ (de lige tal). Idet summen og produktet af lige tal igen er lige tal, og ringegenskaberne nedarves fra $\mathbb{Z}$, må $2\mathbb{Z}$ være en ring. ∘

I den kommende diskussion ønsker vi at udelukke eksempler som $2\mathbb{Z}$. Årsagen er, at der i $2\mathbb{Z}$ mangler et 1. Med 1 i en generel ring menes et element, som opfylder $1a = a1 = a$ for alle a i ringen (sådan et element kaldes en *multiplikativ invers*). Tænk på 1 som tallet 1 fra $\mathbb{Z}$, $\mathbb{R}$ og så videre. Bemærk også i definitionen af en ring, at vi ikke kræver, at multiplikationen $\cdot$ er kommutativ. Dog kommer vi i dette forløb kun til at støde på kommutative ringe, dvs. ringe hvor $\cdot$ er en kommutativ kompositionsregel. Derfor har vi følgende konvention.

Konvention: Alle ringe R i forløbet opfylder, at $\cdot$ er kommutativ, og at der eksisterer et $1 \in R$, så $1a = a1 = a$ for alle $a \in R$.

Lad os nu kigge på en type ring, der vil dukke op ofte i forløbet.

Definition 1.13. Lad R være en ring. Vi lader $R[x]$ betegne mængden af polynomier i variablen x med koefficienter i R. Elementerne i $R[x]$ kan alle skrives på formen

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

hvor $a_0, a_1, ..., a_n \in R$ kaldes polynomiets *koefficienter*. n er et ikke-negativt heltal, der kan variere. Mængden $R[x]$ kan betragtes som en ring med addition og multiplikation givet ved almindelig addition og multiplikation af polynomier. $R[x]$ kaldes *polynomiumsringen* over R.

Eksempel 1.14. Lad os se på $\mathbb{Z}[x]$. Eksempler på elementer i $\mathbb{Z}[x]$ kunne være

$$2x^4 - 5x^3 + x - 8, \quad x^3 + 4x^2 + 2x, \quad 4.$$

$4 \in \mathbb{Z}[x]$ er et eksempel på et såkaldt konstant polynomium. Addition og multiplikation i $\mathbb{Z}[x]$ fungerer præcist, som vi forventer:

$$\begin{aligned}(x^2+3) + (2x^3 - 4x + 1) &= 2x^3 + x^2 - 4x + 4 \\ (x^2+3)(2x^3 - 4x + 1) &= x^2 2x^3 - x^2 4x + x^2 \\ &\quad + 3 \cdot 2x^3 + 3(-4)x + 3 \\ &= 2x^5 + 2x^3 + x^2 - 12x + 3.\end{aligned}$$

Logikken er den samme i ringene $\mathbb{Q}[x]$ og $\mathbb{R}[x]$. ∘

Nu har vi fastlagt, hvad en ring er, og vi har set nogle centrale eksempler. Vi kan nu så småt begynde at klassificere ringe.

Definition 1.15. Lad R være en ring.

- Et element $u \in R$ kaldes en *enhed*, hvis der eksisterer et element $v \in R$, så $uv = vu = 1$. v kaldes da en invers til u (mht. multiplikation).
- Et element $a \in R$ med $a \neq 0$ kaldes en *nuldivisor*, hvis der eksisterer et element $b \in R$, $b \neq 0$, så $ab = ba = 0$.

Vi bemærker, at 1 altid er en enhed, thi $1 \cdot 1 = 1$, og dermed er 1 sin egen invers. Det er dog ikke altid tilfældet, at der findes andre enheder end 1 i en ring. Vi skal se sådan et eksempel om ikke så længe. Med disse definitioner kan vi formulere følgende resultat.

Proposition 1.16. Lad R være en ring.

1. Hvis u er en enhed, er den inverse unik.
2. To elementer a og b i R er enheder hvis og kun hvis ab er en enhed.

3. Et element i R kan ikke både være en nuldivisor og en enhed.

Bevis. 1. Lad u være en enhed med to inverser v og w. Da har vi

$$v = 1v = (wu)v = w(uv) = w1 = w,$$

så v og w er ens. Dermed kan der kun findes én invers.

2. Antag, at a og b er enheder med inverser a^{-1} og b^{-1}. Da har vi

$$(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = a1a^{-1} = aa^{-1} = 1$$

og ligeledes er $(b^{-1}a^{-1})(ab) = 1$ ved at benytte, at R er kommutativ. Dermed er $b^{-1}a^{-1}$ en invers til ab, så ab er en enhed. Antag omvendt, at ab er en enhed med invers $(ab)^{-1}$. Da har vi

$$a(b(ab)^{-1}) = (ab)(ab)^{-1} = 1$$

og igen ved at benytte, at R er kommutativ, fås $(b(ab)^{-1})a = 1$. Altså er $a^{-1} = b(ab)^{-1}$, så a er en enhed. På samme måde vises det, at b er en enhed, thi den har den inverse $a(ab)^{-1}$.

3. Antag for modstrid, at a er både en enhed og en nuldivisor. Da a er en enhed, findes et b, så $ab = ba = 1$. Da a er en nuldivisor, findes et c, så $ac = ca = 0$. Samtidig skal det gælde, at $a, c \neq 0$. Da $0 = ac$ og $ba = 1$, må vi have $b \cdot 0 = (ba)c = 1c$, altså $0 = c$, men dette er en modstrid. Altså kan et element ikke være både en nuldivisor og en enhed.

■

Bemærkning 1.17. I punkt 2 af ovenstående proposition er det nødvendigt, at R er kommutativ. Det tilsvarende resultat for ikke-kommutative ringe er, at a og b er enheder hvis og kun hvis ab og ba er enheder.

Ovenstående resultat fortæller os, at hvis et element u i en ring er en enhed, da er dens invers unikt bestemt. Vi vil betegne denne invers med u^{-1}, hvilket vi også benyttede os af i løbet af beviset. Det

er ikke alle ringe, der har nuldivisorer. Sådanne ringe betragtes ofte som særligt pæne, og de har deres eget navn.

Definition 1.18. En ring R kaldes et *integritetsområde*, hvis R ikke indeholder nogle nuldivisorer.

En ring R er altså et integritetsområde, hvis $ab = 0$ betyder, at enten $a = 0$ eller $b = 0$. Dermed kan man sige, at en ring er et integritetsområde, hvis nulreglen gælder.

Eksempel 1.19. Både $\mathbb{Z}, \mathbb{Q}$ og $\mathbb{R}$ er integritetsområder, da den eneste måde, man kan få $a \cdot b = 0$, er hvis $a = 0$ eller $b = 0$. ∘

Eksempel 1.20. Lad os betragte et mere abstrakt eksempel på en ring. Lad $R = P(\mathbb{N})$ være potensmængden af de naturlige tal, altså mængden af alle delmængder af $\mathbb{N}$. F.eks. er $\{1\}$ og $\{2, 3, 8\}$ elementer i R. Vi udstyrer R med to kompositionsregler, der gør R til en ring. Vi lader additionen være givet ved $A \Delta B = (A \setminus B) \cup (B \setminus A)$ kaldet den *symmetriske differens*, mens multiplikationen er givet ved $A \cap B$, altså fællesmængden. Nogle konkrete eksempler kunne være

$$\{1, 2, 3\} \Delta \{3, 4, 5\} = \{1, 2\} \cup \{4, 5\} = \{1, 2, 4, 5\},$$
$$\{1, 2, 3\} \cap \{3, 4, 5\} = \{3\}.$$

Det er en tidskrævende affære at vise, at de to kompositionsregler Δ og $\cap$ på R faktisk gør R til en ring, så vi springer det over. Læseren kan dog med fordel tegne nogle Venn-diagrammer for at overbevise sig selv om det. Det ses, at $0 = \emptyset$, thi for alle $A \in R$ gælder

$$A \Delta \emptyset = (A \setminus \emptyset) \cup (\emptyset \setminus A) = A \cup \emptyset = A,$$

og ligeledes er $\emptyset \Delta A = A$. Det ses, at $1 = \mathbb{N}$, eftersom det for alle A gælder, at

$$A \cap \mathbb{N} = A = \mathbb{N} \cap A.$$

Vi bemærker, at kun 1 er en enhed, og at der findes et væld af nuldivisorer. Faktisk er alle elementer udover $\mathbb{N}$ en nuldivisor, idet

hvis $A \neq \mathbb{N}$, vil $\mathbb{N} \setminus A \neq \emptyset$ og $A \cap (A \setminus \mathbb{N}) = \emptyset$. Specielt er R ikke et integritetsområde.

$\circ$

Vi kommer til at se flere eksempler på ringe senere, der ikke er integritetsområder.

Proposition 1.21. Lad R være et integritetsområde.

1. Lad $a, b, c \in R$. Hvis $ab = ac$ for $a \neq 0$ i R, da vil $b = c$.
2. $R[x]$ er et integritetsområde.

Bevis. 1. Antag, at $ab = ac$. Da har vi $0 = ab - ac = a(b - c)$. Da R er et integritetsområde, skal vi have $a = 0$ eller $b - c = 0$. Da vi har antaget, at $a \neq 0$, må $b - c = 0$, og dermed er $b = c$ som ønsket.

2. Lad $f(x), g(x) \in R[x]$, og antag, at $f(x), g(x) \neq 0$. Hvis a_n betegner den førende koefficient (koefficienten på den højeste potens af x) for $f(x)$ og b_m den førende koefficient for $g(x)$, da er den førende koefficient for $f(x)g(x)$ lig $a_n b_m$. Da a_n og b_m er forskellige fra 0, er $a_n b_m \neq 0$, thi R er et integritetsområde. Dermed er $f(x)g(x) \neq 0$. Vi har altså vist, at hvis $f(x)$ og $g(x)$ ej er lig 0, da er $f(x)g(x)$ heller ikke lig 0. Kontraponeres udsagnet fås, at $f(x)g(x) = 0$ medfører, at enten $f(x) = 0$ eller $g(x) = 0$. Dette viser, at $R[x]$ er et integritetsområde.

■

Nu har vi studeret nuldivisorer. Lad os vende tilbage til enheder.

Definition 1.22. En ring, hvor alle elementer bortset fra 0 er en enhed, kaldes et *legeme*.

Eksempel 1.23. $\mathbb{Q}$ er et legeme. Hvis $a/b \neq 0$, hvor a og b er heltal, da er b/a en invers, thi

$$\frac{a}{b}\frac{b}{a} = \frac{ab}{ab} = 1.$$

$\mathbb{R}$ er også et legeme. ∘

Eksempel 1.24. $\mathbb{Z}$ er ikke et legeme. De eneste enheder i $\mathbb{Z}$ er 1 og -1. ∘

Når vi skal studere kvotientringe, får vi metoder til at konstruere legemer og integritetsområder. Det følger af Proposition 1.16, at et legeme er et integritetsområde. Den anden vej gælder ikke generelt, som eksemplet med $\mathbb{Z}$ viser. Dog (måske overraskende) gælder den modsatte implikation, når ringen er endelig.

Proposition 1.25. Lad R være et endeligt integritetsområde. Da er R et legeme.

Bevis. Lad $a \in R$ være forskellig fra 0. Vi skal vise, at a har en invers. Betragt afbildningen $f : R \to R$ givet ved $f(x) = ax$. Vi hævder, at f er injektiv. Antag, at $f(x_1) = f(x_2)$. Dette er det samme som $ax_1 = ax_2$. Da $a \neq 0$, giver Proposition 1.21, at $x_1 = x_2$. Dermed er f injektiv. Da R er endelig, og f er en injektiv afbildning fra R til sig selv, må f være surjektiv. Dermed findes et element $b \in R$, så $f(b) = 1$, altså $ab = 1$ for dette element. Dermed er a en enhed. Da a var et arbitrært ikke-nul element, er beviset færdigt. ■

At konstruere nye ringe ud fra gamle

Når matematikere begynder at studere et objekt, vil de som regel være interesserede i, hvordan man ud fra ét eller flere af disse kan konstruere nye objekter af samme type. Ligeledes vil vi spørge os selv, hvordan vi kan lave nye ringe ud fra gamle. Der er flere måder at gøre dette på. Én måde er gennem delringe.

Definition 1.26. Lad R være en ring. En delmængde $S \subseteq R$ kaldes en *delring*, hvis S i sig selv er en ring med kompositionsreglerne fra R.

Eksempel 1.27. $\mathbb{Z}$ er en delring af $\mathbb{Q}$, og $\mathbb{Q}$ er en delring af $\mathbb{R}$. Ligeledes er $\mathbb{Z}[x]$ en delring af $\mathbb{Q}[x]$, der igen er en delring af $\mathbb{R}[x]$. ∘

Eksempel 1.28. Lad R være en ring, og betragt polynomiumsringen $R[x]$. Betragt delmængden S bestående af alle polynomier, der kun har led med lige potenser af x (inklusiv potensen 0, der svarer til konstantleddet). Eksempler på elementer i S er

$$x^2 + 1, \quad 3x^4 - 2x^2 + 10 \quad \text{og} \quad -3x^6 + x^4 + 3x^2.$$

Er S en delring af $R[x]$? Vi ser, at 0 og 1 ligger i S. Hvis $f(x) \in S$, ligger $-f(x)$ også i S. Lad nu $f(x), g(x) \in S$. Vi skal undersøge, om $f(x) + g(x) \in S$ og $f(x)g(x) \in S$. Det er ikke svært at se, at $f(x) + g(x)$ kun indeholder led med en lige potens af x. Det samme gælder om $f(x)g(x)$, da alle potenser af x i produktet fremkommer ved at summere potenser af x i $f(x)$ og $g(x)$, og summen af to lige tal igen er et lige tal. Dermed er S en ring og altså en delring af $R[x]$. Hvad med mængden af polynomier, der kun har ulige potenser af x? Dette er ikke en delring. F.eks. er x et element af denne mængde, men $x \cdot x = x^2$ er ikke, så kompositionsreglen $\cdot$ tager os ud af mængden.

○

En anden måde at konstruere nye ringe ud fra gamle er gennem produktmængder.

Definition 1.29. Lad $R_1, R_2, ..., R_n$ være ringe og betragt produktmængden $R_1 \times R_2 \times \cdots \times R_n$. Ved at lave addition og multiplikation koordinatvist, bliver $R_1 \times R_2 \times \cdots \times R_n$ til en ring, som vi betegner som en *produktring*. For en ring R betegner R^n produktet $R \times R \times \cdots \times R$ med n kopier af R.

Vi lader læseren overbevise sig selv om, at 0 og 1 i $R_1 \times R_2 \times \cdots \times R_n$ er givet ved hhv. $(0, 0, ..., 0)$ og $(1, 1, ..., 1)$.

Eksempel 1.30. $\mathbb{Z}^3 = \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$. Som eksempel kan vi betragte elementerne $(1, -2, 3), (-5, 0, 4) \in \mathbb{Z}^3$. Da har vi

$$(1, -2, 3) + (-5, 0, 4) = (-4, -2, 7),$$
$$(1, -2, 3) \cdot (-5, 0, 4) = (-5, 0, 12).$$

○

Eksempel 1.31. Betragt ringen $\mathbb{Q} \times \mathbb{Z}[x]$. Eksempler på elementer i denne ring er

$$\left(\frac{1}{2}, 3x^2 + 5\right), \quad (5, -3x), \quad \left(-\frac{3}{2}, x^3 + 6x + 2\right).$$

Lad os se på et eksempel på regning i denne ring:

$$\left(-\frac{1}{4}, 4x^2 - 1\right) + \left(\frac{3}{5}, x + 2\right) = \left(\frac{7}{20}, 4x^2 + x + 1\right)$$
$$\left(-\frac{1}{4}, 4x^2 - 1\right) \cdot \left(\frac{3}{5}, x + 2\right) = \left(-\frac{3}{20}, 4x^3 + 8x^2 - x - 2\right).$$

∘

Ringhomomorfier

Nu har vi set en række eksempler på ringe, og vi har fået nogle grundlæggende definitioner på plads. Vi skal nu kort se på, hvordan man relaterer ringe til hinanden. Idéen er at definere en afbildning fra en ring til en anden, der respekterer kompositionsreglerne i de to ringe.

Definition 1.32. Lad R og S være ringe. En *ringhomomorfi* $\varphi : R \to S$ er en afbildning, som opfylder følgende to regler for alle $a, b \in R$:

1. $\varphi(a + b) = \varphi(a) + \varphi(b)$.
2. $\varphi(ab) = \varphi(a)\varphi(b)$.

Bemærkning 1.33. Vi vil som regel bruge græske bogstaver såsom φ (kaldet "phi") og ψ (kaldet "psi") for ringhomomorfier.

Eksempel 1.34. Hvis R og S er ringe, da er afbildningen $\varphi : R \to S$ givet ved $\varphi(a) = 0$ for alle $a \in R$ en ringhomomorfi. Denne afbildning kaldes *nulafbildningen*. ∘

Eksempel 1.35. Betragt afbildningen $\varphi : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ givet ved $\varphi((a,b)) = a + b$. Er φ en ringhomomorfi? For $(a,b),(c,d) \in \mathbb{Z} \times \mathbb{Z}$ har vi

$$\begin{aligned}\varphi((a,b)+(c,d)) &= \varphi((a+c,b+d)) = (a+c)+(b+d)\\ &= (a+b)+(c+d) = \varphi((a,b))+\varphi((c,d)),\end{aligned}$$

så den første betingelse er opfyldt. Dog ser vi, at det går galt for multiplikationen. F.eks. kan vi se på elementerne $(1,0)$ og $(0,1)$. Da har vi

$$\varphi((1,0)\cdot(0,1)) = \varphi((0,0)) = 0+0 = 0,$$

men

$$\varphi((1,0))\varphi((0,1)) = (1+0)(0+1) = 1.$$

Dermed er φ ikke en ringhomomorfi.

∘

Eksempel 1.36. Lad R være en ring og betragt afbildningen $\varphi : R[x] \to R$ givet ved $\varphi(f(x)) = f(0)$. Det ses, at φ til polynomiet $f(x)$ giver os konstantleddet. Vi hævder, at φ er en ringhomomorfi. Lad $f(x), g(x) \in R[x]$. Da har vi

$$\varphi(f(x)+g(x)) = f(0)+g(0) = \varphi(f(x))+\varphi(g(x))$$

og

$$\varphi(f(x)g(x)) = f(0)g(0) = \varphi(f(x))\varphi(g(x)),$$

hvilket viser det ønskede. Som konkret eksempel kunne vi lade $R = \mathbb{Z}$. Da har vi f.eks.

$$\varphi(x^2+1) = 0^2+1 = 1 \quad \text{og} \quad \varphi(5x^3-8x^2+4x+8) = 8.$$

∘

Eksempel 1.37. Hvis S er en delring af R, da vil afbildningen $i : S \to R$ givet ved $i(a) = a$ være en ringhomomorfi. Afbildningen i kaldes *inklusionsafbildningen*. ∘

Når vi kommer til kvotientringe, skal vi se endnu flere eksempler på ringhomomorfier, og det viser sig, at man kan benytte ringhomomorfier til at lave nye ringe. Til dette skal vi bruge følgende definition.

Definition 1.38. Lad $\varphi : R \to S$ være en ringhomomorfi. Da kaldes delmængden af R givet ved

$$\ker \varphi = \{a \in R \mid \varphi(a) = 0\}$$

for *kernen* af φ.

Lidt løst sagt kan man sige, at kernen af en ringhomomorfi er alle de elementer, der sendes til nul. Lad os se på nogle eksempler.

Eksempel 1.39. Hvis $\varphi : R \to S$ er lig nulafbildningen, da er $\ker \varphi = R$. ∘

Eksempel 1.40. Lad R være en ring og $\varphi : R[x] \to R$, $\varphi(f(x)) = f(0)$ som set i et tidligere eksempel. Hvordan ser kernen for φ ud? For et polynomium $f(x)$ er $f(0) = 0$ hvis og kun hvis konstantleddet for $f(x)$ er nul. Dermed er $\ker \varphi$ netop de polynomier, hvis konstantled er nul. Når vi introducerer idealer, kan vi give en mere elegant beskrivelse af $\ker \varphi$. ∘

Vi har følgende elegante kriterium for, hvornår en ringhomomorfi er injektiv.

Proposition 1.41. En ringhomomorfi $\varphi : R \to S$ er injektiv hvis og kun hvis $\ker \varphi = \{0\}$.

Bevis. Antag, at φ er injektiv. Da $\varphi(0) = 0$ (opgave), og φ er injektiv, må 0 være det eneste element, som sendes til 0, og dermed er $\ker \varphi = \{0\}$. Antag omvendt, at $\ker \varphi = \{0\}$. Lad $a, b \in R$ opfylde $\varphi(a) = \varphi(b)$. Da er $0 = \varphi(a) - \varphi(b) = \varphi(a - b)$, og dermed er $a - b \in \ker \varphi$. Men da $\ker \varphi = \{0\}$, må $a - b = 0$ og altså er $a = b$. Dette viser injektivitet. ■

En vigtig funktion for ringhomomorfier er at beskrive, hvordan ringe relaterer sig til hinanden. Hvor meget minder de om hinanden?

Definition 1.42. En bijektiv ringhomomorfi $\varphi : R \to S$ kaldes en *ringisomorfi*. Hvis der findes en ringhomomorfi mellem R og S, da siger vi, at R og S er *isomorfe*, og vi skriver $R \simeq S$.

Senere i forløbet skal vi se eksempler på isomorfier mellem ringe, og der er også givet et i opgaverne. Hvis to ringe er isomorfe, er de essentielt set ens. Elementerne i de to ringe hedder muligvis forskellige ting, men der er lige mange af dem, og regneoperationerne i de to ringe fungerer på præcist samme måde.

2 Idealer og kvotienter

Idealer

Vi skal nu introducere en type ring, der fortjener sit eget afsnit, nemlig kvotientringe. Vi skal først definere nogle særligt pæne delmængder af ringe, nemlig idealer.

Definition 2.1. Lad R være en ring. Et *ideal* I i R er en delmængde af R, som opfylder:

1. For alle $a, b \in I$ er $a + b \in I$.
2. For alle $a \in I$ og $r \in R$ er $ra \in I$.

Eksempel 2.2. For enhver ring R er R og $\{0\}$ begge idealer. ∘

Eksempel 2.3. Betragt $\mathbb{Z}$. Se på delmængden $n\mathbb{Z} = \{nm \mid m \in \mathbb{Z}\}$, altså delmængden af alle heltal i n-tabellen. F.eks. er $3\mathbb{Z} = \{..., -6, -3, 0, 3, 6, ...\}$. $n\mathbb{Z}$ er et ideal, hvilket vi nu viser:

1. Lad $a, b \in n\mathbb{Z}$. Da må $a = nm$ og $b = nk$ for heltal m og k. Da har vi $a + b = nm + nk = n(m + k)$, og $m + k$ er igen et heltal, så $a + b \in n\mathbb{Z}$.
2. Lad $a \in n\mathbb{Z}$ og $r \in \mathbb{Z}$. Da er $a = nm$ for et heltal m, og vi har $ra = (rm)n \in n\mathbb{Z}$ som ønsket.

∘

Lad R være en ring med elementer $a_1, ..., a_n$. Det er ikke svært at se, at mængden

$$a_1R + \cdots a_nR = \{a_1r_1 + \cdots + a_nr_n \mid r_1, ..., r_n \in R\}$$

er et ideal. Vi kalder dette idealet frembragt af $a_1, ..., a_n$. Det simple tilfælde med ét element a giver idealet

$$aR = \{ar \mid r \in R\}$$

altså alle multipla af a med elementer fra R. Sådanne idealer er vigtige nok til at få deres eget navn.

Definition 2.4. Et ideal på formen aR kaldes et *hovedideal*.

Eksempel 2.5. Betragt polynomiumsringen $\mathbb{R}[x]$ og idealet $(x^2+1)\mathbb{R}[x]$. Dette ideal består af alle polynomier $f(x)$, der kan skrives på formen $f(x) = g(x)(x^2+1)$. F.eks. er $x^3+x \in (x^2+1)\mathbb{R}[x]$, idet $x^3+x = x(x^2+1)$. ∘

Eksempel 2.6. Betragt idealet $4\mathbb{Z}+6\mathbb{Z}$ i $\mathbb{Z}$. Elementerne i dette ideal er alle heltal a, der kan skrives som $a = 4x+6y$, hvor $x, y \in \mathbb{Z}$. Vælger vi $x = 2$ og $y = -1$, kan vi se, at $2 = 4\cdot 2+6(-1) \in 4\mathbb{Z}+6\mathbb{Z}$. Dermed må $2\mathbb{Z} \subseteq 4\mathbb{Z}+6\mathbb{Z}$. Dog ser vi også, at alle tal på formen $4x+6y$ må være lige, og altså er $4\mathbb{Z}+6\mathbb{Z} \subseteq 2\mathbb{Z}$. Dermed er $4\mathbb{Z}+6\mathbb{Z} = 2\mathbb{Z}$, og $4\mathbb{Z}+6\mathbb{Z}$ er faktisk et hovedideal, selvom det ikke så sådan ud i begyndelsen. ∘

For legemer viser det sig, at idealer er ganske uinteressante.

Proposition 2.7. En ring R er et legeme hvis og kun hvis de eneste idealer i R er $\{0\}$ og R selv.

Bevis. Antag, at R er et legeme, og lad I være et ideal. Hvis $I = \{0\}$, er vi færdige, så antag $I \neq \{0\}$, og lad $a \in I$ være forskellig fra 0. Fordi R er et legeme, findes et $b \in R$ så $ab = 1$. Altså skal $1 \in I$. Men eftersom alle elementer $c \in R$ kan skrives som $c = c \cdot 1$, må $R \subseteq I$, hvilket viser $I = R$. Antag nu omvendt, at de eneste idealer i R er $\{0\}$ og R. Lad $a \neq 0$ i R. Da må idealet aR være lig R, eftersom $aR \neq \{0\}$. Dermed må $1 \in aR$, altså $1 = ab$ for et $b \in R$. Altså er a en enhed, og da a var et vilkårligt ikke-nul element, må R være et legeme som ønsket. ■

Vi har set en fremgangsmåde til at konstruere idealer, nemlig ved at vælge nogle elementer $a_1, ..., a_n$ og tage mængden bestående af alle elementer, der kan skrives på formen $a_1r_1+\cdots+a_nr_n$, hvor $r_1, ..., r_n$ er

elementer i hele ringen R. En anden mindre eksplicit fremgangsmåde er gennem homomorfier, som det næste resultat siger noget om.

Proposition 2.8. Lad $\varphi : R \to S$ være en ringhomomorfi. Da er $\ker \varphi$ et ideal i R.

Bevis. Lad $a, b \in \ker \varphi$. Da er $\varphi(a+b) = \varphi(a) + \varphi(b) = 0 + 0 = 0$, så $a + b \in \ker \varphi$. Lad nu $a \in \ker \varphi$ og $r \in R$. Da har vi $\varphi(ra) = \varphi(r)\varphi(a) = \varphi(r) \cdot 0 = 0$, så $ra \in \ker \varphi$. Dette viser, at $\ker \varphi$ er et ideal som ønsket. ■

Eksempel 2.9. Betragt ringhomomorfien $\varphi : R[x] \to R$ givet ved $\varphi(f(x)) = f(0)$ fra tidligere. Vi ved, at $\ker \varphi$ er netop de polynomier, hvis konstantled er nul, hvilket svarer til polynomierne, der kan skrives på formen $xg(x)$ for et $g(x) \in R[x]$. Altså er $\ker \varphi = xR[x]$. ∘

Inden vi bevæger os videre til kvotientringe, vil vi nævne nogle flere egenskaber for idealer. Nogle af beviserne skal I selv lave.

Proposition 2.10. Lad I og J være idealer i en ring R.

1. $I + J = \{a + b \mid a \in I, b \in J\}$ er et ideal i R.
2. $IJ = \{ab \mid a \in I, b \in J\}$ er et ideal i R.
3. $I \cap J$ er et ideal i R.
4. Der gælder $IJ \subseteq I \cap J \subseteq I + J$.

Bevis. Vi viser kun den første påstand. De andre overlades til læseren. Lad $c_1, c_2 \in I + J$. Da er $c_1 = a_1 + b_1$ og $c_2 = a_2 + b_2$ for $a_1, a_2 \in I$ og $b_1, b_2 \in J$. Dermed fås

$$c_1 + c_2 = (a_1 + a_2) + (b_1 + b_2) \in I + J.$$

Hvis $c = a + b \in I + J$, $r \in R$, da vil $rc = ra + rb \in I + J$, hvor vi bruger, at I og J begge er idealer. ■

Kvotientringe

Vi har nu de redskaber, der skal til for at konstruere kvotientringe. Kvotientringe kan virke meget mærkelige første gang, man støder på dem. Her kan det hjælpe at have mange eksempler i tankerne. Lad R være en ring og I et ideal. Lad for $a \in R$ mængden $a + I$ være givet ved

$$a + I = \{a + b \mid b \in I\}.$$

Man kan tænke på $a + I$ som, at alle elementer i I bliver "skubbet" med a. F.eks. er

$$2 + 3\mathbb{Z} = \{..., -4, -1, 2, 5, 8, ...\}.$$

Ideen med en kvotientring er, at elementerne er $a + I$, hvor a gennemløber alle elementer i R. Elementerne er altså forskydninger af et bestemt ideal og dermed mængder. Vi betegner mængden af disse forskudte idealer med R/I. Hvordan regner man med disse elementer? Sagt mere formelt, hvordan definerer vi addition og multiplikation i R/I? På følgende måde:

$$(a + I) + (b + I) = (a + b) + I,$$
$$(a + I)(b + I) = (ab) + I.$$

Her bør vi stoppe op og spørge, om disse regneoperationer overhovedet giver mening. Ser vi på eksemplet fra før, kan vi f.eks. se, at

$$2 + 3\mathbb{Z} = \{..., -4, -1, 2, 5, 8, ...\} = 5 + 3\mathbb{Z},$$

så elementet a foran I definerer ikke $a + I$ unikt. Dermed kunne man frygte, at selvom $a + I = b + I$, og $c + I = d + I$, da vil $(a + I) + (b + I) \neq (c + I) + (d + I)$. Heldigvis viser det sig, at dette ikke hænder, som vi skal se om lidt. Vi er dog nødt til at få lidt notation på plads.

Definition 2.11. Lad I være et ideal i en ring R, og lad $a, b \in R$. Vi siger, at $a \equiv b \pmod{I}$, hvis $a - b \in I$. I ord siger vi, at a og b er *kongruente modulo* I.

Denne definition burde være velkendt fra talteori (eller også bliver den det snart). Lad os se hvordan. Se på tilfældet $R = \mathbb{Z}$ og $I = n\mathbb{Z}$. Da er $a \equiv b \pmod{n\mathbb{Z}}$ hvis og kun hvis $a - b \in n\mathbb{Z}$. Dette er ækvivalent med $a - b = nk$ for et heltal k, altså at $n \mid a - b$. Dette er præcist den definition, som $a \equiv b \pmod{n}$ har i talteori. Vi kan altså konkludere, at

$$a \equiv b \pmod{n\mathbb{Z}} \quad \Leftrightarrow \quad a \equiv b \pmod{n}.$$

Lad os nu vende tilbage til kvotientringe. Hvis $a \equiv b \pmod{I}$, altså $a - b \in I$, da har vi

$$b + I = ((a - b) + b) + I = a + I,$$

og det er ikke svært at se, at den omvendte implikation også gælder. Dermed har vi

$$a + I = b + I \quad \Leftrightarrow \quad a \equiv b \pmod{I}.$$

Vi kan nu formulere følgende resultat.

Lemma 2.12. Lad $a_1, a_2, b_1, b_2 \in R$ og I et ideal i R. Hvis $a_1 \equiv a_2 \pmod{I}$ og $b_1 \equiv b_2 \pmod{I}$, da gælder

$$\begin{aligned} a_1 + b_1 &\equiv a_2 + b_2 \pmod{I}, \\ a_1 b_1 &\equiv a_2 b_2 \pmod{I}. \end{aligned}$$

Bevis. Per antagelse har vi $a_1 - a_2, b_1 - b_2 \in I$. Dermed har vi

$$(a_1 + b_1) - (a_2 + b_2) = (a_1 - a_2) + (b_1 - b_2) \in I,$$

thi I er et ideal, hvilket viser første ækvivalens. Ift. multiplikation fås

$$\begin{aligned} a_1 b_1 - a_2 b_2 &= a_1 b_1 - a_2 b_1 + a_2 b_1 - a_2 b_2 \\ &= (a_1 - a_2) b_1 + a_2 (b_1 - b_2) \in I \end{aligned}$$

fordi I er et ideal. Dette færdiggør beviset. ■

Lemmaet fortæller os, at de to regneregler

$$(a+I)+(b+I) = (a+b)+I,$$
$$(a+I)(b+I) = (ab)+I$$

er *veldefinerede*. Dette er blot matematisk jargon for, at man ikke går galt i byen ved at bruge dem. Vi kan nu langt om længe definere en kvotientring.

Definition 2.13. For en ring R og et ideal I i R defineres *kvotientringen* R/I til

$$R/I = \{a+I \mid a \in R\}$$

med regneoperationerne som defineret ovenover.

Alle reglerne for en ring (associativitet, neutralelementer osv.) følger direkte af, at R er en ring til at starte med. F.eks. har vi

$$\begin{aligned}(a+I)+((b+I)+(c+I)) &= (a+I)+((b+c)+I)\\ &= (a+(b+c))+I\\ &= ((a+b)+c))+I\\ &= ((a+b)+I)+(c+I)\\ &= ((a+I)+(b+I))+(c+I),\end{aligned}$$

og læseren er velkommen til at verificere resten af aksiomerne. Bemærk, at neutralelementet for addition er $0+I=I$, og for multiplikation er det $1+I$. Lad os se nogle eksempler.

Eksempel 2.14. Det vigtigste eksempel på en kvotientring er $\mathbb{Z}/n\mathbb{Z}$. Ser vi konkret på $\mathbb{Z}/7\mathbb{Z}$ kan vi f.eks. regne

$$(3+7\mathbb{Z})+(6+7\mathbb{Z}) = 9+7\mathbb{Z} = 2+7\mathbb{Z},$$
$$(3+7\mathbb{Z})(6+7\mathbb{Z}) = 18+7\mathbb{Z} = 4+7\mathbb{Z},$$

idet $9-2=7\in 7\mathbb{Z}$ og $18-4=14\in 7\mathbb{Z}$. ∘

Det kan være tidskrævende at skulle skrive $a + I$ hver gang, man skal skrive et element i R/I. Vi indfører derfor notationen $\overline{a} = a + I$. Denne notation virker naturligvis kun, når I er kendt ud fra konteksten, men dette er som regel tilfældet. I denne notation ville ovenstående eksempel sige

$$\overline{3} + \overline{6} = \overline{9} = \overline{2}$$

og

$$\overline{36} = \overline{18} = \overline{4},$$

som er en del hurtigere at skrive.

Eksempel 2.15. Betragt kvotientringen $\mathbb{Q}[x]/I$, hvor $I = (x^2 + x + 1)\mathbb{Q}[x]$. Hvordan regner man i denne ring? Idéen er, at alle led, hvor $x^2 + x + 1$ optræder, er 0 i $\mathbb{Q}[x]/I$. F.eks. har vi

$$\overline{4x^2 + 3x + 4} = \overline{x^2 + 1 + 3(x^2 + x + 1)} = \overline{x^2 + 1}.$$

Vi kan faktisk gøre dette udtryk endnu simplere. Idet $\overline{x^2 + x + 1} = \overline{0}$ (da $x^2 + x + 1 \in I$), må

$$\overline{x^2 + 1} = \overline{x^2 + x + 1 - x} = \overline{-x}.$$

Faktisk kan vi benytte relationen $\overline{x^2} = \overline{-(x + 1)}$ til at konkludere, at alle elementer i $\mathbb{Q}[x]/I$ er på formen $\overline{ax + b}$ for passende $a, b \in \mathbb{Q}$. Vi kan nemlig bruge $\overline{x^2} = \overline{-(x + 1)}$ til at reducere alle potenser af x større end 1. Lad os tage et simpelt eksempel. Vi har

$$\overline{x^3} = \overline{x(-(x + 1))} = \overline{-x^2 - x} = \overline{x + 1 - x} = \overline{1},$$

så overraskende er $\overline{x^3}$ faktisk lig den multiplikative identitet i $\mathbb{Q}[x]/I$!
∘

Eksempel 2.16. Hvad sker der, hvis vi betragter en kvotientring R/I, hvor $I = R$? Husk, at $\overline{a} = 0$ i R/I hvis og kun hvis $a \in I$. Så

hvis $I = R$, må R/I være lig nulringen. Hvad hvis $I = \{0\}$? Da er $R/I \simeq R$. Vi kan nemlig lave afbildningen

$$\varphi : R \to R/I, \quad \varphi(a) = \overline{a}.$$

Da $\overline{a} = 0$ hvis og kun hvis $a \in I = \{0\}$, er $\ker \varphi = \{0\}$, og dermed er φ injektiv per Proposition 1.41. Surjektiviteten er oplagt, og resultatet følger dermed. Grundet Proposition 2.7 kan vi konkludere, at man ikke kan konstruere nogle spændende kvotientringe ud fra legemer. ∘

Eksempel 2.17. Betragt $\mathbb{Z}[x]/I$, hvor $I = (x^3 + x)\mathbb{Z}[x] + 2\mathbb{Z}[x]$. Vi regner i denne ring på samme måde som i forrige eksempel, men nu er der endnu flere ting, der bliver nul. Vi har nemlig $\overline{p(x)} = \overline{0}$ hvis og kun hvis $p(x) = q(x)(x^3+x)+r(x)\cdot 2$ for polynomier $q(x), r(x) \in \mathbb{Z}[x]$. F.eks. er $\overline{6x^8 - 3x^3 + 5x - 10} = \overline{0}$, idet

$$6x^8 - 3x^3 + 5x - 10 = 1 \cdot (x^3 + x) + (3x^8 - 2x^3 + 2x - 5) \cdot 2.$$

∘

Eksempel 2.18. Dette eksempel er til dem, som har hørt om de komplekse tal $\mathbb{C}$. Det er tal på formen $a + bi$, hvor $a, b \in \mathbb{R}$ og i (kaldet den *imaginære enhed*) opfylder $i^2 = -1$. Med andre ord er i en rod i polynomiet $x^2 + 1$. Vi kan benytte kvotientringe til at konstruere $\mathbb{C}$, nemlig som kvotientringen $\mathbb{R}[x]/(x^2+1)\mathbb{R}[x]$. I ringen $\mathbb{R}[x]/(x^2+1)\mathbb{R}[x]$ har vi nemlig per konstruktion, at $\overline{x^2 + 1} = \overline{0}$, dvs. $\overline{x^2} = \overline{-1}$. Så intuitivt giver det mening, at $\mathbb{R}[x]/(x^2+1)\mathbb{R}[x]$ og $\mathbb{C}$ er så godt som det samme. Lad os vise det mere formelt ved at vise, at der findes en ringisomorfi $\varphi : \mathbb{C} \to \mathbb{R}[x]/(x^2 + 1)\mathbb{R}[x]$. Definér

$$\varphi(a + bi) = \overline{a + bx}.$$

Vi tjekker, at φ er en ringhomomorfi. Lad $a + bi, c + di \in \mathbb{C}$. Da fås

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

For at tjekke, at φ respekterer multiplikationen i de to ringe, regner vi først:

$$\begin{aligned}\varphi((a+bi)(c+di)) &= \varphi((ac-bd)+(ad+bc)i)\\ &= \overline{ac-bd+(ad+bc)x}.\end{aligned}$$

Nu beregner vi $\varphi(a+bi)\varphi(c+di)$:

$$\begin{aligned}\varphi(a+bi)\varphi(c+di) &= \overline{(a+bx)(c+dx)} = \overline{ac+adx+bcx+bdx^2}\\ &= \overline{ac-bd+(ad+bc)x}.\end{aligned}$$

Vi konkluderer, at φ er en ringhomomorfi. Vi mangler nu at vise, at φ er bijektiv, dvs. at den er injektiv og surjektiv. Antag, at $\varphi(a+bi)=0$, altså $a+bx \in (x^2+1)\mathbb{R}[x]$. Dermed findes $p(x)\in\mathbb{R}[x]$, så $a+bx=(x^2+1)g(x)$. $a+bx$ kan maksimalt have grad ét, mens $(x^2+1)g(x)$ har grad større end eller lig to, medmindre $g(x)=0$. Dermed må $g(x)=0$, og $a+bx=0$, hvilket medfører $a+bi=0$. Vi har vist, at $\ker\varphi=\{0\}$, og fra Proposition 1.41 kan vi konkludere, at φ er injektiv. Vi mangler nu at vise surjektiviteten. Lad $\overline{p(x)}\in \mathbb{R}[x]/(x^2+1)\mathbb{R}[x]$ være givet. Da x^2+1 har grad to, kan vi finde et førstegradspolynomium $q(x)=a+bx$, så $\overline{q(x)}=\overline{p(x)}$ ved at benytte $\overline{x^2}=\overline{-1}$ til at reducere alle potenser af x større end 1. Da har vi $\varphi(a+bi)=\overline{q(x)}=\overline{p(x)}$, og vi konkluderer, at φ er surjektiv. Vi har dermed vist, at

$$\mathbb{C}\simeq\mathbb{R}[x]/(x^2+1)\mathbb{R}[x].$$

∘

Der er en såkaldt *kanonisk* afbildning fra en ring R til enhver kvotientring R/I af R.

Proposition 2.19. Lad R være en ring og I et ideal. Vi har en ringhomomorfi

$$\pi : R\to R/I,\quad \pi(a)=\overline{a}$$

med $\ker\pi = I$.

Bevis. Vi overlader det som en øvelse til læseren at vise, at π er en ringhomomorfi. At $\ker \pi = I$ følger direkte af, at $\pi(a) = \overline{a} = \overline{0}$ hvis og kun hvis $a \in I$ per konstruktion. ■

Vi har nu konstrueret kvotientringe og set en række eksempler på dem. Vi vender nu kort tilbage til idealer. Der findes nemlig nogle interessante egenskaber for idealer, der kan oversættes til egenskaber for den tilhørende kvotientring.

Definition 2.20. Lad R være en ring.

- Et ideal $P \neq R$ kaldes et *primideal*, hvis det for alle $a, b \in R$ gælder, at $ab \in P$ medfører $a \in P$ eller $b \in P$.

- Et ideal $M \neq R$ kaldes et *maksimalideal* hvis det for ethvert ideal $I \neq R$ med $M \subseteq I$ gælder, at $I = M$.

Primidealer og maksimalidealer spiller ikke den store rolle i forløbet, men et ringteoriforløb ville ikke være fuldendt uden at nævne dem. Er der en sammenhæng mellem primtal og primidealer? Ja! Det viser sig, at et ideal i $\mathbb{Z}$ er et primideal hvis og kun hvis det er på formen $p\mathbb{Z}$, hvor p er et primtal. Lad os vise den ene implikation, nemlig at $p\mathbb{Z}$ er et primideal, hvis p er et primtal. Lad $a, b \in \mathbb{Z}$ med $ab \in p\mathbb{Z}$, altså $ab = pk$ for et heltal k. Dette er ækvivalent med, at $p \mid ab$. Men da ved vi (Euklids lemma), at $p \mid a$ eller $p \mid b$, hvilket er det samme som $a \in p\mathbb{Z}$ eller $b \in p\mathbb{Z}$ som ønsket.

Eksempel 2.21. Betragt $\mathbb{Z}[x]$. Vi hævder, at $x\mathbb{Z}[x]$ er et primideal, men ikke et maksimalideal. Vi ser, at $x\mathbb{Z}[x]$ er strengt indeholdt i idealet $x\mathbb{Z}[x] + 2\mathbb{Z}[x] \neq \mathbb{Z}[x]$ (1 er ikke indeholdt i venstresiden). Dermed er $x\mathbb{Z}[x]$ ikke et maksimalideal. For at se, at det er et primideal, antag $p(x)q(x) \in x\mathbb{Z}[x]$, dvs. $p(x)q(x) = xr(x)$ for et $r(x) \in \mathbb{Z}[x]$. Vi har da $p(0)q(0) = 0$, dvs. $p(0) = 0$ eller $q(0) = 0$. Men dette betyder, at enten $p(x) \in x\mathbb{Z}[x]$ eller $q(x) \in x\mathbb{Z}[x]$, da $x\mathbb{Z}[x]$ netop er polynomierne med konstantled lig nul. ○

Vi har følgende sammenhæng mellem et primideal og dens kvotientring.

Proposition 2.22. Lad R være en ring. P er et primideal hvis og kun hvis R/P er et integritetsområde.

Bevis. Husk, at et integritetsområde R er en ring, hvor $ab = 0$ medfører $a = 0$ eller $b = 0$. P er et primideal hvis og kun hvis $ab \in P$ medfører $a \in P$ eller $b \in P$. Dette er ækvivalent med $\overline{ab} = \overline{0}$ medfører $\overline{a} = \overline{0}$ eller $\overline{b} = \overline{0}$. Men dette er ækvivalent med, at R/P er et integritetsområde som ønsket. ■

Vi har et tilsvarende resultat for maksimalidealer. Dog har vi ikke redskaberne til at bevise det.

Proposition 2.23. Lad R være en ring. M er et maksimalideal hvis og kun hvis R/M er et legeme.

Bevis. Se f.eks. Proposition 12 på side 254 i [8]. Resultatet bygger på isomorfisætningerne for ringe, som vi ikke har tid til at komme ind på i forløbet her. ■

Korollar 2.24. Et maksimalideal er et primideal.

Bevis. Lad M være et maksimalideal i R. Da er R/M et legeme per ovenstående proposition. Et legeme er altid et integritetsområde (idet alle ikke-nul elementer er enheder, kan der ikke være nogle nuldivisorer), og altså er R/M et integritetsområde. Vi konkluderer, at M er et primideal. ■

Lad os tage et kig på kvotientringen $\mathbb{Z}/n\mathbb{Z}$. Vi kan bruge de ovenstående resultater til at sige noget om egenskaberne for denne ring alt efter n's værdi. Først skal vi dog vise et hjælperesultat, der bygger på talteori.

Lemma 2.25. Betragt kvotientringen $\mathbb{Z}/n\mathbb{Z}$.

1. $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$ er en enhed hvis og kun hvis a og n er indbyrdes primiske.

2. Alle ikke-nul $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$ er enten en enhed eller en nuldivisor.

Bevis. 1. $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$ er per definition en enhed, hvis der findes $\overline{b} \in \mathbb{Z}/n\mathbb{Z}$, så $\overline{ab} = \overline{1}$. Dette er ækvivalent med $ab - 1 \in n\mathbb{Z}$ dvs. $ab - 1 = kn$ for et $k \in \mathbb{Z}$. Dette omskrives til $ab - kn = 1$, og per Bezout's lemma er dette ækvivalent med, at a og n er indbyrdes primiske.

2. Lad $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$ være forskellig fra $\overline{0}$. Hvis $\overline{a}$ ikke er en enhed, kan a ikke være indbyrdes primisk med n, altså findes $k > 1$ som deler både a og n. Da har vi $\overline{n/k} \neq \overline{0}$, men $\overline{a(n/k)} = \overline{n(a/k)} = \overline{0}$, så $\overline{a}$ må være en nuldivisor.

■

Proposition 2.26. $\mathbb{Z}/n\mathbb{Z}$ er et legeme hvis og kun hvis n er et primtal.

Bevis. Alle ikke-nul elementer i $\mathbb{Z}/n\mathbb{Z}$ er givet ved $\overline{1}, \overline{2}, ..., \overline{n-1}$ (da $1, 2, ..., n-1$ er de mulige rester ved division med n). Per forrige lemma er $\mathbb{Z}/n\mathbb{Z}$ da et legeme netop hvis $1, 2, ..., n-1$ alle er indbyrdes primiske med n. Dette er ækvivalent med, at n er et primtal, hvilket viser det ønskede. ■

Bemærkning 2.27. Husk, at et endeligt integritetsområde er et legeme jævnfør Proposition 1.25. Dermed er $\mathbb{Z}/n\mathbb{Z}$ et integritetsområde hvis og kun hvis det er et legeme.

Eksempel 2.28. $\mathbb{Z}/15\mathbb{Z}$ er ikke et legeme, da 15 ikke er et primtal. Mere konkret kan dette også ses ved, at $\overline{3} \neq \overline{0}$ og $\overline{5} \neq \overline{0}$, men $\overline{3 \cdot 5} = \overline{15} = \overline{0}$, så $\overline{3}$ er en nuldivisor. ∘

Eksempel 2.29. $\mathbb{Z}/23\mathbb{Z}$ er et legeme, thi 23 er et primtal. Givet et element $\overline{a} \in \mathbb{Z}/23\mathbb{Z}$ forskelligt fra nul, hvordan bestemmes den inverse? Svaret er Euklids algoritme. Vi bestemmer heltal x og y, så

$ax + 23y = 1$. Lad os tage et eksempel med $a = 5$. Euklids algoritme giver

$$\begin{aligned} 23 &= 4 \cdot 5 + 3 \\ 5 &= 1 \cdot 3 + 2 \\ 3 &= 1 \cdot 2 + 1 \end{aligned}$$

Vi kan nu trævle algoritmen op baglæns. Vi har

$$\begin{aligned} 1 &= 3 - 2 = 3 - (5 - 3) = 2 \cdot 3 - 5 \\ &= 2 \cdot (23 - 4 \cdot 5) - 5 \\ &= -9 \cdot 5 + 2 \cdot 23. \end{aligned}$$

Vi kan nu aflæse tallet x til at være $x = -9$. Altså er $\overline{5}^{-1} = \overline{-9}$. Vi kan naturligvis også tjekke efter ved at regne

$$\overline{5 \cdot (-9)} = \overline{-45} = \overline{1},$$

så vores svar er korrekt. ∘

Eksempel 2.30. Lad os igen vise, at $x\mathbb{Z}[x]$ er et primideal, men ikke et maksimalideal i $\mathbb{Z}[x]$. Vi hævder, at

$$\mathbb{Z}[x]/x\mathbb{Z}[x] \simeq \mathbb{Z}.$$

Definér $\varphi : \mathbb{Z} \to \mathbb{Z}[x]/x\mathbb{Z}[x]$ ved sammensætningen $\varphi = \pi \circ i$, hvor $i : \mathbb{Z} \to \mathbb{Z}[x]$ er inklusionen $i(a) = a$. Konkret er $\varphi(a) = \overline{a}$. Vi ved, at φ er en ringhomomorfi, da den er en sammensætning af to ringhomomorfier (se opgaverne). Da x er et førstegradspolynomium, kan alle elementer i $\mathbb{Z}[x]/x\mathbb{Z}[x]$ skrives på formen $\overline{a}$ for et $a \in \mathbb{Z}$, hvilket viser, at φ er surjektiv. Antag nu, at $\varphi(a) = \overline{0}$. Da er $a \in x\mathbb{Z}[x]$. Men a er et heltal, så dette kan kun lade sig gøre, hvis $a = 0$. Altså er φ en isomorfi som ønsket. Idet $\mathbb{Z}$ er et integritetsområde, men ikke et legeme, følger det nu, at $x\mathbb{Z}[x]$ er et primideal, men ikke et maksimalideal. ∘

Eksempel 2.31. Med præcist samme fremgangsmåde som i forrige eksempel kan man vise

$$\mathbb{Q}[x]/x\mathbb{Q}[x] \simeq \mathbb{Q}.$$

Idet $\mathbb{Q}$ er et legeme, må $x\mathbb{Q}[x]$ være et maksimalideal og et primideal.
$\circ$

3 Euklidiske ringe

Vi har indtil videre set på to klasser af (kommutative) ringe, nemlig integritetsområder og legemer. Vi ved, at alle integritetsområder er legemer, så der er en klar rangorden indtil videre. I dette afsnit skal vi klassificere integritetsområder yderligere, og vi skal få flere redskaber til at regne på dem.

Euklidiske ringe

I kender euklidisk division fra talteori. Det viser sig, at man kan foretage euklidisk division i en hel klasse af ringe, der ikke så overraskende kaldes euklidiske.

Definition 3.1. Lad R være et integritetsområde. En afbildning $N : R \to \mathbb{N} \cup \{0\}$ kaldes en *norm*.

Definition 3.2. Et integritetsområde R kaldes *euklidisk*, såfremt der findes en norm N på R, sådan at for alle to elementer $a, b \in R$ med $b \neq 0$ kan man finde $q, r \in R$ med

$$a = qb + r \quad \text{hvor} \quad r = 0 \text{ eller } N(r) < N(b).$$

I så fald betegnes q som *kvotienten* og r som *resten*.

Definitionen siger blot, at det er muligt at foretage euklidisk division i R, hvor N dikterer størrelsesforholdet. Lad os starte med at få på plads, at legemer atter er ganske uinteressante fra et ringteoretisk perspektiv.

Proposition 3.3. Et legeme er euklidisk uanset valget af norm.

Bevis. Lad K være et legeme og $a, b \in K$ med $b \neq 0$. Da K er et legeme, findes b^{-1}, så $bb^{-1} = 1$. Dermed er $a = qb + 0$ for $q = ab^{-1}$. Dette virker, uanset hvad normen vælges til. ■

Lad os nu se på nogle interessante eksempler på euklidiske ringe. Husk, at den absolutte værdi $|a|$ af a er givet ved a, hvis $a \geq 0$ og $-a$, hvis $a < 0$. Med andre ord, $|a|$ fjerner blot et eventuelt fortegn på a.

Proposition 3.4. $\mathbb{Z}$ er euklidisk med normen $N(a) = |a|$.

Bevis. Lad a og b være heltal med $b \neq 0$. Vi deler op i to tilfælde:

- $b > 0$: Ved at tegne en tegning ser man, at ethvert heltal skal ligge i præcist ét af intervallerne $[nb, (n+1)b), n \in \mathbb{Z}$. Altså må $a \in [qb, (q+1)b)$ for et $q \in \mathbb{Z}$. Da må $r = a - qb < b = |b|$ som ønsket.
- $b < 0$: Fra forrige tilfælde ved vi, at der findes $q', r \in \mathbb{Z}$, så $a = q'(-b) + r$ og $r < -b = |b|$, idet $-b > 0$. Vælg nu $q = -q'$, da har vi fundet et par q, r med de ønskede egenskaber.

■

Eksempel 3.5. Betragt $a = 124$ og $b = 18$. Ved at prøve os frem, ser vi, at 18 deler 124 seks gange, og $124 - 6 \cdot 18 = 16 < 18$. Dermed kan vi skrive

$$124 = 6 \cdot 18 + 16.$$

Dette er dog ikke den eneste mulighed. Vi kunne også skrive

$$124 = 7 \cdot 18 - 2$$

og $|-2| < 18$, så dette er også en valid opskrivning. Resten og kvotienten bliver dog unik, hvis vi også kræver, at resten er positiv. Men dette er ikke et krav, som kommer af normen $N(a) = |a|$. ∘

Eksempel 3.6. Lad $a = 89$ og $b = -7$. Da har vi

$$89 = (-12)(-7) + 5,$$

og $5 < |-7|$. På samme vis som i forrige eksempel kunne vi i stedet have valgt

$$89 = (-13)(-7) - 2,$$

idet $|-2| < |-7|$. ∘

En anden interessant type euklidisk ring er polynomiumsringe, hvor koefficienterne udgør et legeme. Lad os først indføre følgende betegnelse for graden af et polynomium.

Definition 3.7. For et polynomium $p(x) \in R[x]$ (R er en ring) lader vi $\deg p(x)$ betegne graden af $p(x)$, altså den største potens af x, der indgår i $p(x)$. Vi definerer $\deg 0 = 0$.

Navnet deg kommer fra det engelske "degree". I skal ikke spekulere for længe over, hvorfor $\deg 0 = 0$. Dette er blot en konvention, som sikrer, at nulpolynomiet altid har den lavest mulige grad.

Eksempel 3.8. For $p(x) = x^2 + 1$ har vi $\deg p(x) = 2$. For $q(x) = 5$ har vi $\deg q(x) = 0$, idet $5 = 5x^0$, så den højeste potens af x i $q(x)$ er nul. ∘

Proposition 3.9. Lad K være et legeme. Da er $K[x]$ euklidisk med norm $N(p(x)) = \deg p(x)$.

Bevis. Lad $a(x), b(x) \in K[x]$ med $b(x) \neq 0$. Vi fører beviset med induktion på $n = \deg a(x)$.

- $n = 0$: I dette tilfælde er $a(x)$ konstant, $a(x) = a_0$. Vælg da $q(x) = 0$ og $r(x) = a_0$.
- $n > 0$: Antag påstanden gælder for alle polynomier af grad lavere end n. Lad $\deg b(x) = m$. Hvis $n < m$ ($a(x)$ har lavere grad end $b(x)$), vælg da $q(x) = 0$ og $r(x) = a(x)$. Antag nu $m \leq n$. Skriv

$$a(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$
$$b(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0.$$

  Betragt nu polynomiet $c(x) = a(x) - \frac{a_n}{b_m} x^{n-m} b(x)$. $\deg c(x) < n$ idet det første led af polynomiet, vi trækker fra, er lig $a_n x^n$,

og alle andre led har lavere grad end n. Vi kan da bruge induktionsantagelsen til at sige, at der findes polynomier $q_1(x)$ og $r(x)$, så

$$\begin{aligned} &c(x) = q_1(x)b(x) + r(x), \\ &\text{hvor} \quad r(x) = 0 \quad \text{eller} \quad \deg r(x) < \deg b(x). \end{aligned}$$

Lad nu $q(x) = q_1(x) + \frac{a_n}{b_m}x^{n-m}$, da har vi

$$\begin{aligned} &a(x) = q(x)b(x) + r(x), \\ &\text{hvor} \quad r(x) = 0 \quad \text{eller} \quad \deg r(x) < \deg b(x), \end{aligned}$$

som ønsket.

■

Beviset for ovenstående sætning er ikke-konstruktivt. Det fortæller os ikke, hvordan vi foretager euklidisk division med polynomier. Vi illustrerer en metode (kaldet *polynomiumsdivision*) med en række eksempler.

Eksempel 3.10. Lad os starte simpelt og lave division med $a(x) = x^2 - 1$ og $b(x) = x - 1$ (f.eks. i $\mathbb{Q}[x]$ eller $\mathbb{R}[x]$). I polynomiumsdivision starter man med at skrive problemet op således:

$$x - 1\big)\overline{\quad x^2 \qquad - 1}$$

Polynomiet til venstre er det, man deler med. Nu ser vi på leddet af højeste grad i polynomiet til venstre, nemlig x. Dette led deler x^2 x gange, så vi skriver x over stregen:

$$x - 1\big)\overline{\qquad\qquad\quad x \qquad\quad}$$
$$\qquad\quad x^2 \qquad - 1$$

Vi trækker nu $x(x-1) = x^2 - x$ fra polynomiet under stregen.

$$\begin{array}{r|l} & \quad x \\ \hline x-1 & \quad x^2 \quad -1 \\ & -x^2+x \\ \hline & \quad x-1 \end{array}$$

Da $x^2 - 1 - (x^2 - x) = x - 1$, skrives dette under stregen. Vi starter nu processen forfra og ser, hvor mange gange x går op i x, nemlig én gang. Derfor tilføjes $+1$ over stregen.

$$\begin{array}{r|l} & \quad x+1 \\ \hline x-1 & \quad x^2 \quad -1 \\ & -x^2+x \\ \hline & \quad x-1 \end{array}$$

Vi trækker nu $1 \cdot (x - 1)$ fra og får

$$\begin{array}{r|l} & \quad x+1 \\ \hline x-1 & \quad x^2 \quad -1 \\ & -x^2+x \\ \hline & \quad x-1 \\ & -x+1 \\ \hline & \quad 0 \end{array}$$

Da 0 har lavere grad end $x - 1$, stopper processen. Vi kan nu aflæse svaret på følgende måde: Det sidste polynomium, altså 0, er resten. Polynomiet over stregen, altså $x + 1$, er kvotienten. Dermed har vi

$$x^2 + 1 = (x + 1)(x - 1) + 0.$$

Dette er også nemt at verificere med håndkraft. ∘

Eksempel 3.11. Vi tager et lidt sværere eksempel. Lad $a(x) = x^3 + 4x + 2$ og $b(x) = x - 3$. Vi opskriver problemet:

$$\begin{array}{r|l} & \quad x^3 \qquad +4x \quad +2 \\ \hline x-3 & \end{array}$$

x deler x^3 x^2 gange, så vi noterer x^2 over stregen.

$$\begin{array}{rr|l} & & x^2 \\ \hline x-3) & x^3 & +4x\ +2 \end{array}$$

Vi trækker nu $x^2(x-3) = x^3 - 3x^2$ fra og får

$$\begin{array}{rr|l} & & x^2 \\ \hline x-3) & x^3 & +4x\ +2 \\ & -x^3+3x^2 & \\ \hline & 3x^2 & +4x \end{array}$$

Vi gentager og ser, at x deler $3x^2$ $3x$ gange. $3x$ tilføjes altså over stregen, og vi får

$$\begin{array}{rr|l} & & x^2+3x \\ \hline x-3) & x^3 & +4x\ +2 \\ & -x^3+3x^2 & \\ \hline & 3x^2 & +4x \end{array}$$

Vi trækker $3x(x-3) = 3x^2 - 9x$ fra og får

$$\begin{array}{rr|l} & & x^2+3x \\ \hline x-3) & x^3 & +4x\ +2 \\ & -x^3+3x^2 & \\ \hline & 3x^2 & +4x \\ & -3x^2 & +9x \\ \hline & & 13x\ +2 \end{array}$$

x deler $13x$ 13 gange, så 13 tilføjes over stregen.

$$\begin{array}{rr|l} & & x^2+3x+13 \\ \hline x-3) & x^3 & +4x\ +2 \\ & -x^3+3x^2 & \\ \hline & 3x^2 & +4x \\ & -3x^2 & +9x \\ \hline & & 13x\ +2 \end{array}$$

$13(x-3) = 13 - 39$ trækkes fra, og vi får den endelige opskrivning

$$\begin{array}{rl|l} & & x^2 + 3x + 13 \\ \hline x-3 & x^3 \quad + 4x + 2 & \\ & -x^3 + 3x^2 & \\ \hline & 3x^2 + 4x & \\ & -3x^2 + 9x & \\ \hline & 13x + 2 & \\ & -13x + 39 & \\ \hline & 41 & \end{array}$$

Da 41 har grad lavere end $x - 3$, stopper processen, og vi har det endelige svar

$$x^3 + 4x + 2 = (x^2 + 3x + 13)(x - 3) + 41.$$

$\circ$

Euklids algoritme

I kender Euklids algoritme fra talteori. Den præcist samme fremgangsmåde fungerer i alle euklidiske ringe. Vi beskriver først algoritmen formelt, og derefter giver vi nogle eksempler og nogle teoretiske konsekvenser af den. Først er vi dog nødt til at definere en største fælles divisor i generelle ringe.

Definition 3.12. Lad R være en ring og $a, b \in R$ med $b \neq 0$.

- b kaldes en *divisor* i a, hvis der findes et $c \in R$, så $a = bc$. Vi skriver i så fald $b \mid a$. I dette tilfælde kaldes a også et *multiplum* af b.

- En *største fælles divisor* for a og b er et element $d \in R$, så $d \mid a$ og $d \mid b$, og hvis $d' \in R$ er et andet element med $d' \mid a$ og $d' \mid b$, da vil $d' \mid d$. Vi betegner en største fælles divisor med $\gcd(a, b)$.

Bemærk, at vi skriver *en* største fælles divisor og ikke *den* største fælles divisor. En største fælles divisor behøver ikke at være unik. Faktisk er den det ret sjældent! Et simpelt eksempel er, at både 2 og -2 er største fælles divisorer for 4 og 6 i $\mathbb{Z}$. Mere generelt kan vi lave følgende observation.

Proposition 3.13. Hvis $a, b \in R$ og d er en største fælles divisor for a og b, da er ud for u en enhed også en største fælles divisor for a og b.

Bevis. Dette følger direkte af, at d deler et element $c \in R$ hvis og kun hvis ud gør det. Vi kan nemlig bemærke, at $c = de = ud(u^{-1}e)$. ■

Den mere interessante implikation går den anden vej. Den viser sig at gælde i et integritetsområde.

Proposition 3.14. Lad R være et integritetsområde og $a, b \in R$. Hvis både d og d' er største fælles divisorer, da er $d = ud'$ for en enhed u. Specielt er $dR = d'R$.

Bevis. Lad både d og d' være største fælles divisorer for a og b. Da d er en største fælles divisor, må $d' \mid d$, dvs. der findes $c \in R$ så $d = cd'$. Analogt findes $e \in R$, så $d' = ed$. Dermed har vi $d = ced$, altså $d(1 - ec) = 0$. $d \neq 0$ per antagelse, så idet R er et integritetsområde, må $ec = 1$, så e og c er enheder, hvilket viser første udsagn. Den anden del er vist i opgave 6.21. ■

Følgende lemma er vigtigt ift. at bevise korrekthed af Euklids algoritme.

Lemma 3.15. Lad $a, b \in R$, hvor R er euklidisk, og $b \neq 0$. Foretag euklidisk division og skriv

$$a = qb + r.$$

Da er $\gcd(a, b) = \gcd(b, r)$.

Bevis. Lad d være en største fælles divisor for a og b, d' en største fælles divisor for b og r. Da d' deler b og r, må d' også dele $qb+r=a$. Altså er d' en divisor i både a og b, så $d' \mid d$. Omvendt, da d deler a og b, må d dele b og $r=a-bq$, så $d \mid d'$. Altså er $d=ud'$ for en enhed u, hvilket viser det ønskede. ■

Lad os nu gennemgå Euklids algoritme. Lad R være euklidisk og $a, b \in R$. Foretag euklidisk division iterativt

$$\begin{aligned} a &= q_0 b + r_0 \\ b &= q_1 r_0 + r_1 \\ r_0 &= q_2 r_1 + r_2 \\ &\vdots \\ r_{n-2} &= q_n r_{n-1} + r_n \\ r_{n-1} &= q_{n+1} r_n, \end{aligned}$$

hvor r_n er den sidste ikke-nul rest. Følgende sætning fortæller os, at Euklids algoritme finder en største fælles divisor.

Sætning 3.16. Lad R være en euklidisk ring og $a, b \in R$ med $b \neq 0$. Lad $d = r_n$, altså den sidste ikke-nul rest i Euklids algoritme anvendt på a og b. Da er d en største fælles divisor i a og b.

Bevis. Vi skal først vise, at algoritmen terminerer (afslutter). Hvis N betegner den anvendte norm på R, da har vi $N(r_0) > N(r_1) > \cdots$, og da disse er ikke-negative heltal, kan denne sekvens ikke fortsætte for evigt, altså må algoritmen terminere. At r_n er en største fælles divisor følger ved at benytte Lemma 3.15 iterativt:

$$\gcd(a,b) = \gcd(b,r_0) = \gcd(r_1,r_0) = \cdots = \gcd(r_n,0) = r_n.$$

■

Eksempel 3.17. Betragt $a = 68$ og $b = 12$ i $\mathbb{Z}$. Vi benytter Euklids algoritme

$$\begin{aligned} 68 &= 5 \cdot 12 + 8 \\ 12 &= 1 \cdot 8 + 4 \\ 8 &= 2 \cdot 4, \end{aligned}$$

så $\gcd(68, 12) = 4$. ∘

Eksempel 3.18. Betragt $\mathbb{Q}[x]$ eller $\mathbb{R}[x]$. Vi ønsker at bestemme $\gcd(x^4 + 2x^3 - x + 2, x^2 + 4)$. Vi benytter Euklids algoritme

$$\begin{aligned} x^4 + 2x^3 - x + 2 &= (x^2 + 2x - 4)(x^2 + 4) + (-9x + 18) \\ x^2 + 4 &= \left(-\frac{1}{9}x - \frac{2}{9}\right)(-9x + 18) + 8 \\ -9x + 18 &= \left(-\frac{9}{8}x + \frac{9}{4}\right) \cdot 8 \end{aligned}$$

Dermed er $\gcd(x^4 + 2x^3 - x + 2, x^2 + 4) = 8$. Bemærk, at 8 er en enhed, så faktisk kunne vi lige så godt have skrevet $\gcd(x^4 + 2x^3 - x + 2, x^2 + 4) = 1$. ∘

Der er en elegant sammenhæng mellem idealer og største fælles divisorer.

Sætning 3.19. Lad R være en euklidisk ring, $a, b \in R$ med $b \neq 0$ og d en største fælles divisor fra Euklids algoritme for a og b. Da er

$$dR = aR + bR,$$

specielt eksisterer der x og y i R, så

$$ax + by = d.$$

Bevis. Vi skal vise to inklusioner. Vi viser først $aR + bR \subseteq dR$. Idet $d \mid a$ og $d \mid b$, må der findes $c, e \in R$ så $a = dc$ og $b = de$. Dette betyder, at $aR \subseteq dR$ og $bR \subseteq dR$, hvilket medfører $aR + bR \subseteq dR$.

Vi skal nu vise den anden inklusion, nemlig $dR \subseteq aR + bR$. Her er det tilstrækkeligt at vise, at $d \in aR + bR$, altså at $d = ax + by$ for passende $x, y \in R$. For at vise dette, benytter vi Euklids algoritme. Vi har $d = r_{n-2} - q_n r_{n-1}$ og ved at trævle algoritmen op bagfra (formelt ville man bruge induktion) kan vi skrive enhver rest ud fra de tidligere rester, indtil vi når a og b. Dermed har vi skrevet d på formen $d = ax + by$. ■

Bemærkning 3.20. For $\mathbb{Z}$ kaldes ovenstående resultat som regel *Bezout's identitet*.

Eksempel 3.21. Betragt $a = 136$ og $b = 48$. Vi ønsker at bestemme en største fælles divisor d og $x, y \in \mathbb{Z}$, så $d = 136x + 48y$. Først benytter vi Euklids algoritme

$$\begin{aligned} 136 &= 2 \cdot 48 + 40 \\ 48 &= 1 \cdot 40 + 8 \\ 40 &= 5 \cdot 8. \end{aligned}$$

Vi ser, at $d = \gcd(136, 48) = 8$. Vi trævler nu algoritmen op bagfra

$$8 = 48 - 40 = 48 - (136 - 2 \cdot 48) = 48 \cdot 3 + 136(-1),$$

så vi kan vælge $x = -1, y = 3$. ∘

Tilfældet, hvor $\gcd(a, b)$ er en enhed fortjener sit eget navn.

Definition 3.22. Lad R være en euklidisk ring. a og b kaldes *indbyrdes primiske*, hvis 1 er en største fælles divisor.

Vi ser, at a og b i en euklidisk ring R er indbyrdes primiske hvis og kun hvis $aR + bR = R$. Med andre ord, hvis alle $c \in R$ kan skrives på formen $c = ax + by$ for $x, y \in R$.

Eksempel 3.23. I et tidligere eksempel viste vi, at $x^4 + 2x^3 - x + 2$ og $x^2 + 4$ er indbyrdes primiske i $\mathbb{Q}[x]$ (eller $\mathbb{R}[x]$). ∘

4 Faktorisering og klassifikation af ringe

Irreducible elementer og primelementer

I heltallene $\mathbb{Z}$ ved vi, at vi kan faktorisere alle elementer (bortset fra 0, 1 og -1) til et produkt af primtal. Dette resultat kaldes *Aritmetikkens fundamentalsætning*. Vi skal nu se, hvordan dette koncept kan generaliseres til ringe (specifikt integritetsområder) udover $\mathbb{Z}$.

Definition 4.1. Lad R være et integritetsområde.

- Lad $a \in R$ være forskellig fra 0 ikke en enhed. a kaldes *irreducibel*, hvis $a = bc$ for $b, c \in R$ medfører, at b eller c er en enhed.
- Et ikke-nul element $p \in R$ kaldes et *primelement*, hvis pR er et primideal.
- Hvis $a = ub$ for $b \in R$ og en enhed u, da kaldes a og b *associerede*.

Skrevet ud i detaljer siger definitionen, at p er et primelement hvis det for $a, b \in R$ og $p \mid ab$ gælder, at $p \mid a$ eller $p \mid b$. Hvad er forskellen på et primelement og et irreducibelt element? I dette forløb er der faktisk ingen forskel, men det tager lidt arbejde at se hvorfor. Vi starter med at vise, at et primelement altid er irreducibelt. Den anden implikation kan findes i det supplerende materiale.

Proposition 4.2. Lad R være et integritetsområde. Hvis $p \in R$ er et primelement, er p irreducibelt.

Bevis. Lad p være et primelement, og antag $p = ab$ for $a, b \in R$. pR er et primideal, og $ab = p \in pR$, så $a \in pR$ eller $b \in pR$. Vi kan uden tab af generalitet antage $a \in pR$, så $a = pc$ for et $c \in R$. Da har vi altså $p = pcb$ og altså $p(1 - cb) = 0$. $p \neq 0$, så $1 - cb = 0$ (fordi R er

et integritetsområde) svarende til $cb = 1$. Altså er b en enhed, og vi konkluderer, at p er irreducibelt. ■

Definition 4.3. Et integritetsområde R kaldes *faktoriel* eller *UFD*, hvis ethvert element $a \in R$, som ikke er nul eller en enhed, kan skrives som et produkt af irreducible elementer p_i,

$$a = p_1 \cdots p_n,$$

og denne opskrivning er unik, hvis man ser bort fra associerede. Med andre ord, hvis $a = q_1 \cdots q_m$, hvor q_i er irreducible, da er $n = m$, og p_i er associeret til et q_j for alle i.

UFD kommer fra den engelske betegnelse for faktoriel, nemlig *unique factorization domain*.

Eksempel 4.4. For at illustrere hvordan definitionen skal forstås, se da på $\mathbb{Z}$ (vi skal se, at $\mathbb{Z}$ er faktoriel om ikke så længe). Bemærk, at a og b i $\mathbb{Z}$ er associerede netop hvis $a = \pm b$, altså er a og b kun afviger med et fortegn. Definitionen siger her konkret, at vi f.eks. ikke skelner mellem de to faktoriseringer

$$6 = 2 \cdot 3 \quad \text{og} \quad 6 = (-3)(-2).$$

Godt nok er de forskellige, men de er essentielt set det samme. Der kan kun være forskelle ift. rækkefølgen og eventuelle fortegn. ○

Eksempel 4.5. Lad os se et eksempel, hvor unikheden er mindre oplagt. Se på ringen $\mathbb{Q}[x]$ (som også viser sig at være faktoriel) og elementet $x^2 - 1$. Vi har tidligere set, at vi kan faktorisere $x^2 - 1$ som

$$x^2 - 1 = (x - 1)(x + 1).$$

Men vi kunne lige så godt have skrevet

$$x^2 - 1 = \frac{1}{2}(x - 1)2(x + 1).$$

Dog skelner vi ikke mellem disse to faktoriseringer, da $\frac{1}{2}(x - 1)$ er associeret til $x - 1$, eftersom $\frac{1}{2}$ er en enhed. Ligeledes er $2(x + 1)$ associeret til $x + 1$. ○

Indtil videre har vi ikke verificeret, at ovenstående eksempler faktisk er faktorielle ringe. Vi venter med dette til senere, hvor vi gennemgår de nødvendige teoretiske værktøjer. Vi skal nu se på nogle teknikker til at faktorisere, hvor vi fokuserer på polynomiumsringe $K[x]$, hvor K er et legeme.

Faktorisering i polynomiumsringe

Faktorisering af polynomier af grad to eller tre viser sig at være forholdsvist nemt. Her kommer polynomiumsdivision også i spil. For to polynomier $p(x)$ og $q(x)$ siger vi, at $q(x)$ er en *faktor* i $p(x)$, hvis $q(x) \mid p(x)$.

Proposition 4.6. Lad K være et legeme og $p(x) \in K[x]$. $x - a$ er en faktor i $p(x)$ hvis og kun hvis $p(a) = 0$.

Bevis. Antag, at $x-a$ er en faktor i $p(x)$ og skriv $p(x) = (x-a)q(x)$. Da har vi $p(a) = (a-a)q(a) = 0$ som ønsket. Antag omvendt, at $p(a) = 0$. Foretag polynomiumsdivision med $x - a$ og skriv

$$p(x) = q(x)(x - a) + r(x)$$

hvor $\deg r(x) < \deg(x - a) = 1$. Altså er $\deg r(x) = 0$, så $r(x)$ er et konstant polynomium. Da $p(a) = 0$, har vi $0 = r(a)$, men idet $r(x)$ er konstant, må $r(x) = 0$. Det følger, at $x-a$ er en faktor i $p(x)$ som ønsket. ■

Eksempel 4.7. Betragt $x^2 - 4 \in \mathbb{Q}[x]$. Det er ikke svært at se, at $x = 2$ og $x = -2$ er rødderne i polynomiet, så $x - 2$ og $x + 2$ er faktorer. ○

Ovenstående resultat gør os i stand til at afgøre, hvornår polynomier af grad to eller tre er irreducible.

Proposition 4.8. Lad K være et legeme. Et polynomium af grad to eller tre i $K[x]$ er irreducibelt hvis og kun hvis det ikke har en rod i K.

Bevis. Lad $p(x) \in K[x]$ have grad to eller tre. Hvis $p(x)$ har en rod a, har det faktoren $x - a$, så $p(x) = (x - a)q(x)$ for et polynomium $q(x)$ af grad et eller to. Dermed har vi en faktorisering af $p(x)$, så $p(x)$ er ej irreducibelt. Antag omvendt, at $p(x)$ ikke er irreducibelt, og skriv $p(x) = q(x)r(x)$, hvor $\deg q(x), \deg r(x) < \deg p(x)$. Da $p(x)$ har grad to eller tre, skal enten $q(x)$ eller $r(x)$ have grad et. Antag uden tab af generalitet, at $q(x)$ har grad et, altså $q(x) = ax + b$. Da er

$$p(x) = (ax + b)r(x) = \left(x + ba^{-1}\right) ar(x),$$

og vi ser, at $-ba^{-1}$ er en rod i $p(x)$. ■

Eksempel 4.9. $x^3 - 27 \in \mathbb{Q}[x]$ er ej irreducibelt, da $x = 3$ er en rod. $x^2 + 1$ er irreducibelt i $\mathbb{Q}[x]$, eftersom det ikke har nogle rødder i $\mathbb{Q}$. ∘

Eksempel 4.10. Resultatet kan kun bruges på polynomier af grad to eller tre. Se f.eks. på $x^4 + 2x^2 + 1 \in \mathbb{Q}[x]$. Dette polynomium kan skrives som

$$x^4 + 2x^2 + 1 = (x^2 + 1)^2,$$

så det er ikke irreducibelt, men det har ingen rødder i $\mathbb{Q}$. ∘

Hvordan bestemmer man rødder i polynomier? For andengradspolynomier kan man bruge den sædvanlige løsningsformel, men hvad hvis graden er højere end to? Som regel er det faktisk en god strategi at gætte sig frem. Følgende resultat fortæller en, hvordan man kan gætte smart i tilfældet med $p(x) \in \mathbb{Z}[x]$.

Proposition 4.11. Lad $p(x) = a_n x^n + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$. Hvis $a/b \in \mathbb{Q}$ er en forkortet brøk, og a/b er en rod i $p(x)$, da vil $a \mid a_0$ og $b \mid a_n$.

Bevis. Vi har per antagelse

$$0 = p(a/b) = a_n \left(\frac{a}{b}\right)^n + \cdots + a_1 \frac{a}{b} + a_0.$$

Ganges igennem med b^n fås

$$0 = a_n a^n + \cdots + a_1 a b^{n-1} + a_0 b^n.$$

Da a deler begge sider samt alle led på højre side udover $a_0 b^n$, må a også dele $a_0 b^n$. Da a og b er indbyrdes primiske, må a dele a_0 som ønsket. På præcist samme måde ses, at $b \mid a_n$. ■

Eksempel 4.12. Betragt $x^3 - x + 3 \in \mathbb{Q}[x]$. Per forrige proposition er ± 3 og ± 1 de eneste mulige rødder i $\mathbb{Q}$. Vi har $3^3 - 3 + 3 = 27$, $(-3)^3 - (-3) + 3 = -21$, $1^3 - 1 + 3 = 3$ og $(-1)^3 - 1 + 3 = 1$, så $x^3 - x + 3$ har ingen rationale rødder. Dermed er $x^3 - x + 3$ irreducibelt. ○

Eksempel 4.13. Lad os vise, at $\sqrt{2}$ er irrational, altså at $\sqrt{2}$ ikke er i $\mathbb{Q}$. Polynomiet $x^2 - 2$ har $\sqrt{2}$ som rod. Ifølge ovenstående proposition er de eneste mulige rationale rødder ± 1 og ± 2, men ingen af disse er rødder i polynomiet. Altså kan $\sqrt{2}$ ikke være rational. ○

Vi giver endnu et kriterium for irreducibilitet, som til tider kan kræve lidt kreativitet at benytte.

Proposition 4.14 (**Eisensteins irreducibilitetskriterium**)**.** Lad R være et integritetsområde og P et primideal. Betragt et polynomium $p(x) = x^n + a_{n-1}x^{n-1} + \cdots a_1 x + a_0$, hvor $a_{n-1}, ..., a_1, a_0 \in P$ og $a_0 \notin P^2$. Da er $p(x)$ irreducibel i $R[x]$.

Bevis. Antag for modstrid, at $p(x)$ er reducibel (ikke irreducibel), og skriv $p(x) = q(x)r(x)$, hvor $q(x)$ og $r(x)$ er ikke-konstante polynomier. Grundet vores antagelser om koefficienterne får vi $\overline{x^n} = \overline{p(x)} = \overline{q(x)r(x)}$ i $(R/P)[x]$. Fordi $(R/P)[x]$ er et integritetsområde (thi P er et primideal), må konstantleddene for $\overline{q(x)}$ og $\overline{r(x)}$ begge være nul (overvej). Men det svarer til, at konstantleddene i $q(x)$ og $r(x)$ begge er i P. Idet konstantleddet i $p(x)$, a_0, er produktet af konstantleddene i $q(x)$ og $r(x)$, må $a_0 \in P^2$, men dette er en modstrid. Vi konkluderer, at $p(x)$ er irreducibel. ■

Tilfældet med $\mathbb{Z}[x]$ er så typisk, at vi lader det få sit eget korollar.

Korollar 4.15. Lad p være et primtal i $\mathbb{Z}$ og $q(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \in \mathbb{Z}[x]$, hvor $p \mid a_i$ for alle $i = 0, 1, ..., n-1$, men $p^2 \nmid a_0$. Da er $q(x)$ irreducibel.

Bevis. Brug forrige proposition med $R = \mathbb{Z}[x]$, $P = pR$. ■

Eksempel 4.16. Betragt polynomiet $x^8 + 4x^7 + 12x^6 - 38x^3 + 10x^2 - 8x + 6 \in \mathbb{Z}[x]$. Vi ser, at primtallet 2 deler alle koefficienter bortset fra den første, og 2 deler kun 6 én gang. Dermed er polynomiet irreducibelt. ○

Følgende sætning giver en nyttig sammenhæng mellem polynomier med heltalskoefficienter og polynomier med rationale koefficienter.

Proposition 4.17. Lad $p(x) \in \mathbb{Z}[x]$. Hvis $p(x)$ er irreducibelt i $\mathbb{Q}[x]$, er $p(x)$ det også i $\mathbb{Z}[x]$.

Bevis. Dette resultat er et specialtilfælde af et resultat kaldet Gauss' lemma. Se [8] Proposition 5 på side 303 for et elementært bevis. ■

Eksempel 4.18. Har polynomiet $x^4 + 10x^2 + 15$ nogle rødder i $\mathbb{Q}$? Ved at bruge Eisensteins kriterium med $p = 5$ ser vi, at polynomiet er irreducibelt i $\mathbb{Z}[x]$. Dermed er det også irreducibelt i $\mathbb{Q}[x]$ og har altså ingen rationale rødder. ○

Lad os give nogle eksempler på, hvordan nogle konkrete polynomier kan faktoriseres med de værktøjer, vi har fået etableret.

Eksempel 4.19. Lad os faktorisere $p(x) = x^3 - x^2 - 3x + 3$ i $\mathbb{Q}[x]$. Det er oplagt at tjekke for rødder. Vi ser, at de mulige rationale rødder er ± 3 og ± 1. Prøver man sig frem, ser man, at 1 er en rod, mens de andre muligheder ikke er. Altså er $x - 1$ en faktor. Vi bruger nu polynomiumsdivision med $x - 1$ og får

$$
\begin{array}{rr|l}
 & x^2 & -3 \\
\hline
x-1 & x^3 - x^2 - 3x + 3 & \\
 & -x^3 + x^2 & \\
\hline
 & & -3x + 3 \\
 & & 3x - 3 \\
\hline
 & & 0
\end{array}
$$

hvoraf vi har $x^3 - x^2 - 3x + 3 = (x-1)(x^2-3)$. $x^2 - 3$ ses at være irreducibelt (f.eks. per Eisenstein), så dette er altså faktoriseringen af $x^3 - x^2 - 3x + 3$ i $\mathbb{Q}[x]$. ∘

Eksempel 4.20. Lad os faktorisere $x^4 - x^3 - 5x^2 - x - 6$ i $\mathbb{Q}[x]$. Igen kan vi tjekke for rationale rødder. Alle mulige rationale rødder er ± 1, ± 2, ± 3 og ± 6. Ved at sætte ind i polynomiet får vi, at -2 og 3 er rødder. Dermed er $x + 2$ og $x - 3$ faktorer. Vi laver nu polynomiumsdivision med $x + 2$:

$$
\begin{array}{rr|l}
 & x^3 - 3x^2 & +x - 3 \\
\hline
x+2 & x^4 - x^3 - 5x^2 & -x - 6 \\
 & -x^4 - 2x^3 & \\
\hline
 & -3x^3 - 5x^2 & \\
 & 3x^3 + 6x^2 & \\
\hline
 & x^2 & -x \\
 & -x^2 - 2x & \\
\hline
 & & -3x - 6 \\
 & & 3x + 6 \\
\hline
 & & 0
\end{array}
$$

Dermed har vi $x^3-3x^2+x-3 = (x+2)(x^3-3x^2+x-3)$. Da 3 også er en rod, må $x-3$ være en faktor i x^3-3x^2+x-3. Polynomiumsdivision giver $x^3-3x^2+x-3 = (x+2)(x-3)(x^2+1)$, hvilket er den endelige faktorisering. ∘

Eksempel 4.21. Lad os tage et andet eksempel end $\mathbb{Q}[x]$. Betragt $p(x) = x^3 + x^2 + x + 1$ i $\mathbb{Z}/2\mathbb{Z}[x]$. Bemærk, at $\mathbb{Z}/2\mathbb{Z}$ er et legeme, da

2 er et primtal. Der er kun to elementer i $\mathbb{Z}/2\mathbb{Z}$, nemlig 0 og 1 (vi undlader her at skrive stregen over elementerne). Vi har

$$p(1) = 1 + 1 + 1 + 1 = 0,$$

så 1 er en rod. Dermed er $x - 1 = x + 1$ en faktor. Polynomiumsdivision giver

$$\begin{array}{r|l} & x^2 \quad\quad + 1 \\ \hline x+1 & x^3 + x^2 + x + 1 \\ & \underline{-x^3 - x^2} \\ & \quad\quad\quad x + 1 \\ & \quad\quad\quad \underline{-x - 1} \\ & \quad\quad\quad\quad 0 \end{array}$$

så $x^3 + x^2 + x + 1 = (x + 1)(x^2 + 1)$. Bemærk, at 1 er en rod i $x^2 + 1$, så $x^2 + 1$ er reducibelt. Man kan udføre polynomiumsdivision igen og se, at $(x+1)^2 = x^2 + 1$, så den endelige faktorisering af $x^3 + x^2 + x + 1$ er

$$x^3 + x^2 + x + 1 = (x + 1)^3.$$

∘

Klassifikation af ringe og opsamling

Vi har set flere slags ringe undervejs i forløbet. Vi har primært arbejdet med integritetsområder, og vi har studeret nogle særlige typer integritetsområder, nemlig euklidiske ringe, faktorielle ringe og legemer. Vi har set, at legemer er euklidiske, men hvad mere kan vi sige? I dette afsnit skal vi sætte system i de forskellige ringe, vi har set.

Definition 4.22. En ring R kaldes et *hovedidealområde* eller et *PID*, hvis alle idealer er hovedidealer, altså hvis alle idealer er på formen aR for et element $a \in R$.

PID er den engelske betegnelse for et hovedidealområde. Det er en forkortelse for *principal ideal domain*. På engelsk hedder et hovedideal et *principal ideal*. Det viser sig, at hovedidealområder er den type ring, der forbinder euklidiske ringe med faktorielle ringe.

Proposition 4.23. En euklidisk ring er et hovedidealområde.

Bevis. Lad R være en euklidisk ring med norm N. Lad I være et ideal. Husk, at N antager værdier i $\mathbb{N} \cup \{0\}$, som er begrænset nedadtil. Vi kan dermed finde et element $a \in I$ med minimal norm, altså hvor $N(a)$ er mindst blandt alle $N(b)$, hvor $b \in I$. Lad nu $b \in I$ være arbitrær. Foretag euklidisk division

$$b = qa + r, \quad \text{hvor} \quad r = 0 \quad \text{eller} \quad N(r) < N(a).$$

Hvis $r \neq 0$, må $N(r) < N(a)$. Idet $r = b - aq \in I$, har vi altså bestemt et element i I med norm strengt mindre end normen af a. Dette er en modstrid, og dermed må $b = qa \in aR$. Da b var arbitrær, må $I \subseteq aR$. Den anden inklusion er triviel, og dermed er $I = aR$, hvilket beviser sætningen. ■

Sætning 4.24. Et hovedidealområde er faktoriel.

Bevis. Se det supplerende materiale. ■

Korollar 4.25 (**Aritmetikkens fundamentalsætning**)**.** $\mathbb{Z}$ er faktoriel.

Bevis. Vi har tidligere vist, at $\mathbb{Z}$ er euklidisk. Dermed er $\mathbb{Z}$ også et hovedidealområde per ovenstående proposition. Men da er $\mathbb{Z}$ faktoriel per sætningen. ■

Vi har nu et hierarki for ringe som følger:

$$\begin{aligned} R \text{ er et legeme } &\Rightarrow R \text{ er euklidisk } \Rightarrow R \text{ er PID} \\ &\Rightarrow R \text{ er UFD } \Rightarrow R \text{ er et integritetsområde} \end{aligned}$$

Ingen af disse implikationer er biimplikationer. Vi har allerede set, at $\mathbb{Z}$ er euklidisk, men ikke et legeme (andre eksempler er $K[x]$, hvor K er et legeme). De andre manglende biimplikationer kræver lidt mere forklaring.

Sætning 4.26. Lad R være en ring. Da er R faktoriel hvis og kun hvis $R[x]$ er det.

Bevis. Se Theorem 7 på side 304 i [8]. ■

Eksempel 4.27. Betragt $\mathbb{Z}[x]$. Per ovenstående sætning er $\mathbb{Z}[x]$ faktoriel, da $\mathbb{Z}$ er det. Dog er $\mathbb{Z}[x]$ ikke et hovedidealområde. Vi viser dette ved at vise, at $2\mathbb{Z}[x] + x\mathbb{Z}[x]$ ikke er et hovedideal. Antag for modstrid, at $2\mathbb{Z}[x] + x\mathbb{Z}[x] = p(x)\mathbb{Z}[x]$ for et $p(x) \in \mathbb{Z}[x]$. Da $2 \in p(x)\mathbb{Z}[x]$ må $p(x) \mid 2$. Men dette er kun muligt hvis $p(x) = \pm 1$ eller $p(x) = \pm 2$. Vi kan udelukke $p(x) = \pm 2$, da $p(x) \mid x$ også. Dermed skal $p(x) = \pm 1$, og altså skulle vi have $2\mathbb{Z}[x] + x\mathbb{Z}[x] = \mathbb{Z}[x]$. Specielt skal det være muligt at skrive $1 = 2a(x) + xb(x)$ for $a(x), b(x) \in \mathbb{Z}[x]$. Men indsætter vi 0, får vi $1 = 2a(0)$, altså $a(0) = 1/2$, hvilket er umuligt, da $a(x)$ har heltalskoefficienter. Vi har altså en modstrid, og vi konkluderer, at $2\mathbb{Z}[x] + R\mathbb{Z}[x]$ ikke er et hovedideal. ∘

Eksempel 4.28. Betragt ringen

$$\mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} \mid a, b \in \mathbb{Z}\},$$

hvor $\sqrt{-5}$ er et element, som opfylder $\sqrt{-5}^2 = -5$. Addition og multiplikation er defineret som sædvanligt. Vi lader læseren overbevise sig selv om, at $\mathbb{Z}[\sqrt{-5}]$ er lukket under addition og multiplikation, og at $\mathbb{Z}[\sqrt{-5}]$ opfylder alle reglerne for en ring. $\mathbb{Z}[\sqrt{-5}]$ ses at være et integritetsområde, men $\mathbb{Z}[\sqrt{-5}]$ er ikke UFD. Dette kan ses ved, at vi har

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}),$$

og disse faktoriseringer viser sig at være forskellige forstået på den måde, at de to elementer på den ene side er ikke associeret til nogle af dem på den anden side. ∘

Den vakse læser vil se, at vi mangler ét modeksempel. Vi mangler et hovedidealområde, der ikke er euklidisk. Man kan give et forholdsvist elementært eksempel, men det kræver meget arbejde at vise, at det er PID, men ikke euklidisk. Vi overlader det til læseren at undersøge denne sag.

5 Supplerende materiale

Noetherske ringe og et bevis for sætning 4.24

Da vi klassificerede ringe, kom vi ind på hovedidealområder, hvori alle idealer var frembragt af ét element. En mindre streng tilgang ville være at kræve, at alle idealer var frembragt af et *endeligt* antal elementer. Alle ringe, man støder på i praksis, har denne egenskab. Først skal vi dog definere, hvad det vil sige at frembringe et ideal i helt generel forstand.

Definition 5.1. Lad R være en ring og $A \subseteq R$ en delmængde. Da betegner

$$AR = \{a_1 r_1 + \cdots a_n r_n \mid a_i \in A, r_i \in R, n \in \mathbb{N}\}$$

idealet frembragt af A i R.

Definitionen siger, at AR består af alle elementer, der kan skrives som en endelig sum på formen $a_1 r_1 + \cdots a_n r_n$, hvor a'erne kommer fra A, og r'erne kommer fra R. Bemærk, at summen kan være arbitrært stor, men ikke uendelig.

Definition 5.2. En ring R kaldes *noethersk*, hvis enhver voksende kæde af idealer

$$I_1 \subseteq I_2 \subseteq ...$$

terminerer, dvs. der findes et $N \in \mathbb{N}$, så $I_n = I_N$ for alle $n \geq N$.

Noetherske ringe er opkaldt efter den tyske matematiker Emmy Noether. Følgende sætning karakteriserer noetherske ringe.

Sætning 5.3. Følgende udsagn er ækvivalente for en ring R:

(i) R er noethersk.

(ii) Alle idealer i R er endeligt frembragte. Med andre ord, ethvert ideal i R er på formen $a_1 R + \cdots a_n R$.

Bevis. Vi viser først, at (i) medfører (ii). Antag, at R er noethersk, og lad I være et ideal. Lav nu en kæde af idealer som følger: Vælg et vilkårligt $a_1 \in I$. Hvis $I_1 = a_1R = I$, er vi færdige. Ellers må der findes et $a_2 \in I$, men $a_2 \notin a_1R$. Sæt da $I_2 = a_1R + a_2R$. Gentag nu processen igen og igen, så vi får $I_1 \subseteq I_2 \subseteq ....$ Per antagelse stopper denne process efter et endeligt antal trin, n. Da vil $I = a_1R + \cdots a_nR$, og I er endeligt frembragt som ønsket. Vi viser nu, at (ii) medfører (i), så antag at alle idealer i R er endeligt frembragte. Lad

$$I_1 \subseteq I_2 \subseteq ...$$

være en voksende kæde af idealer. Vi lader læseren vise, at i dette tilfælde vil $\cup_{i=1}^{\infty} I_i$ være et ideal. Dermed er

$$\bigcup_{i=1}^{\infty} I_i = a_1R + \cdots a_nR$$

for nogle $a_i \in R$ per antagelse. Men alle disse a_i må ligge i et bestemt ideal I_N i kæden, eftersom vi kun har endeligt mange af dem, og kæden er voksende. Dermed har vi $I_n = I_N$ for alle $n \geq N$ som ønsket. ■

Korollar 5.4. Et hovedidealområde er en noethersk ring.

Bevis. Alle idealer i et hovedidealområde er endeligt frembragte, da de per definition er frembragt af ét element. ■

Eksempel 5.5. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{Q}[x]$ og $\mathbb{R}[x]$ er noetherske ringe. ○

Der findes også noetherske ringe, der ikke er hovedidealområder. Følgende sætning tillader os at konstruere nye noetherske ringe ud fra gamle.

Sætning 5.6. Antag, at R er en noethersk ring.

1. R/I er noethersk for ethvert ideal I i R.
2. $R[x]$ er noethersk. Dette resultat kaldes *Hilberts basissætning.*

Bevis. Beviset for første punkt overlades til læseren. Andet punkt er Theorem 7.5 i [9], hvor der findes et bevis. ■

Eksempel 5.7. Vi har tidligere set, at $\mathbb{Z}[x]$ ikke er PID. Dog er $\mathbb{Z}[x]$ noethersk per Hilberts basissætning. ○

Eksempel 5.8. Betragt ringen $S = R[x_1, x_2, ...]$ (R er ikke nulringen), altså en polynomiumsring i uendeligt mange variable. Denne ring er ej noethersk, da vi har en kæde

$$x_1 S \subseteq x_1 S + x_2 R \subseteq ...,$$

og denne kæde terminerer aldrig. ○

Vi vil nu bevise sætning 4.24, nemlig at et PID er UFD.

Lemma 5.9. I et hovedidealområde er et element irreducibelt hvis og kun hvis det er et primelement.

Bevis. I alle integritetsområder er et primelement irreducibelt, se proposition 4.2, så vi skal kun vise den anden implikation. Lad p være irreducibelt. Vi skal vise, at pR er et primideal. Vi gør dette ved at vise, at pR er et maksimalideal. Antag, at M er et ideal, der indeholder pR. $M = mR$ for et element m, da R er PID. Altså har vi $p = am$ for et $a \in R$. Men p er irreducibel per antagelse, så a eller m er en enhed. Hvis a er en enhed, er $pR = M$. Hvis m er en enhed, er $M = R$. Altså er det kun R og pR selv, der indeholder pR, hvilket viser, at pR er et maksimalideal. ■

Bevis for sætning 4.24. Lad R være et hovedidealområde, og lad $a \in R$ være et element forskelligt fra 0, som ikke er en enhed. Vi viser først, at a kan skrives som et produkt af irreducible elementer. Hvis a er irreducibel, er vi færdige, så antag, at det ikke er tilfældet. Da kan vi skrive $a = a_1 a_2$, hvor hverken a_1 eller a_2 er en enhed. Hvis a_1 og a_2 er irreducible er vi færdige, så antag, at a_1 ikke er irreducibel. Da kan vi skrive $a_1 = a_{11} a_{12}$, hvor hverken a_{11} eller a_{12} er irreducible. Vi kan

gentage denne procedure igen og igen og opnå en strengt voksende kæde af idealer

$$aR \subseteq a_1R \subseteq a_{11}R \subseteq ...$$

Da R er noethersk, må denne kæde terminere efter endeligt mange skridt, og dermed har vi vist, at vi kan skrive a som et produkt af irreducible elementer. Nu skal vi vise, at opskrivningen essentielt set er unik. Antag derfor, at vi har to opskrivninger af a som et produkt af irreducible elementer,

$$p_1 \cdots p_n = a = q_1 \cdots q_m.$$

Da p_1 deler venstresiden, deler p_1 også højresiden. p_1 er et primelement per ovenstående lemma, og dermed må p_1 dele mindst ét af q_j'erne. Ved om nødvendigt at ændre nummereringen, kan vi antage, at p_1 deler q_1. Per irreducibilitet, må p_1 og q_1 være associerede, $q_1 = up_1$. Da vi er i et integritetsområde, kan vi fjerne p_1 fra begge sider og få

$$p_2 \cdots p_n = uq_2 \cdots q_m.$$

Gentager vi proceduren igen og igen, er det klart, at der er lige mange faktorer på hver side, og at de er associeret til hinanden som ønsket. Beviset er dermed færdigt. ■

Læg mærke til, hvordan sidste del af beviset (omkring unikhed) minder meget om beviset for aritmetikkens fundamentalsætning i talteori.

Perspektiver og videre læsning

Kapitel 7-9 i [8] omhandler ringteori, og flere af beviserne er fra disse kapitler. En anden mere beregningsfokuseret tilgang kan findes i kapitel 9 af [10]. Når man har fået styr på basal ringteori, kan man gå i flere retninger. En af disse er kommutativ algebra, hvor [9] er en fantastisk introduktion, der også er velegnet til selvstudie. Hvis læseren er blevet interesseret i de gaussiske heltal, der blev introduceret i en række opgaver, er [11] en god kilde.

6 Opgaver

Opgaver til introduktion

- **Opgave 6.1:**
  Diskutér følgende begreber med din sidemakker. Kom med mindst ét eksempel til hvert begreb. Det er selvfølgelig tilladt at kigge i kompendiet undervejs.

  1) Kompositionsregel

  2) Ring

  3) Enhed

  4) Nuldivisor

  5) Integritetsområde

  6) Legeme

  7) Produktring

  8) Ringhomomorfi

- **Opgave 6.2:**
  Betragt ringen $\mathbb{R} \times \mathbb{Z}[x]$.

  1) Find en nuldivisor i denne ring.

  2) Find en enhed (udover $(1, 1)$) i denne ring.

- **Opgave 6.3:**
  Lad R være en ring og u en enhed. Bevis, at $-u$ er en enhed (hvor $-u = (-1) \cdot u$).

- **Opgave 6.4:**
  Lad R være en ring, som ikke er nulringen. Overvej, om $R \times R$ kan være et integritetsområde.

•• **Opgave 6.5:**
  Lad R være en ring og betragt $(a, b) \in R \times R$. Hvad skal gælde om a og b for, at (a, b) er en enhed? Hvad skal gælde om a og b for, at (a, b) er en nuldivisor?

•• **Opgave 6.6:**
Betragt mængden $R = \{f : [0,1] \to \mathbb{R} \mid f \text{ er en funktion}\}$, altså mængden af alle funktioner fra $[0,1]$ til $\mathbb{R}$. For $f, g \in R$ definerer vi $f+g$ og fg til at være funktionerne givet ved

$$(f+g)(x) = f(x) + g(x), \quad (fg)(x) = f(x)g(x).$$

1) Overbevis dig selv om, at R er en ring med de givne regneoperationer. Lad $f(x) = \cos x$ og $g(x) = 2x + e^x$. Bestem funktionerne $f+g$ og fg.

2) Lad $f \in R$ med $f(x) \neq 0$ for alle $x \in [0,1]$. Vis, at f er en enhed. Hvad er f^{-1}?

3) Antag, at $f \in \mathbb{R}$ opfylder $f(x_0) = 0$ for et $x_0 \in [0,1]$. Vis, at f er en nuldivisor.

I resten af opgaven betragter vi for et fast $a \in [0,1]$ afbildingen

$$\varphi_a : R \to \mathbb{R}, \quad \varphi_a(f) = f(a).$$

4) Lad i denne delopgave $a = 0$. Udregn $\varphi_a(f)$ for hhv. $f(x) = e^x$ og $f(x) = x^2 + 4x - 3$.

5) Bevis, at φ_a er en ringhomomorfi (for alle valg af a).

6) Forklar med dine egne ord, hvad $\ker \varphi_a$ er.

•• **Opgave 6.7:**
Lad $\varphi : R \to S$ være en ringhomomorfi.

1) Bevis, at $\varphi(0) = 0$. [Vink: $0 = 0 + 0$]

2) Bevis, at $\varphi(-a) = -\varphi(a)$ for alle $a \in R$.

•• **Opgave 6.8:**
Lad R, S og T være ringe. Antag, at $\varphi : R \to S$ og $\psi : S \to T$ er ringhomomorfier. Bevis, at $\psi \circ \varphi : R \to T$ er en ringhomomorfi.

•• **Opgave 6.9:**
Lad R være et integritetsområde.

1) Bevis, at $a^2 = 1$ medfører $a = 1$ eller $a = -1$. [Vink: Brug omskrivningen $a^2 - 1 = (a+1)(a-1)$]

2) Giv et eksempel på en ring R, så $a^2 = 1$, men $a \neq 1, -1$. [Vink: produktringe]

•• **Opgave 6.10:**
Denne opgave er til dem, som kender til differentialregning. Betragt polynomiumsringen $\mathbb{R}[x]$ og afbildingen $\varphi : \mathbb{R}[x] \to \mathbb{R}[x]$ givet ved

$$\varphi(f(x)) = f'(x).$$

Er φ en ringhomomorfi? Hvis ikke, opfylder den så en af egenskaberne?

••• **Opgave 6.11:**
I denne opgave introducerer vi en ring $\mathbb{Z}[i]$ kaldet de *gaussiske heltal*. De er givet ved

$$\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\},$$

hvor $i^2 = -1$. Et eksempel på et element kunne være $2 + 3i$. Regneoperationerne er præcist som med almindelige tal, man skal blot huske, at $i^2 = -1$.

1) Lad $\alpha = 2 + 3i$ og $\beta = -1 + 4i$. Vis ved udregning, at vi har

$$\alpha + \beta = 1 + 7i \quad \text{og} \quad \alpha\beta = -14 + 5i.$$

Hvis $\alpha = a + bi$, definerer vi *konjugationen* af α som $\alpha^* = a - bi$.

2) Lad $\alpha = a + bi$ og $\beta = c + di$ være elementer i $\mathbb{Z}[i]$. Bevis, at $(\alpha\beta)^* = \alpha^*\beta^*$. [Vink: regn venstre- og højresiden ud for sig og se, at de er ens]

3) Definér normen $N : \mathbb{Z}[i] \to \mathbb{N} \cup \{0\}$ ved $N(a + bi) = a^2 + b^2$. Vis, at $N(\alpha) = \alpha\alpha^*$. Vis derefter, at N er multiplikativ, altså at $N(\alpha\beta) = N(\alpha)N(\beta)$. Bemærk, at vi ikke behøver, at a og b er heltal

for at vise dette. Dette vil vi udnytte senere. [Vink: til sidste del er det smart at bruge forrige delopgave]

4) Vi skal nu undersøge enhederne i $\mathbb{Z}[i]$. Bevis, at $\alpha \in \mathbb{Z}[i]$ er en enhed hvis og kun hvis $N(\alpha) = 1$. [Vink: Hvis $a, b \in \mathbb{N}$ med $ab = 1$, da må $a = b = 1$]

5) Bevis, at enhederne i $\mathbb{Z}[i]$ er ± 1 og $\pm i$.

••• **Opgave 6.12:**
Lad R være en ring.

1) Bevis, at delmængden $S = \{(a, a) \mid a \in R\}$ er en delring af $R \times R$.

2) Bevis, at $S \simeq R$.

••• **Opgave 6.13:**
Lad R være en ring. Et element a kaldes *nilpotent*, hvis der findes en potens k, så $a^k = 0$. Antag, at a er nilpotent.

1) Bevis, at $a = 0$ eller at a er en nuldivisor.

2) Lad $b \in R$ være vilkårlig. Vis, at ab er nilpotent.

3) Bevis, at $1 + a$ er en enhed. [Vink: Brug $1 - x^k = (1 - x)(1 + x + x^2 + \cdots + x^{k-1})$]

Opgaver til idealer og kvotienter

• **Opgave 6.14:**
Diskutér følgende begreber med din sidemakker. Kom med mindst ét eksempel til hvert begreb. Det er selvfølgelig tilladt at kigge i kompendiet undervejs.

1) Ideal (herunder hovedideal)

2) Kvotientring

3) Primideal

• **Opgave 6.15:**
Hvilke af følgende kvotientringe er legemer?

1) $\mathbb{Z}/4\mathbb{Z}$

2) $\mathbb{Z}/11\mathbb{Z}$

3) $\mathbb{Z}/19\mathbb{Z}$

4) $\mathbb{Z}/10\mathbb{Z}$

- **Opgave 6.16:**
  Er $x^2\mathbb{Q}[x]$ et primideal i $\mathbb{Q}[x]$? Er $\mathbb{Q}[x]/x^2\mathbb{Q}[x]$ et integritetsområde? Et legeme?

- **Opgave 6.17:**
  Betragt ringen $\mathbb{Z}/10\mathbb{Z}$. Elementerne i denne ring er $\overline{0}, \overline{1}, ..., \overline{9}$. For hvert element herunder, bestem hvilken af $\overline{0}, \overline{1}, ..., \overline{9}$ det er lig.

  1) $\overline{20}$

  2) $\overline{31}$

  3) $\overline{17}$

  4) $\overline{19}$

  5) $\overline{1005}$

- **Opgave 6.18:**
  Betragt kvotientringen $\mathbb{Z}/12\mathbb{Z}$.

  1) Find alle enheder i denne ring.

  2) Find alle nuldivisorer i denne ring.

•• **Opgave 6.19:**
  Betragt kvotientringen $\mathbb{Z}/17\mathbb{Z}$. Denne ring er et legeme, thi 17 er et primtal, og dermed har alle $\overline{a} \neq \overline{0}$ en (multiplikativ) invers.

  1) Bestem den inverse til $\overline{3}$.

  2) Bestem den inverse til $\overline{16}$.

•• **Opgave 6.20:**
  Betragt kvotientringen $\mathbb{Q}[x]/I$ med $I = (x^3-x-2)\mathbb{Q}[x]$. Vis følgende:

  1) $\overline{x^3} = \overline{x+2}$

  2) $\overline{x^7} = \overline{4x^2+5x+2}$

•• **Opgave 6.21:**
Lad R være et integritetsområde og $a, b \in R$. Vis, at $aR = bR$ hvis og kun hvis $a = ub$ for en enhed u.

•• **Opgave 6.22:**
Hvad er idealet i $3\mathbb{Z} + 5\mathbb{Z}$ i $\mathbb{Z}$ lig?

•• **Opgave 6.23:**
Lad R være en ring.

1) Bevis, at R er et integritetsområde hvis og kun hvis $\{0\}$ er et primideal.

2) Bevis, at R er et legeme hvis og kun hvis $\{0\}$ er et maksimalideal.

•• **Opgave 6.24:**
Lad I være et ideal i en ring R. Bevis, at $I = R$ hvis og kun hvis I indeholder en enhed.

••• **Opgave 6.25:**
Lad $\varphi : K \to R$ være en ringhomomorfi, hvor K er et legeme. Bevis, at φ enten er injektiv eller er lig nulafbildningen. [Vink: Benyt Proposition 1.41, 2.7 og 2.8]

••• **Opgave 6.26:**
Lad R være en ring og I et ideal. Delmængden

$$\sqrt{I} = \{a \in R \mid a^n \in I \text{ for et } n \in \mathbb{N}\}$$

kaldes *radikalet* af I.

1) Vis, at $\sqrt{I}$ er et ideal med $I \subseteq \sqrt{I}$.

2) I kaldes *radikalt*, hvis $I = \sqrt{I}$. Bevis, at ethvert primideal P er radikalt.

3) Giv et eksempel på et radikalt og et ikke-radikalt ideal. Kan du finde et radikalt ideal, der ikke er et primideal?

Opgaver til euklidiske ringe

- **Opgave 6.27:**
  Diskutér følgende begreber med din sidemakker. Kom med mindst ét eksempel til hvert begreb. Det er selvfølgelig tilladt at kigge i kompendiet undervejs.

  1) Norm

  2) Euklidisk ring

  3) Divisor

  4) Største fælles divisor

  5) Indbyrdes primiske

- **Opgave 6.28:**
  Brug Euklids algoritme til at bestemme en største fælles divisor d af 245 og 65 i $\mathbb{Z}$. Bestem herefter x og y, så $d = 245x + 65y$.

- **Opgave 6.29:**
  Foretag polynomiumsdivision med $x^2 + 2x + 5$ og $x - 3$.

- **Opgave 6.30:**
  Vis, at $\mathbb{Z} = 17\mathbb{Z} + 90\mathbb{Z}$.

- **Opgave 6.31:**
  Bestem et $d \in \mathbb{Z}$, så $d\mathbb{Z} = 77\mathbb{Z} + 126\mathbb{Z}$.

•• **Opgave 6.32:**
  Bestem $\gcd(x^3 + 2x^2 - 3x + 4, x - 1)$ i $\mathbb{Q}[x]$ med Euklids algoritme.

•• **Opgave 6.33:**
  Bestem $\gcd(x^4 - 2x^2 - 2x - 4, x^2 - 4)$ i $\mathbb{Q}[x]$ med Euklids algoritme.

••• **Opgave 6.34:**
  Denne opgave bygger videre på opgave 6.11 omkring de gaussiske heltal $\mathbb{Z}[i]$.

  1) Lav opgave 6.11, hvis du ikke allerede har lavet den.

I denne opgave vil vi vise, at $\mathbb{Z}[i]$ er euklidisk med normen N. Vi vil faktisk vise, at for alle $\alpha, \beta \in \mathbb{Z}[i]$ findes $\gamma, \rho \in \mathbb{Z}[i]$ med $\alpha = \gamma\beta + \rho$, hvor $N(\rho) \leq (1/2)N(\beta)$.

2) Lad $\alpha, \beta \in \mathbb{Z}[i]$ med $\beta \neq 0$. Vis, at man kan skrive

$$\frac{\alpha}{\beta} = \frac{x}{N(\beta)} + \frac{y}{N(\beta)}i$$

for passende $x, y \in \mathbb{Z}$. [Vink: gang tæller og nævner med β^*]

3) Per euklidisk division (i $\mathbb{Z}$) kan vi skrive

$$x = N(\beta)q_1 + r_1, \quad y = N(\beta)q_2 + r_2,$$

hvor $|r_1|, |r_2| \leq (1/2)N(\beta)$. Forklar hvorfor. [Vink: Husk, at vi har to valg for kvotienten og resten, når vi laver euklidisk division $a = qb+r$. Hvis det ene valg af r ikke opfylder $|r| \leq (1/2)b$, hvad så med den anden?]

4) Vis ved udregning, at

$$\frac{\alpha}{\beta} = q_1 + q_2 i + \frac{r_1 + r_2 i}{N(\beta)}.$$

5) Vi definerer nu $\gamma = q_1 + q_2 i$. Vis, at

$$\alpha - \beta\gamma = \frac{r_1 + r_2 i}{\beta^*}.$$

[Vink: Husk fra tidligere, at $N(\beta) = \beta\beta^*$]

6) Lad nu $\rho = \alpha - \beta\gamma$. Vi mangler at vise $N(\rho) \leq (1/2)N(\beta)$. Ved at tage normer har vi

$$N(\rho) = \frac{r_1^2 + r_2^2}{N(\beta)}.$$

Vis herfra, at $N(\rho) \leq (1/2)N(\beta)$.

Opgaver til faktorisering og klassifikation af ringe

- **Opgave 6.35:**
  Diskutér følgende begreber med din sidemakker. Kom med mindst ét eksempel til hvert begreb. Det er selvfølgelig tilladt at kigge i kompendiet undervejs.

  1) Irreducibel

  2) Primelement

  3) Associerede elementer

  4) Faktoriel ring/UFD

  5) Hovedidealområde/PID

- **Opgave 6.36:**
  Giv et eksempel på et irreducibelt polynomium i $\mathbb{Q}[x]$, der ikke er irreducibelt i $\mathbb{R}[x]$.

- **Opgave 6.37:**
  Vis, at følgende polynomier er irreducible i $\mathbb{Z}[x]$.

  1) $x^2 + 4x - 2$

  2) $x^6 + 18x^5 + 9x^6 + 12$

  3) $x^3 - 33x^2 + 77x - 44$

- •• **Opgave 6.38:**
  Faktorisér følgende polynomier over $\mathbb{Q}[x]$.

  1) $x^3 + 2x^2 - 3x$

  2) $x^3 + 6x^2 + 12x + 8$

- •• **Opgave 6.39:**
  Faktorisér følgende polynomier over $\mathbb{Z}/5\mathbb{Z}$.

  1) $x^2 + 1$

  2) $2x^3 + 2x + 4$

  3) $x^4 + 4$

•• **Opgave 6.40:**
Find alle andengradspolynomier over $\mathbb{Z}/2\mathbb{Z}$. Hvor mange er der? Hvilke er irreducible?

•• **Opgave 6.41:**
Bevis, at $\sqrt{3}$ er irrational. [Vink: se eksempel 4.13]

••• **Opgave 6.42:**
Lad K være et legeme. Bevis, at der findes uendeligt mange irreducible elementer i $K[x]$. [Vink: Hvis $p_1(x), ..., p_n(x)$ er irreducible, betragt da $p_1(x) \cdots p_n(x) + 1$. Da $K[x]$ er UFD, kan vi faktorisere i irreducible elementer og få $p_1(x) \cdots p_n(x) + 1 = q_1(x) \cdots q_m(x)$. Kan $q_i(x) = p_j(x)$ for nogle i, j?]

••• **Opgave 6.43:**
I denne opgave skal vi yderligere studere de gaussiske heltal fra opgave 6.11 og 6.34.

1) Lav opgave 6.11, såfremt du ikke allerede har lavet den.

Opgave 6.34 viser, at $\mathbb{Z}[i]$ er euklidisk og dermed UFD. I denne opgave vil vi blive klogere på irreducible elementer og faktorisering i $\mathbb{Z}[i]$. Det er ikke nødvendigt at lave opgave 6.34 for at løse de kommende opgaver.

2) Lad $\alpha \in \mathbb{Z}[i]$. Bevis, at hvis $N(\alpha)$ er et primtal i $\mathbb{Z}$, da er α irreducibel. [Vink: Husk, at N er multiplikativ]

3) Brug resultatet fra forrige delopgave til at vise, at $1 + i$, $1 - 2i$, $2 + 3i$ og $1 + 4i$ er irreducible i $\mathbb{Z}[i]$.

4) Man kan bevise, at de irreducible elementer i $\mathbb{Z}[i]$ er alle de elementer, der er associeret til én af følgende:

(i) $1 + i$

(ii) Primtal $p \in \mathbb{Z}$ med $p \equiv 3 \pmod 4$

(iii) $a + bi$ eller $a - bi$ med $p = a^2 + b^2$ et primtal med $p \equiv 1 \pmod 4$

Afgør, hvilke af følgende elementer, der er irreducible.

- $1 - i$
- $3 + 7i$
- 19
- $-5 - 4i$
- $-11i$
- $2 + 5i$
- $-1 - 2i$
- $13i$

5) Faktorisér $4 + 6i$ i et produkt af irreducible elementer i $\mathbb{Z}[i]$.

7 Projekt: En spøjs talring

I dette projekt kigger vi på ringen $\mathbb{Z}[\sqrt{-2}] = \{a + b\sqrt{-2} \mid a, b \in \mathbb{Z}\}$. Denne ring består af elementer på formen $a+b\sqrt{-2}$, hvor $\sqrt{-2}$ er et tal, som opfylder $\sqrt{-2}^2 = -2$. Man kan vise, at sådan et tal findes, og I skal blot tage det for givet. Addition og multiplikation i ringen er defineret, som man forventer. F.eks. er

$$\begin{aligned}(2 - \sqrt{-2}) + (-3 + 4\sqrt{-2}) &= -1 + 3\sqrt{-2} \\ (2 - \sqrt{-2})(-3 + 4\sqrt{-2}) &= 2 \cdot (-3) + 2 \cdot 4\sqrt{-2} \\ &\quad - \sqrt{-2}(-3) - 4\sqrt{-2}^2 \\ &= -6 + 8\sqrt{-2} + 3\sqrt{-2} - 4(-2) \\ &= 2 + 11\sqrt{-2}\end{aligned}$$

Vi vil i det følgende betegne elementer i $\mathbb{Z}[\sqrt{-2}]$ med græske bogstaver såsom α, β, γ osv. og almindelige heltal med latinske bogstaver.

Opgave 1: Lad $\alpha = 2+3\sqrt{-2}$ og $\beta = -4+2\sqrt{-2}$. Beregn $\alpha+\beta$, $\alpha - \beta$ og $\alpha \cdot \beta$.

For et $\alpha = a + b\sqrt{-2}$, lader vi $\overline{\alpha} = a - b\sqrt{-2}$ betegne den *konjugerede* til α. Hvis f.eks. $\alpha = 2 + 3\sqrt{-2}$, er $\overline{\alpha} = 2 - 3\sqrt{-2}$.

Opgave 2: Lad $\alpha = a + b\sqrt{-2}$ og $\beta = c + d\sqrt{-2}$. Bevis, at $\overline{\alpha\beta} = \overline{\alpha} \cdot \overline{\beta}$. Vink: Skriv venstresiden og højresiden ud for sig og se, at de er ens.

Vi indfører nu normen N på $\mathbb{Z}[\sqrt{-2}]$ givet ved $N(a + b\sqrt{-2}) = a^2 + 2b^2$.

Opgave 3: Udregn $N(2 + 3\sqrt{-2})$ og $N(-5 + 7\sqrt{-2})$.

Opgave 4: Bevis, at $N(\alpha) = \alpha\overline{\alpha}$.

Opgave 5: Brug opgave 2 og 4 til at vise, at $N(\alpha\beta) = N(\alpha)N(\beta)$. Vi siger, at N er *multiplikativ*.

Vi vil nu vise, at $\mathbb{Z}[\sqrt{-2}]$ er euklidisk i en række trin. Vi skal vise, at for alle α, β med $\beta \neq 0$ eksisterer der $\gamma, \rho \in \mathbb{Z}[\sqrt{-2}]$, så

$$\alpha = \beta\gamma + \rho, \quad \text{hvor} \quad N(\rho) < N(\beta).$$

Opgave 6: Vis, at vi kan skrive

$$\frac{\alpha}{\beta} = \frac{a + b\sqrt{-2}}{N(\beta)} = \frac{a}{N(\beta)} + \frac{b}{N(\beta)}\sqrt{-2}$$

for nogle heltal a, b. Vink: Gang tæller og nævner i α/β med $\overline{\beta}$.

Opgave 7: Forklar, hvorfor vi kan vælge heltal c og d, så afstanden fra c til $a/N(\beta)$ er $\leq 1/2$, og afstanden fra d til $b/N(\beta)$ er $\leq 1/2$.

Lad nu $\gamma = c + d\sqrt{-2}$ og $\rho = \alpha - \beta\gamma$. Det er klart, at $\alpha = \beta\gamma + \rho$. Vi skal vise, at $N(\rho) < N(\beta)$. Vi vil faktisk vise, at $N(\rho) \leq \frac{3}{4}N(\beta)$.

Opgave 8: Vis, at

$$N(\rho) = N(\beta)\left(\left(\frac{a}{N(\beta)} - c\right)^2 + 2\left(\frac{b}{N(\beta)} - d\right)^2\right).$$

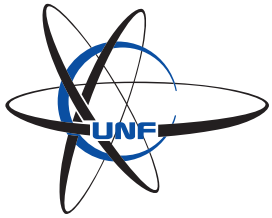
Vink: Skriv først

$$\rho = \beta\left(\frac{\alpha}{\beta} - \gamma\right)$$

og brug, at N er multiplikativ.

Opgave 9: Vis nu $N(\rho) \leq \frac{3}{4}N(\beta)$ og konkludér, at $\mathbb{Z}[\sqrt{-2}]$ er en euklidisk ring.

Opgave 10: Foretag euklidisk division med $\alpha = 2 + 3\sqrt{-2}$ og $\beta = 1 - \sqrt{-2}$.



Gauge Symmetrier 4

1 Introduktion

Naturen er fyldt med mønstre manifesteret gennem fysiske love. Men en fysisk lov beskriver ikke blot de observerbare fænomener som vi støder på i vores hverdag. I sidste ende er de rodfæstet i dybere strukturer som ligger til grund for universets historie og dets indhold. Ligesom ord i et sprog kan oversættes mellem sprog, er det ikke selve bogstaverne der bærer en betydning. Det er indholdet som overlever oversættelse fra et sprog til et andet. Fysiske love er ligeledes universelle i den forstand at det hverken er ordene, begreberne, eller systemet som bærer betydning. Det er den underliggende struktur og de mønstre som bevares mellem forskellige kontekster. Et fremragende eksempel på denne universalitet er Emmy Noethers fundamentale sætning om bevarelseslove og symmetrier. Historisk set blev denne vigtige sætning først præsenteret af Felix Klein den 7. juli i 1918 ved Königliche Gesellschaft der Wissenschaften i Göttingen, baseret på Noethers artikel "Invariante Variationsprobleme". Den forbinder symmetrier i et fysisk system med bevarelse af fysiske størrelser og er blevet en hjørnesten i moderne teoretisk fysik – alt fra partikelfysik til kosmologi. Hele dette forløb har til opgave at redegøre matematikken bag Noethers famøse sætning.

Et klassisk eksempel på denne sammenhæng er energibevarelse, som følger af en tidssymmetri i et system. Dette er ikke nødvendigvis trivielt. Men det betyder, at hvis et system forbliver uændret under

Figur 4.1: Emmy Noether.

tidsforskydninger, vil dets energi være konstant. Tilsvarende følger bevarelse af impuls fra systemets rumlige symmetri, mens bevarelsen af impulsmoment er forbundet med dets rotationsmæssige symmetri. Altså kan vi sige meget fundamentale ting om fysiske systemer, uden at vide noget som helst andet end dets symmetrier. Det er en voldsom stærk superkraft!

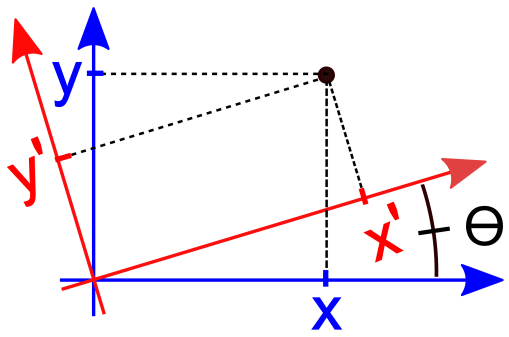
Symmetri		Bevarelse
Tid	$\longleftrightarrow$	Energi
Translation	$\longleftrightarrow$	Impuls
Rotation	$\longleftrightarrow$	Impulsmoment
Gauge symmetri (fx Standardmodellen $U(1) \times SU(2) \times SU(3)$)	$\longleftrightarrow$	Ladning
Bølgefunktionens fase i kvantemekanik	$\longleftrightarrow$	Sandsynlighedsdensitet
Overflade-tyngdekraft på en begivenhedshorisont	$\longleftrightarrow$	Entropien af det sorte hul

Inden for fysikkens verden skal symmetri forstås en form for invarians under en vis transformation. Med andre ord er der ingen forskel på før- og efter-billedet. I matematikkens verden kan dette repræsenteres ved hjælp af gruppeteori: En gruppe er en mængde af transformationer som bevarer et bestemt sæt egenskaber. Vores opgave bliver derfor at tage vores matematiske viden og se dem med fysiske briller, hvor visse transformationer rent faktisk har fysisk konsekvenser. For eksempel, hvis vi betragter et system af partikler, kan vi analysere, hvordan de ændrer sig når vi udfører matematiske transformationer på dem. Men på et abstrakt niveau kan vi også bruge de matematiske resultater fra gruppeteori til at fortælle os noget mere om fysikken som vi måske ikke ville forvente. Et af projekterne i dette forløb bliver at *forudsige* antallet af elementarpartikler (1 foton, 3 W/Z bosoner, 8 gluoner) – kun ved brug af gruppeteori! What!?!? Det viser bare, når man virkelig dykker dybt ned, at matematik og fundamental fysik holder hånd i hånd.

2 Transformationer og Gruppevirkninger

Transformationer

Når man skal løse et problem i fysik går man som regel til værks ved at vælge en matematisk model. Heldigvis har man en del frihed i det her valg: Der er mange forskellige måder at se en given situation. For eksempel via valg af koordinatsystem eller observatør. Man kalder dette en form for "frihedsgrad", som man har til rådighed til at løse sine problemer med. Men det vigtige er blot at man kan transformerer mellem de forskellige perspektiver på en måde der bevare den underliggende fysik. Lad os tage et eksempel!



Figur 4.2: To koordinat systemer der beskriver det samme system.

Eksempel 2.1. Tag to partikler med masse m_1 og m_2 med hastighed u_1 og u_2 (før de kolliderer). Vi kan frit vælge vores reference punkt til at være i massecentrum, som har position R. Lad derudover r_1 og r_2 være positionerne for de to partikler på et givet tidspunkt t før kollisionen. Da har vi at R bevæger sig med en hastighed, set fra

laboratoriets reference punkt,

$$\begin{aligned} V &= \frac{\mathrm{d}R}{\mathrm{d}t} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m_1 r_1 + m_2 r_2}{m_1 + m_2}\right) \\ &= \frac{m_1 u_1 + m_2 u_2}{m_1 + m_2}\,. \end{aligned}$$

Set fra R har vi derfor at de to partikler bevæger sig med hastighed $u_i' = u_i - V$. Vi ser

$$m_1 u_1' + m_2 u_2' = 0\,. \tag{4.1}$$

Vi har her implicit brugt at fysikens love er ens uanset hvordan man placerer sit koordinatsystem eller hvor hurtigt man bevæger sig. Analysen gør sig også gældende bagefter kollisionen hvor hastigheden er v_1 og v_2 set fra laboratoriets synspunkt er $v_i' = v_i - V$ samt

$$m_1 v_1' + m_2 v_2' = 0\,. \tag{4.2}$$

Men da har vi $m_1 u_1' + m_2 u_2' = m_1 v_1' + m_2 v_2'$ hvilket også må gælde sæt fra laboratorier reference punktet så

$$m_1 u_1 + m_2 u_2 = m_1 v_1 + m_2 v_2\,. \tag{4.3}$$

Vi aflæser at den totale impuls er bevaret! (givet vi kan lave de transformationer vi har gjort brug af). Det er nu meget mere realistisk at give sig i kast med at beregne hastighederne v_1 og v_2. ∘

Vi har i ovenstående brugt transformationer til at gøre vores liv lettere. Men de har ikke været helt vilkårlige. De er en slags *symmetrier* af det system vi arbejder med, i den forstand at efter vi har anvendt en transformation, er det essentielt det samme system vi kigger på. Lidt ligesom hvis man roterer en firkant 90 grader, at så ligner den stadig en firkant. Vi kan se at symmetrier i fysik er et kraftfuldt værktøj til at reducere kompleksiteten og udgør et brugbart regneværktøj. Vi har faktisk set at en symmetri i dette tilfælde

medførte at en egenskab er bevaret – nemlig impulsen. Motiveret af dette vil vi forsøge at præcisere det er der rent faktisk er sket; hvad menes der overhovedet med en symmetri, en fysisk system og bevarede størrelser. Hvordan kan disse idéer generaliseres?

Grupper

Centralt i det matematiske studie af symmetrier ligger konceptet af en gruppe. En gruppe er essentielt en samling af symmetrier. Definitionen er som følger:

Definition 2.2. En **gruppe** er en mængde G samt en kompositionsregel $\cdot\colon G \times G \to G$, der tager to elementer g,$h \in G$ og giver et nyt element $g \cdot h$, som overholder følgende regler:

1. Den er associativ: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.
2. Den indholder et identitetselement $e \in G$ således at $e \cdot g = g \cdot e = g \ \forall g \in G$.
3. For et givent element $g \in G$ findes der et invers element $h \in G$ således at $g \cdot h = h \cdot g = e$. Ofte skrives dette som $h = g^{-1}$.

Definitionen kan forståes på følgende måde: Man skal tænke på G som en samling af symmetrier af et eller andet objekt. For eksempel symmetrierne af en firkant. Man tænke på symmetrierne som at svare til en eller anden transformation af objektet der bevarer det. At man kan sammensætte dem $g \cdot h$ svarer da til at anvende transformationen h og derefter g. Associativitet svarer ligeledes til at det er rækkefølgen man anvender transformationerne som er væsentligt og ikke selve rækkefølgenden de sammensættes. Eksistensen af identiteten svarer til at "gøre ingenting" altid er en symmetri. Og inversen fortæller at enhver symmetri i virkeligheden er en bijektion. Altså hvis man har roteret firkanten 90 grader, må man kunne rotere den tilbage.

Fremover vil vi dog undgå at bruge prikken som kompositionsregel og bare skrive gh i stedet for $g \cdot h$. Kompositionsregelen lades ofte være implicit.

Definition her er abstraheret i den forstand at den ikke længere har noget direkte at gøre med symmetrier af et objekt. Det er relevant for os at kunne tænke mere præcist på hvordan symmetriere virker på et objekt.

Definition 2.3. Lad G være en gruppe og X en mængde. En gruppevirkning af G på X er et valg af transformationer $\lambda_g \colon X \to X$ for enhver $g \in G$ således at

$$\lambda_e = \mathrm{Id}_X \quad \text{og} \quad \lambda_g \circ \lambda_h = \lambda_{gh}\,.$$

Ofte kan man vælge at skrive $\lambda_g(x) = g.x$, $g \cdot x$ eller blot gx. Identitetsafbildningen $\mathrm{Id}_X \colon X \to X$ er den som sender alle elementer til sig selv, $\mathrm{Id}_X(x) = x$. Typisk har X også noget ekstra struktur man ønsker bevaret. I vores tilfælde vil X repræsentere de forskellige tilstande et fysisk system kan have, med dynamik som struktur (mere om dette senere). Vores gruppe skal derfor bevare dynamikken.

Lie grupper og Lie algebraer

I fysik er de symmetrier som vi oftest arbejder med kontinuerte. Det betyder også at vi har evnen til at variere transformationer på en kontinuert måde. Tag for eksempel en rotation i $\mathbb{R}^2$ med vinkel θ, som kan skrives på følgende form:

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\,.$$

Denne måde at skrive en $n \times m$ tabel, hvor n er antallet af rækker og m antallet af søjler (i dette tilfælde 2×2), kaldes også en **matrix**. En (kolonne-)vektor er for eksempel også en matrix, på formen $n \times 1$. Vi vil snart demonstrere hvordan man laver beregninger med matricer og hvordan de relateres til transformationer. Dét at transformationen er kontinuert implicerer at afbildningerne

$$(g,h) \mapsto gh \quad \text{og} \quad g \mapsto g^{-1}.$$

også skal være kontinuerte. Med 'kontinuert' menes der at en "lille" ændring i input giver en tilsvarende "lille" ændring i output. Faktisk vil vi gerne kunne regne ud *hvor meget* ændrer sig – det kalder man jo bare at differentiere! Derfor bruger man i fysik hvad man kalder Lie grupper.

Definition 2.4. En **Lie gruppe** er en gruppe G med nogle ekstra krav:

1. G er hvad man kalder en mangfoldighed. Det vil sige at hvis man zoomer nok ind ligner den bare $\mathbb{R}^n$. På pæn vis kan vi derfor bruge vores viden om differentialregning på den.

2. Multiplikation $(g,h) \mapsto gh$ og at tage invers $g \mapsto g^{-1}$ er differentiable afbildninger.

Eksempel 2.5. Et eksempel på en Lie gruppe kunne være $\mathbb{R}$ med $+$ som operation. Argumentet overlader vi til Opgave 1. ∘

Vi vil gerne beskrive grupper bestående af matricer, så lad os gennemgå hvordan man ganger to matricer sammen.

Metode 2.6. Lad A og B være to $n \times n$ matricer. At gange dem sammen, AB, minder meget om at tage prikproduktet mellem to vektorer. Dette kan ses hvis vi skriver A som en søjle af rækkevektorer og B som en række af søjlevektorer.

$$A = \begin{pmatrix} — & \vec{a}_1 & — \\ — & \vec{a}_2 & — \\ & \vdots & \\ — & \vec{a}_n & — \end{pmatrix}, \qquad B = \begin{pmatrix} | & | & & | \\ \vec{b}_1 & \vec{b}_2 & \dots & \vec{b}_n \\ | & | & & | \end{pmatrix}$$

Hver vektor $\vec{a}_i$, $\vec{b}_i$ består af n tal. Produktet AB svarer bare til at

man tager prikproduktet af de forskellige vektorer.

$$AB = \begin{pmatrix} — & \vec{a}_1 & — \\ — & \vec{a}_2 & — \\ & \vdots & \\ — & \vec{a}_n & — \end{pmatrix} \begin{pmatrix} \downarrow & \downarrow & & \downarrow \\ \vec{b}_1 & \vec{b}_2 & \dots & \vec{b}_n \\ | & | & & | \end{pmatrix}$$
$$= \begin{pmatrix} \vec{a}_1 \cdot \vec{b}_1 & \cdots & \vec{a}_1 \cdot \vec{b}_n \\ \vdots & \ddots & \vdots \\ \vec{a}_n \cdot \vec{b}_1 & \cdots & \vec{a}_n \cdot \vec{b}_n \end{pmatrix}.$$

Eksempel 2.7. Et mere avanceret eksempel kunne være mængden af 2×2 matricer som har en invers, sammen med matrixmultiplikation som kompositionsregel.

$$\mathrm{GL}(2,\mathbb{R}) \coloneqq \left\{ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| \, ad - bc \neq 0 \right\} \tag{4.4}$$

Gruppeoperationen er givet ved

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} w & x \\ y & z \end{pmatrix} = \begin{pmatrix} aw + by & ax + bz \\ cw + dy & cx + dz \end{pmatrix} \tag{4.5}$$

Og inverser er givet ved

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{4.6}$$

Begge operationer ser differentiable ud (det er de faktisk også). Og mængden selv ligner ret meget $\mathbb{R}^4$. Så det er faktisk en Lie gruppe. ∘

Det er generelt nemmest at forstille sig en Lie gruppe som at være en samling af $n \times n$ matricer med indgange i enten $\mathbb{R}$ eller $\mathbb{C}$ der overholder en eller anden regel. Næsten alle Lie grupper vi kommer til at støde på er af denne form.

Definition 2.8. Lad $\mathcal{M}$ være en mangfoldighed. Lad $C^\infty(\mathcal{M})$ være alle glatte funktioner $f\colon \mathcal{M} \to \mathbb{R}$. En tangentvektor $\mathcal{V}_p$ ved $p \in \mathcal{M}$ er en funktion $\mathcal{V}_p\colon C^\infty(\mathcal{M}) \to \mathbb{R}$ som overholder kædereglen og er lineær:

$$\mathcal{V}_p(\varphi\psi) = \varphi(p)\mathcal{V}_p(\psi) + \mathcal{V}_p(\varphi)\psi(p) \tag{4.7}$$

$$\mathcal{V}_p(a\varphi + b\psi) = a\mathcal{V}_p(\varphi) + b\mathcal{V}_p(\psi). \tag{4.8}$$

Samlingen af sådanne tangent vektorer skrives som $T_p\mathcal{M}$ og kaldes **tangentrummet** ved p. Et **vektorfelt** på $\mathcal{M}$ er en funktion $\mathcal{V}\colon p \mapsto \mathcal{V}_p \in T_p\mathcal{M}$.

Husk: Tangentvektorer spiser funktioner og giver et tal. Tallet $\mathcal{V}_p(f)$ svar til hvor hurtigt f ændrer sig i vektorens retning.

Definition 2.9. Lad $f\colon \mathcal{M} \to \mathcal{N}$ være en glat afbildning mellem to mangfoldigheder. For $q = f(p)$ definerer vi afbildningen, kaldet **differentialet**

$$\mathrm{d}f_p\colon T_p\mathcal{M} \to T_q\mathcal{N}$$

ved at sætte

$$\mathrm{d}f_p(\mathcal{V}_p)(\varphi) = \mathcal{V}_p(\varphi \circ f)$$

Tangentrummet giver et billede af hvordan mangfoldigheden ser ud helt tæt på og ligeledes giver differentialet en idé om hvordan funktioner ser ud helt tæt på.

Definition 2.10. Lie algebraen af en Lie gruppe er tangent rummet $\mathfrak{g} = T_eG$. Vi kan altid udvide et element $X \in \mathfrak{g}$ til et vektorfelt som følger: Lad $L_g\colon G \to G$ være givet ved $L_g(h) = gh$. Sæt da

$$X_g = \mathrm{d}(L_g)_e\,(X)\,. \tag{4.9}$$

Alternativt:

Definition 2.11. En Lie algebra er et tangentrum sammen med en operation $[\cdot,\cdot]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, kaldet **Lie parentesen**, som opfylder følgende egenskaber:

- Anti-kommutativitet
$$[X,Y] = -[Y,X]$$
- Bilinearitet
$$[aX + bY,Z] = a[X,Y] + b[Y,Z]$$
$$[z,aX + bY] = a[Z,X] + b[Z,Y]$$
hvor $a,b \in \mathbb{R}$.
- Jacobi identiteten
$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$$

Definition 2.12. En **en-parameter undergruppe** af en Lie gruppe G er en differentierbar afbildning $\gamma\colon \mathbb{R} \to G$ således at

$$\gamma(0) = e \quad \text{og} \quad \gamma(t+s) = \gamma(t)\gamma(s)\,. \tag{4.10}$$

Dette er præcis hvad motiverer Lie grupper. Vi har en måde at gå kontinuerligt fra at gøre intet til g. En en-parameter undergruppe er unikt bestemt af dens værdi $\gamma'(0)$. På den måde måler en Lie algebra $\mathfrak{g}$ små ændringer, i den forstand at elementerne af $\mathfrak{g}$ er præcis de mulige start-hastigheder for at gå fra at gøre intet til at gøre noget. For at give lidt uformel intuition lad $0 < \varepsilon \ll 1$ være et "lille" tal så $\varepsilon^2 \approx 0$. Lad derudover $\gamma'(0) = X$ og $\gamma(\varepsilon) = g$. Vi laver en naiv og formelt forkert første-ordens (Taylor-)approksimation

$$\begin{aligned} g &= \gamma(\varepsilon) \\ &= \gamma(0) + \varepsilon\gamma'(0) + \mathcal{O}(\varepsilon^2) \\ &= e + \varepsilon X + \mathcal{O}(\varepsilon^2) \\ &\approx e + \varepsilon X \end{aligned}$$

Så "små" transformationer eller gruppeelementer tæt på identiten svarer i virkligheden bare til elementer af $\mathfrak{g}$. Dette er moralen med Lie grupper og Lie algebraer.

3 Konfigurationsrum og Dynamik

Når vi taler om fysik i konteksten af matematik, er det ikke bare nok at lave en pæn tegning og skrive nogle formler og funktioner ned. Ligesom vi lærte i Kapitel 0, er funktioner nemlig defineret som afbildninger på mængder. Derfor skal vi spørge os selv: Hvilken mængde definerer et fysisk system?

Forestil en partikel der flyver rundt i rummet. For at sige noget om hvad den laver, skal man som minimum sige noget om dens position og hastighed – det vil sige to tal for hver dimension. Læg mærke til at der for hver position x_i langs en akse også skal findes et variabel p_i som fortæller én hvor hurtig ændringen i den position er langs den samme retning. Den tilsvarende hastighed p_i kalder man for det **kanoniske momentum** til x_i. Hver af de forskellige x_i'er kaldes en **frihedsgrad** og antallet af frihedsgrader kalder vi N. Derfor skriver vi typisk indekset $i = 1,2,\ldots,N$.[1] Et øjebliksbillede af et fysisk system er derfor et punkt i et koordinatsystem med $2N$ akser, fordi man skal specificere N positioner og N kanoniske momenta.

Definition 3.1. Et **konfigurationsrum** $\mathcal{K}$ er en mængde

$$\mathcal{K} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_N \times \mathcal{P}_1 \times \cdots \times \mathcal{P}_N \,,$$

hvor N er antallet af **frihedsgrader**. For hver frihedsgrad er der to mængder, $\mathcal{X}_i$ og $\mathcal{P}_i$ for $i = 1,\ldots,N$, som beskriver de værdier som positioner og kanoniske momenta kan tage.

Definition 3.2. Elementerne af et konfigurationsrum kaldes for **konfigurationer**.

$$(x_1,\ldots,x_N,p_1,\ldots,p_N) \in \mathcal{K}$$

Bemærkning 3.3. I klassisk fysik arbejder vi altid med systemer hvor positioner og hastigheder i princippet kan være et hvilket som

[1] For eksempel for $N = 3$ er du måske mere vant til at skrive $(x_1,x_2,x_3) = (x,y,z)$, som er et element i $\mathbb{R}^3$.

helst reelt tal. Derfor er det underforstået at

$$\mathcal{X}_i = \mathcal{P}_i = \mathbb{R}$$

for alle $i = 1, \ldots, N$. Fremover vil vi nogle gange skrive

$$\mathcal{K} = \mathbb{R}^{2N}\,.$$

Der er dog nogle tilfælde hvor det er værd at angive andre mængder. For eksempel hvis en partikel befinder sig på en cirkel $\mathcal{X} = S^1$ eller langs en kurve, som kan parametiseres med et interval $\mathcal{X} = [0,1]$.

Definition 3.4. Et **system** $\mathcal{S} = (H, \mathcal{K})$ med N frihedsgrader er et ordnet par, hvor $\mathcal{K}$ er et konfigurationsrum. Funktionen

$$H \colon \mathcal{K} \to \mathbb{R}$$
$$(x_1, \ldots, x_N, p_1, \ldots, p_N) \mapsto H(x_1, \ldots, x_N, p_1, \ldots, p_N)$$

kaldet en **Hamiltonian**, indkoder systemets dynamik. Dens værdi kan forstås som "energien" af systemet i tilstanden $(x_1, \ldots, x_N, p_1, \ldots, p_N)$. Vi vil senere se, hvorfor denne beskrivelse giver mening.

Proposition 3.5. Lad $\mathcal{S} = (H, \mathcal{K})$ være et system med N frihedsgrader og $k \in \mathcal{K}$ være en konfiguration. Dynamikken, dvs. formlerne der beskriver hvordan positioner $x_i(t)$ og kanoniske momenta $p_i(t)$ ændrer sig med tiden, er beskrevet af **Hamiltons ligninger**.

$$\boxed{\frac{\partial x_i}{\partial t} = \frac{\partial H}{\partial p_i}\,, \qquad \frac{\partial p_i}{\partial t} = -\frac{\partial H}{\partial x_i}\,.} \tag{4.11}$$

Symbolet ∂ kaldes det **partielle differentiale** (eller partielle afledede). Det er næsten ligesom et normalt differentiale "d", men med et lille twist: Når en funktion afhænger af flere variabler, differentierer man *kun* én af dem! Lad os se på et eksempel:

$$\begin{aligned} \frac{\partial}{\partial x}\Big(x(t)^2 + x(t)p(t) - p(t)^2\Big) &= 2x(t) + p(t) \\ \frac{\partial}{\partial p}\Big(x(t)^2 + x(t)p(t) - p(t)^2\Big) &= x(t) - 2p(t) \end{aligned} \tag{4.12}$$

Det partielle differentiale bruges også når man anvender kædereglen på funktioner af flere variable.

$$\frac{\mathrm{d}f(x(t),p(t))}{\mathrm{d}t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial p}\frac{\partial p}{\partial t} \tag{4.13}$$

Eksempel 3.6. En partikel i én dimension med masse m kan beskrives med Hamiltonianen

$$H(x,p) = \frac{p^2}{2m} + V(x)\,. \tag{4.14}$$

Funktionen $V(x)$ kaldes for et potentiale og beskriver hvor meget potentiel energi partiklen har. Den første Hamiltonligning giver

$$\frac{\partial x}{\partial t} = \frac{\partial H}{\partial p} = \frac{p}{m}\,. \tag{4.15}$$

Hvis vi omskriver dette får man

$$p = m\frac{\partial x}{\partial t} = mv\,. \tag{4.16}$$

Dette er præcis definitionen af impuls (også kaldet 'momentum') som vi kender det fra fysik. Den anden Hamiltonligning giver

$$\frac{\partial p}{\partial t} = -\frac{\partial V}{\partial x}\,. \tag{4.17}$$

Vi kan bruge det forrige resultat til at regne ud at højre side bare er masse ganget acceleration. Derfor kan vi regne ud at

$$F = -\frac{\partial V}{\partial x} = ma \tag{4.18}$$

bare er kræften som påvirker partiklen. Med andre ord er en kraft bare en *ændring i potentiel energi*. Hamiltonligningerne producerer Newtons 2. lov! ∘

Bemærkning 3.7. Hvis vi indsætter $p = mv$ i (4.14) ses det tydeligt at Hamiltonianen er den samlede energi i systemet.

$$\begin{aligned} H(x,p) &= \frac{1}{2}mv^2 + V(x) \\ &= \begin{matrix}\text{Kinetisk}\\ \text{energi}\end{matrix} + \begin{matrix}\text{Potentiel}\\ \text{energi}\end{matrix} \end{aligned} \tag{4.19}$$

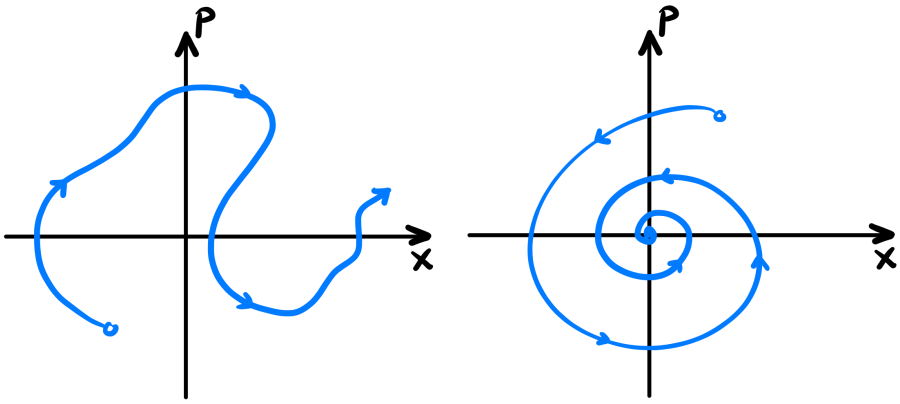
Definition 3.8. Lad $\mathcal{S} = (H,\mathcal{K})$ være et system. En kontinuert kurve i konfigurationsrummet

$$\gamma = \left\{ k(t) = \big(x_i(t),p_i(t)\big) \in \mathcal{K} \;\middle|\; t \in [0,1] \right\} \subset \mathcal{K} \tag{4.20}$$

som overholder Hamiltons ligninger beskriver systemets tilstand som funktion af tid t. Dette kaldes også systemets **historie**.[2] Hvis der eksisterer en tid T således at

$$k(T) = k(T') \qquad \forall T' > T\,, \tag{4.21}$$

så siges det at system er nået **ligevægt**.



Figur 4.3: Skitse af en historie og en ligevægt.

Sætning 3.9. Lad $\mathcal{S} = (H,\mathcal{K})$ være et system. Der er energibevarelse i systemet hvis og kun hvis H kun *eksplicit* er funktion af tid t.

$$\boxed{\frac{\partial H}{\partial t} = 0 \qquad \Longleftrightarrow \qquad \text{Energibevarelse}} \tag{4.22}$$

Med andre ord, hvis man skriver H som en funktion, inkluderer udtrykket udelukkende positioner x_i og kanoniske momenta p_i.

[2] I fysikliteraturen kaldes kurven kun bestående af positioner $(x_1(t),\ldots,x_N(t))$ også for systemets "worldline".

Eksempel 3.10. Eksempler på Hamiltonians hvor de respektive systemer har/ikke har energibevarelse:

$$\begin{aligned} \text{Energibevarelse:} \qquad & H = \frac{p^2}{2m} + \frac{1}{2}m^2\omega^2 x^2 \\ \text{Ikke energibevarelse:} \qquad & H = t\frac{p^3}{2m^2 x} \end{aligned} \tag{4.23}$$

∘

Vi kan forstå forbindelsen mellem energibevarelse og tidsafhængighed ud fra vores fortolkning af H som en energi. Tydeligvis, hvis H ændrer sig med tid, så må energien af systemet også ændre sig. Men H kan jo afhænge af både $x(t)$ og $p(t)$, som hver afhænger af tid. Så hvorfor er det kun den *eksplicitte* tids-afhængighed, som betyder noget? Med andre ord, hvorfor taler Sætning 3.9 om $\frac{\partial H}{\partial t}$ og ikke det totale differentiale $\frac{\mathrm{d}H}{\mathrm{d}t}$? For at se hvorfor det giver mening, lad som om $H = H(t,x(t),p(t))$ er en funktion af både tid, position og kanonisk momentum. Nu kan vi bruge kædereglen og Hamiltons ligninger til at beregne det totale differentiale.

$$\begin{aligned} \frac{\mathrm{d}H}{\mathrm{d}t} &= \frac{\partial H}{\partial t} + \frac{\partial H}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial H}{\partial p}\frac{\partial p}{\partial t} \\ &= \frac{\partial H}{\partial t} - \frac{\partial p}{\partial t}\frac{\partial x}{\partial t} + \frac{\partial x}{\partial t}\frac{\partial p}{\partial t} \\ &= \frac{\partial H}{\partial t} \end{aligned} \tag{4.24}$$

Ergo, hvis H ikke eksplicit afhænger af tid, så afhænger den slet ikke af tid! Hvad kan vi fortolke, ud fra et fysisk perspektiv, hvis energien ikke er bevaret?

1) **Der mangler noget.**
   Hvis energi ikke er bevaret kan det være tegn på at man har glemt at tilføje et udtryk til sin Hamiltonian. Det kan være at der mangler nogle partikler, kræfter, eller interaktioner.

2) **Hamiltonianen er i virkeligheden ikke energien.**
   Næsten det samme som det første punkt, med den forskel at

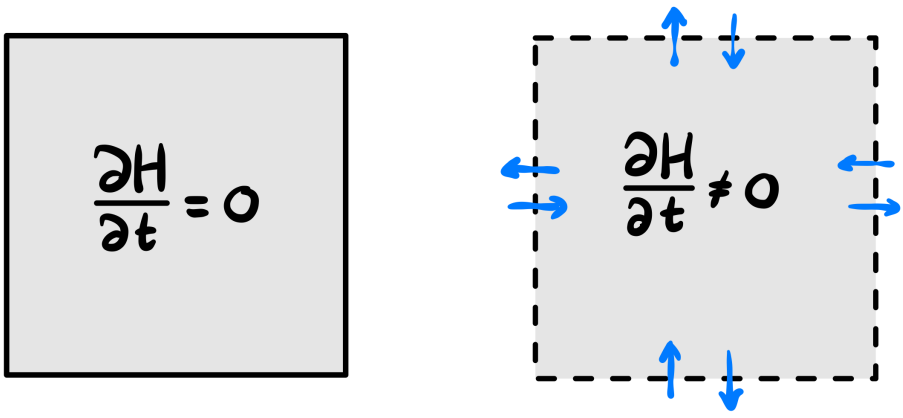
vi har lavet en forkert antagelse om, hvad energien af systemet i virkeligheden svarer til. Nogle gange kan en funktion H godt give den samme dynamik, men den giver ikke den rigtige fortolkning af systemets energi. Opgave 5.13 giver et eksempel når et system er i bevægelse.

3) Man accepterer at energi ikke er bevaret.
Energibevarelse er i sidste ende bare en ideel antagelse, som ikke nødvendigvis behøver at passe i virkeligheden. To årsager kan være:

- Hvis et system $\mathcal{S}$ er et udsnit af et større system $\mathcal{S}'$ med mindst lige så mange frihedsgrader, er energien af $\mathcal{S}$ ikke altid bevaret. Set fra perspektivet af $\mathcal{S}$ forstås det som at systemet udveksler energi med "omgivelserne". Men fra perspektivet af $\mathcal{S}'$ kan den totale energi godt være bevaret. Nogle fysiske eksempler er en utæt beholder, turbulens, friktion og kosmisk inflation.

- Energi er bare ikke bevaret. Punktum. Et funky eksempel er universet. Den korte forklaring er, at fordi universet udvider sig, så er den totale energi ikke bevaret (men det er meget forsimplet og ikke fuldstændig korrekt). Den længere forklaring har noget at gøre med såkaldte 'tidslige Killing vektorfelter'[3] og hvordan de genererer symmetrier. Selvom vi ikke kommer til at lære om hvad 'tidslig' og 'Killing' betyder, kommer vi senere ind på vektorfelter og hvordan de giver anledning til symmetrier. Det er ret cool.

[3]Nej, den har hverken katteører eller siger miav. Det er opkaldt efter Wilhelm Killing, en tysk matematiker fra 1800-tallet.



Figur 4.4: Systemer med/uden energibevarelse. Det ene system er komplet mens det andet udveksler energi med omgivelserne.

Dynamik er bare (symplektisk) geometri

Indtil videre har vi lært at Hamiltonianen H og Hamiltons lignigner beskriver hvordan en konfiguration k udvikler sig i konfigurationsrummet $\mathcal{K}$. Men som vi bemærkede i forrige afsnit om transformationer, kan man beskrive det samme system med forskellige koordinater. Hvordan kan vi forstå dette i et matematisk sprog? Den struktur vi gerne vil bevare er dynamikken af systemet, dvs. Hamiltons ligninger, som relaterer x,p med $\frac{\partial x}{\partial t}, \frac{\partial p}{\partial t}$. Vi kan vælge at skrive dem på følgende form:

$$\begin{pmatrix} \frac{\partial x}{\partial t} \\ \frac{\partial p}{\partial t} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial H}{\partial x} \\ \frac{\partial H}{\partial p} \end{pmatrix} \tag{4.25}$$

$$\frac{\partial k}{\partial t} = \Omega \cdot \nabla H$$

Symbolet ∇ (udtalt 'nabla') er en vektor som differentierer hvad der står foran den med hensyn til koordinaterne i konfigurationsrummet. ∇H fortæller derfor noget om *hvor meget* H ændrer sig i en bestemt retning i $\mathcal{K}$. Bevægelsesligningerne som beskriver relationen mellem

$\frac{\partial x}{\partial t}, \frac{\partial p}{\partial t}$ og ∇H afhænger derfor af objektet Ω. Dette er strukturen som skal bevares, hvis vi skal sørge for at dynamikken ikke ændrer sig!

Definition 3.11. Den **symplektiske struktur** er en $2N \times 2N$ matrix som definerer dynamikken af et system $\mathcal{S}$ med N frihedsgrader. For $N = 1$ frihedsgrad har vi

$$\boxed{\Omega \coloneqq \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.} \tag{4.26}$$

For N frihedsgrader,

$$\Omega \coloneqq \begin{pmatrix} & & & 1 & & \\ & 0 & & & \ddots & \\ & & & & & 1 \\ -1 & & & & & \\ & \ddots & & & 0 & \\ & & -1 & & & \end{pmatrix}. \tag{4.27}$$

De transformationer der bevarer den symplektiske struktur definerer den symplektiske gruppe.

Definition 3.12. Den **symplektiske gruppe** $\mathrm{Sp}(2N)$ er gruppen af transformationer som bevarer den symplektiske struktur for et system med N frihedsgrader.

$$\mathrm{Sp}(2N) \coloneqq \left\{ A = 2N \times 2N \text{ matrix} \mid A^T \Omega A = \Omega \right\} \tag{4.28}$$

Eksempel 3.13. Den symplektiske gruppe $\mathrm{Sp}(2)$ indeholder bl.a. følgende tre elementer

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right\}. \tag{4.29}$$

For eksempel kan vi se at

$$\begin{aligned} A^T\Omega A &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1\cdot 0 + 1\cdot(-1) & 1\cdot 1 + 1\cdot 0 \\ 0\cdot 0 + 1\cdot(-1) & 0\cdot 1 + 1\cdot 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} (-1)\cdot 1 + 1\cdot 1 & (-1)\cdot 0 + 1\cdot 1 \\ (-1)\cdot 1 + 0\cdot 1 & (-1)\cdot 0 + 0\cdot 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \end{aligned} \tag{4.30}$$

∘

Nu når vi ved hvilke transformationer det er der bevarer den symplektiske struktur, kan vi nu spørge os selv, hvordan vi må transformere koordinaterne x,p. Det svarer til at vi skal bestemme de transformationer A, hvor

$$k \mapsto k' = Ak$$

bevarer bevægelsesligningerne.

$$\begin{aligned} k = \Omega\cdot\nabla H \quad \longmapsto \quad k' &= Ak \\ &= A\Omega\cdot\nabla H \\ &= A\Omega A^T\cdot\nabla' H \end{aligned} \tag{4.31}$$

I det sidste trin har vi anvendt følgende resultat.

$$k' = Ak \iff A^T\nabla' = \nabla \tag{4.32}$$

Hvis dynamikken i koordinaterne k' skal give de samme ligninger, kræver det at

$$A\Omega A^T \stackrel{!}{=} \Omega\,, \tag{4.33}$$

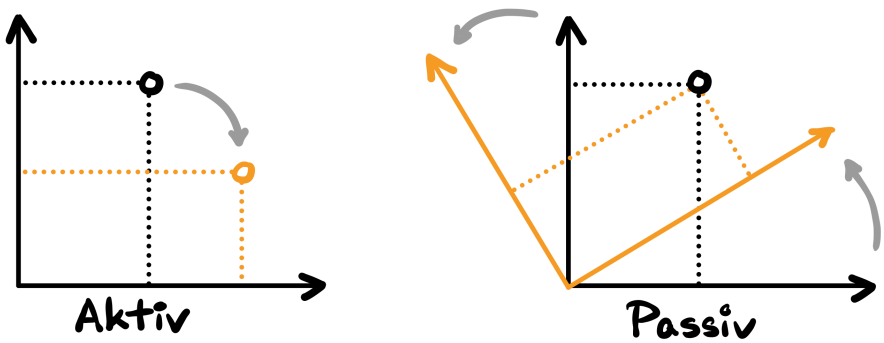
hvilket bare er definitionen på en symplektisk matrix!

Definition 3.14. De transformationer som bevarer dynamikken kaldes **kanoniske transformation** eller en **symplektomorfisme**. Det er en afbildning $k \mapsto Ak$ hvor $A \in \mathrm{Sp}(2N)$.

Hvorfor og hvordan kan man tænke på transformationer som dynamik på et konfigurationsrum? Dét som en kanonisk transformation gør, er i virkeligheden at afbilde punkter $x_i \mapsto x_i'$ og $p_i \mapsto p_i'$. Du kan vælge at fortolke dette på to forskellige måder.

Aktiv transformation: Alle punkter $k \in \mathcal{K}$ *flytter sig* "aktivt" til nye punkter $k' = Ak \in \mathcal{K}$ i et fast koordinatsystem.

Passiv transformation: Alle punkter holdes fast (de er "passive"), men *akserne* i koordinatsystemet flytter sig i modsat retning.



Figur 4.5: Aktiv og passiv transformation

Dynamik er bare et vektorfelt

Måske er kanoniske transformationer som bevarer den symplektiske struktur ikke den mest intuitive måde at tænke på dynamik af et fysisk system... Heldigvis kan vi beskrive dynamik fra et andet perspektiv som minder meget om hvad man laver i fysik: Vi tænker på dynamik i form af vektorer og vektorfelter, som beskriver retningen ting bevæger sig i. Dog er der én forskel! I fysik tænker vi altid på

dynamik som vektorer i *rummet* $(x_1,\ldots,x_N) \in \mathcal{X}$, hvor selve vektoren siger noget om hastigheden (eller impulsen) $(p_1,\ldots,p_N) \in \mathcal{P}$ i ét bestemt punkt. Fordelen ved dette billede er at det er meget intuitivt; du kan fornemme i hvilken retning punkter vil bevæge sig. *Men* det siger intet om, hvordan hastigheder ændrer sig! Derfor skal vi forstå vektorer på en anderledes måde – nemlig som vektorer i *konfigurationsrummet* $\mathcal{K}$. En vektor i et punkt i $\mathcal{K}$ spiller i virkeligheden spillet "Du giver mig en konfiguration og jeg fortæller dig, hvilken konfiguration systemet snart kommer til at være i". Men det siger kun noget om, hvordan én bestemt konfiguration kommer til at ændre sig. For at beskrive hele dynamikken skal vi bruge en vektor i hvert eneste punkt i $\mathcal{K}$... Det er præcis et vektor*felt*!

En anden detalje er, at vi gerne vil tænke på vektorfelter som en form for operation som måler en *ændring* af noget. Betragt en funktion $f\colon \mathcal{K} \to \mathbb{R}$. Vi kan spørge, hvor meget $f(x,p)$ ændrer sig i x-retningen. Vi kan beskrive denne retning som en vektor

$$\vec{V} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Men vi ved allerede hvor meget $f(x,p)$ ændrer sig med x. Det er præcis differentialet $\frac{\partial f}{\partial x}$! Hvad nu hvis vi gangede vektoren $\vec{V}$ med to? Intuitivt fårstår vi det som at "hastigheden" er dobbelt så stor. Så med andre ord er ændringen af $f(x,p)$ bare det dobbelte af, hvad vi havde før, $2\frac{\partial f}{\partial x}$. Hvad med en vektor som peger både langs x og p?

$$\vec{V} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Heldigvis er det nemt: Den totale ændring er bare summen af ændringerne i hver af de to retninger x og p, $\frac{\partial f}{\partial x} + \frac{\partial f}{\partial p}$.

Pointen er, at *for enhver vektor findes der et differentiale som måler hvor meget noget ændrer sig langs vektoren.*

$$\vec{V} = \begin{pmatrix} a \\ b \end{pmatrix} \quad \longleftrightarrow \quad \mathcal{V} = a\frac{\partial}{\partial x} + b\frac{\partial}{\partial p} \tag{4.34}$$

Definition 3.15. Et **vektorfelt** er en afbildning $\mathcal{V}\colon f \mapsto \mathcal{V}(f)$ som tager funktioner $f\colon \mathcal{K} \to \mathbb{R}$ og sender dem til andre funktioner $\mathcal{V}(f)\colon \mathcal{K} \to \mathbb{R}$. De måler essentielt hvor meget $f(k)$ ændrer sig langs en bestemt retning i $\mathcal{K}$.

Bemærkning 3.16. Afbildningen $\mathcal{V}(f)$ som måler hvor meget f ændrer sig langs vektorfeltet $\mathcal{V}$ kaldes også for **Lie differentialet**

$$\mathcal{L}_{\mathcal{V}}(f) = \mathcal{V}(f)\,. \tag{4.35}$$

Lad os gøre det mere tydeligt, hvordan vektoren $\vec{V}$ relaterer sig til vektorfeltet $\mathcal{V}$ beskrevet som et differentiale. Til det skal vi bruge ∇ som vi beskrev tidligere i afsnittet. Lad $f\colon \mathcal{K} \to \mathbb{R}$ være en funktion. Husk at ∇f, også kaldet **gradienten** af f, giver en vektor i den retning hvor f stiger allerhurtigst. Betragt nu prikproduktet

$$\vec{V} \cdot \nabla f = \begin{pmatrix} a & b \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial p} \end{pmatrix} = a\frac{\partial f}{\partial x} + b\frac{\partial f}{\partial p}\,. \tag{4.36}$$

Det er præcis det samme resultat som hvis man evaluerede vektorfeltet $\mathcal{V}$ svarende til $\vec{V}$ på f!

$$\boxed{\mathcal{L}_{\mathcal{V}}(\bullet) = \mathcal{V}(\bullet) = \left(\vec{V} \cdot \nabla\right)(\bullet)} \tag{4.37}$$

Nu kan også forstå meget bedre, hvorfor $\mathcal{V}$ svarer til en ændring *langs* vektoren $\vec{V}$. Det er fordi ∇ måler ændringen af funktionen f, og prikproduktet $\vec{V} \cdot \nabla$ derefter måler, hvor stor den ændring er i retningen af $\vec{V}$ (det er "projektionen" langs $\vec{V}$).

Med vektorfelter som værktøjskasse bliver det meget nemt at forstå dynamikken af et system. I en fysisk forstånd vil vi nemlig gerne undersøge, hvordan konfigurationer, funktioner på $\mathcal{K}$ osv. ændrer sig med *tid*. Det viser sig at der er nogle meget specielle vektorfelter som beskriver netop dette! Med andre ord, vil vi se på vektorfelter som peger i samme retning som hvis man lader uret tikke og betragter, hvordan konfigurationer bevæger sig.

Definition 3.17. Et vektorfelt som kan skrives på formen

$$\mathcal{V}_X = \frac{\partial X}{\partial p}\frac{\partial}{\partial x} - \frac{\partial X}{\partial x}\frac{\partial}{\partial p} \tag{4.38}$$

hvor $X\colon \mathcal{K} \to \mathbb{R}$ er en funktion, kaldes et **Hamilton vektorfelt**.

Proposition 3.18. For alle funktioner på et konfigurationsrum findes der et tilsvarende Hamilton vektorfelt, $X \mapsto \mathcal{V}_X$.

Eksempel 3.19. Hamilton vektorfeltet svarende til Hamiltonianen H.

$$\mathcal{V}_H = \frac{\partial H}{\partial p}\frac{\partial}{\partial x} - \frac{\partial H}{\partial x}\frac{\partial}{\partial p} = \frac{\partial x}{\partial t}\frac{\partial}{\partial x} + \frac{\partial p}{\partial t}\frac{\partial}{\partial p} = \frac{\mathrm{d}}{\mathrm{d}t} \tag{4.39}$$

Derfra kan vi se at

$$\mathcal{V}_H(f) = \frac{\mathrm{d}f}{\mathrm{d}t}\,. \tag{4.40}$$

Hamiltonianen giver dermed anledning til et Hamilton vektorfelt som beskriver, hvordan funktioner ændrer sig med tid! ∘

Eksempel 3.20. I stedet for funktioner, lad os se hvad det samme vektorfelt gør ved konfigurationer $\mathcal{K} \ni k = (x,p)$, som kan skrives som en vektor $\begin{pmatrix} x \\ p \end{pmatrix}$. Da er

$$\begin{aligned} \mathcal{V}_H(k) &= \left(\frac{\partial H}{\partial p}\frac{\partial}{\partial x} - \frac{\partial H}{\partial x}\frac{\partial}{\partial p}\right)\begin{pmatrix} x \\ p \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial H}{\partial p} \\ -\frac{\partial H}{\partial x} \end{pmatrix} \\ &= \Omega \cdot \nabla H \\ &= \frac{\partial}{\partial t}\begin{pmatrix} x \\ p \end{pmatrix} = \frac{\partial k}{\partial t} \end{aligned} \tag{4.41}$$

Dette er præcis Hamiltons ligninger! ∘

Eksempel 3.21. Sidst, men ikke mindst, lad os se hvordan et Hamilton vektorfelt opererer på Hamiltonianen.

$$\begin{aligned}\mathcal{V}_f(H) &= \frac{\partial f}{\partial p}\frac{\partial H}{\partial x} - \frac{\partial f}{\partial x}\frac{\partial H}{\partial p} \\ &= -\frac{\partial f}{\partial p}\frac{\partial p}{\partial t} - \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} \\ &= -\left(\frac{\partial x}{\partial t}\frac{\partial f}{\partial x} + \frac{\partial p}{\partial t}\frac{\partial f}{\partial p}\right) \\ &= -\frac{\mathrm{d}f}{\mathrm{d}t}\end{aligned} \tag{4.42}$$

Hamiltonianen ændrer sig langs et Hamilton vektorfelt, hvis funktionen den er defineret fra, er tidsafhængig. Allerede her ser vi tegn på et meget vigtigt faktum. Nemlig at hvis f er en konstant funktion af tid, så er Hamiltonianen (og dermed energien af systemet), bevaret. Vi vender tilbage til dette is Afsnit 4. ∘

Vi har nu beskrevet hvordan systemet ændrer sig ud fra vektorfelter, men ikke hvordan systemet ser ud *efter* det har ændret sig. Til det skal vi undersøge vektorfelter på konfigurationer $k \in \mathcal{K}$. Her skal vi være forsigtige, fordi et vektorfelt $\mathcal{V}_X(k)$ fortæller kun hvordan k ændrer sig i én instans. Med andre ord, efter der er gået noget tid og $k \mapsto k'$, er der en ny vektor $\mathcal{V}_X(k')$, som vi skal følge. Hver gang konfigurationen opdateres, skal vi derfor i princippet beregne en ny vektor. Men det lyder rigtig indviklet og unødvendigt! Er der en måde hvorpå vi kan beregne 'store spring' efter der er gået 'lang tid'? Heldigvis er svaret Ja!™

Lemma 3.22. Eksponentialfunktionen kan skrives som en uendelig sum

$$\boxed{e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{1}{2}x^2 + \cdots} \tag{4.43}$$

For $|x| \ll 1$ kan det være brugbart at approksimere $e^x \approx 1 + x$.

Definition 3.23. Et **flow** af et vektorfelt $\mathcal{V}$ på et konfigurationsrum $\mathcal{K}$ er en afbildning

$$\begin{aligned} \Phi_{\mathcal{V}} \colon \mathbb{R} \times \mathcal{K} &\to \mathcal{K} \\ (t,k) &\mapsto \Phi_{\mathcal{V}}(t,k) \end{aligned} \tag{4.44}$$

som opfylder følgende egenskaber:

$$\Phi_{\mathcal{V}}(0,k) = k \tag{4.45}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi_{\mathcal{V}}(t,k) = \mathcal{V}\left(\Phi_{\mathcal{V}}(t,k)\right) \tag{4.46}$$

Hvad er forskellen mellem et flow og et vektorfelt? Et vektorfelt beskriver retningen som konfigurationer bevæger sig i. Et flow beslutte hvor konfigurationerne ender efter tiden t. Ligning 4.45 siger ikke andet end at konfigurationer ikke har bevæget sig endnu når tiden er $t = 0$. Ligning 4.46 beskriver ligeledes at banen som konfigurationer bevæger sig i ("flowet") følger vektorfeltet. Hvis man sætter $t = 0$ får man ligeledes, at starthastigheden af k netop er $\mathcal{V}(k)$.

Bemærkning 3.24. Et flow kan skrives som en afbildning

$$\begin{aligned} e^{t\mathcal{V}} \colon \mathcal{K} &\to \mathcal{K} \\ k &\mapsto \Phi_{\mathcal{V}}(t,k) \end{aligned} \tag{4.47}$$

kaldet **eksponentialafbildningen**. Den flytter alle konfigurationer til nye konfigurationer svarende til efter der er gået en tid t.

Eksponentialafbildningen er vildt brugbar, fordi den kan beskrive både 'små' og 'store' spring! Før, da vi kun havde at gøre med vektorfeltet, kunne vi nemlig kun se hvad der skete med k i meget små bidder. Men vi kan nu se eksplicit, hvorfor $e^{t\mathcal{V}}$ giver et flow:

$$e^{t\mathcal{V}}(k)\big|_{t=0} = k \tag{4.48}$$

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left[e^{t\mathcal{V}}(k)\right]\Big|_{t=0} &= \frac{\mathrm{d}}{\mathrm{d}t}\left[1 + t\mathcal{V} + \frac{1}{2}t^2\mathcal{V}^2 + \cdots\right](k)\Big|_{t=0} \\ &= \left[\mathcal{V} + t\mathcal{V}^2 + \cdots\right](k)\big|_{t=0} \\ &= \mathcal{V}(k) \end{aligned} \tag{4.49}$$

Dynamik er bare Lie Grupper/Algebra

Helt i starten af dette forløb diskuterede vi hvordan matematiske grupper kan forstås i form a transformationer. En gruppe virker på et konfigurationsrum ved at flytte alle punkter til andre punkter. Dette er en *global* effekt fordi gruppeelementet kan påvirke hele konfigurationsrummet på samme tid; $G \ni g \colon \mathcal{K} \to \mathcal{K}$. Men det giver os kun et før/efter billede – ikke selve dynamikken! For at se hvad der sker løbende skal vi derfor kigge på små transformationer, eller, set med fysiske briller, "små tidsintervaller". Det er præcis dét vores Lie algebra er til!

For at forstå hvordan en gruppe G påvirker et konfigurationsrum $\mathcal{K}$, er det indsigtsfuldt at se på funktioner på $\mathcal{K}$, $f \colon \mathcal{K} \to \mathbb{R}$. Det er fordi vi kan tegne forbindelser mellem funktioner og Hamilton vektorfelter, som vi indså tidligere.

Eksempel 3.25. Lad $f \colon \mathbb{R} \to \mathbb{R}$ være en funktion givet ved $f(x) = x^2$. En anden måde at tænke på funktionen er som en mængde $F \subset \mathbb{R}^2$

$$F = \left\{(x,y) \in \mathbb{R}^2 \mid y = f(x)\right\} .$$

Lad $g \in G$ være et gruppeelement som virker på $x \in \mathbb{R}$ ved $g \cdot x = x + 1$ (en translation). Så hvis g virker på mængden F svarer det til at flytte alle $x \mapsto x + 1$.

$$\begin{aligned} g \cdot F &= \left\{(x+1,y) \in \mathbb{R}^2 \mid y = f(x)\right\} \\ &= \left\{(x,y) \in \mathbb{R}^2 \mid y = f(x-1)\right\} \end{aligned}$$

Men bemærk at betingelsen $y = f(x)$ er den samme! Hvis man visualiserer ovenstående mængde som et koordinatsystem med kurven $f(x) = x^2$, er det som at vi har flyttet koordinatsystemet. Men vi kunne ligeledes vælge at beholde koordinatsystemet, men flytte kurven – bare i den *modsatte* retning. ∘

Ovenstående eksempel viser en meget generel egenskab for grup-

pevirkninger på funktioner på $\mathcal{K}$. Lad $g \in G$, $f\colon \mathcal{K} \to \mathbb{R}$ og $k \in \mathcal{K}$.

$$\boxed{(g \cdot f)(k) = f\left(g^{-1} \cdot k\right)} \tag{4.50}$$

Definition 3.26. Lad G være en (matrix) Lie gruppe. En **Lie algebra** $\mathfrak{g}$ består af de elementer (matricer) $L \in \mathfrak{g}$ således at

$$e^{tL} \in G\,, \qquad t \in \mathbb{R}\,.$$

Med andre ord giver eksponentialafbildningen af Lie algebra elementer, elementer af Lie gruppen.

En Lie algebra definerer små transformationer af en Lie gruppe og giver ligeledes anledning til et vektorfelt på konfigurationsrummet. Betragt en gruppevirkning på funktion og tag tidsdifferentialet evalueret ved $t = 0$ (med andre ord, 'starthastigheden' af gruppevirkningen)

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(g \cdot f)(k)\bigg|_{t=0} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{tL} \cdot f\right)(k)\bigg|_{t=0} \\ &= (L \cdot f)(k)\end{aligned}$$

Men også, ved brug af Ligning (4.50),

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{tL} \cdot f\right)(k)\bigg|_{t=0} &= \frac{\mathrm{d}}{\mathrm{d}t} f\left(e^{-tL} \cdot k\right)\bigg|_{t=0} \\ &= -\mathcal{V}_L\left(f(k)\right)\,.\end{aligned}$$

Derfor er der en direkte sammenhæng mellem Lie algebra elementer og vektorfelter

$$\mathfrak{g} \ni L \quad \longleftrightarrow \quad \mathcal{V}_L(\bullet)\,.$$

Hvis $\mathcal{V}_L$ også er et Hamilton vektorfelt, kan vi danne forbindelser mellem Lie algebra elementer og funktioner på $\mathcal{K}$.

Definition 3.27. Lad $\mathfrak{g}$ være en Lie algebra og $\mathcal{K}$ et konfigurationsrum. Et **co-moment** er en afbildning

$$\widehat{\mu}\colon L \longmapsto f_\mu$$

der tager et Lie algebra element $L \in \mathfrak{g}$ og sender det til en funktion $f_\mu \colon \mathcal{K} \to \mathbb{R}$, hvis

$$\mathcal{V}_L = \mathcal{V}_{f_\mu} .$$

Et **moment** er den tilsvarende afbildning

$$\mu \colon k \longmapsto L^* .$$

Her skal "L^*" læses som et objekt der tager et Lie algebra element og spytter et tal ud. Med andre ord, $\mu(k) \colon \mathfrak{g} \to \mathbb{R}$. Momentet og co-momentet opfylder

$$\mu(k)(L) = \widehat{\mu}(L)(k) = f_\mu(k) .$$

Eksempel 3.28. Betragt gruppen af translationer $G = \mathbb{R}$ som virker på $\mathcal{K} = \mathbb{R}^2$ via

$$g \cdot (x,p) = (x + a,p) , \quad a \in \mathbb{R} .$$

Det tilsvarende Lie algebra element og vektorfelt er

$$L = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \quad \Longleftrightarrow \quad \mathcal{V}_L = a\frac{\partial}{\partial x} .$$

For at finde momentet og co-momentet skal vi derfor spørge os selv, hvilken funktion $f_\mu(k)$ giver anledning til et vektorfelt $\mathcal{V}_{f_\mu} = \mathcal{V}_L$. Det rigtige svar viser sig at være

$$f_\mu(k) = f_\mu(x,p) = ap$$

fordi

$$\mathcal{V}_{f_\mu} = \frac{\partial f_\mu}{\partial p}\frac{\partial}{\partial x} - \frac{\partial f_\mu}{\partial x}\frac{\partial}{\partial p} = a\frac{\partial}{\partial x} .$$

Nu kan vi bestemme momentet og co-momentet. Først, husk at co-momentet $\widehat{\mu}$ er en funktion som tager et $L \in \mathfrak{g}$ (i dette tilfælde svarer det til at vælge et $a \in \mathbb{R}$) og giver dig et f_μ. Vi kan derfor aflæse at $\widehat{\mu}(L) = f_\mu = a$. Ligeledes er momentet en funktion som tager et $k \in \mathcal{K}$ og giver dig et objekt som spiser et L (eller a her). Så $\mu(k) = p$. ∘

4 Noethers Sætning

Nu er vi i stand til at formulere og forstå Noethers sætning. Den overordnede idé er at der kan eksistere et flow langs et vektorfelt hvor Hamiltonianen af et system er konstant. Fra det kan vi vise at der eksisterer et objekt som er konstant langs systemets historie (flowet svarende til $\mathcal{V}_H$). Med andre ord har vi en bevarelseslov! Vi vil demonstrere i flere eksempler og opgaver hvordan dette fungerer.

Sætning 4.1 (**Noethers Sætning (Version 1)**). Lad $\mathcal{S} = (H,\mathcal{K} = (\mathcal{X},\mathcal{P}))$ være et system og G være en Lie gruppe med tilsvarende gruppevirkning på $\mathcal{X}$, som bevarer den symplektiske struktur (ergo er Hamiltons ligninger bevarede). Comomentet $\widehat{\mu}(L)\colon \mathcal{K} \to \mathbb{R}$ er konstant, hvis og kun hvis, Hamiltonianen H er bevaret langs vektorfeltet $\mathcal{V}_L$, hvor $L \in \mathfrak{g}$.

Bevis. Vores strategi er følgende: Vi vil gerne beskrive små transformationer på baggrund af gruppen G. Derefter skal vi finde en funktion som giver anledning til et Hamilton vektorfelt som bevarer Hamiltonianen H. Til sidst vil vi vise, at denne funktion kan relateres til et (co-)moment, ligesom i Definition 3.27.

Betragt en kurve gennem punktet $x \in \mathcal{X}$, beskrevet via gruppevirkningen af G. Så $\gamma\colon \mathbb{R} \times \mathcal{X} \to \mathcal{X}$ og $\gamma(t,x) = g(t) \cdot x$, hvor $g(t) \in G$ og $g(0) = e$. Den lille transformation gennem x er præcis

$$\delta x = \frac{\mathrm{d}}{\mathrm{d}t} e^{tL} \cdot x \bigg|_{t=0} = L \cdot x$$

for et $L \in \mathfrak{g}$. Kig nu på funktionen

$$\begin{aligned} f_\mu\colon \mathcal{K} &\to \mathbb{R} \\ k &\mapsto p \cdot \delta x \end{aligned}$$

som giver et Hamilton vektorfelt

$$\begin{aligned}\mathcal{V}_{f_\mu} &= \sum_{i=1}^{N} \frac{\partial f_\mu}{\partial p_i}\frac{\partial}{\partial x_i} - \frac{\partial f_\mu}{\partial x_i}\frac{\partial}{\partial p_i} \\ &= \sum_{i=1}^{N} \delta x_i \frac{\partial}{\partial x_i} - p_i \frac{\partial \delta x_i}{\partial x_i}\frac{\partial}{\partial p_i}\end{aligned}$$

Vi vil gerne undersøge hvordan f_μ ændrer sig med tid. Først bemærk at det kanoniske momentum også ændrer sig langs vektorfeltet

$$\delta p_i = \mathcal{V}_{f_\mu}(p_i) = -p_i \frac{\partial \delta x_i}{\partial x_i}\,.$$

Ændringen i Hamiltonianen langs vektorfeltet er

$$\begin{aligned}\mathcal{V}_{f_\mu}(H) &= \sum_{i=1}^{N} \frac{\partial f_\mu}{\partial p_i}\frac{\partial H}{\partial x_i} - \frac{\partial f_\mu}{\partial x_i}\frac{\partial H}{\partial p_i} \\ &= \sum_{i=1} \frac{\partial H}{\partial x_i}\delta x_i + \frac{\partial H}{\partial p_i}\delta p_i \\ &= \delta H \\ &= \frac{\mathrm{d}}{\mathrm{d}t} H(x(t),p(t))\bigg|_{t=0}\end{aligned}$$

Vi er ude efter situationen hvor $\delta H = 0$. Betragt nu følgende co-moment for L,

$$\widehat{\mu}(L) = f_\mu\,.$$

Da er $\mathcal{V}_L = \mathcal{V}_{f_\mu}$ og ligeledes $\mathcal{V}_L(H) = \delta H$. For at se hvordan $\widehat{\mu}(L)$ ændrer sig med tid skal vi tage differentialet. Men husk fra Eksempel 3.19 at dette bare er Hamiltonian vektorfeltet svarende til Hamiltonianen H.

$$\begin{aligned}\frac{\mathrm{d}\widehat{\mu}(L)}{\mathrm{d}t} &= \mathcal{V}_H(\widehat{\mu}(L)) \\ &= \mathcal{V}_H(f_\mu) \\ &= -\mathcal{V}_{f_\mu}(H) \\ &= -\mathcal{V}_L(H) \\ &= -\delta H\end{aligned}$$

Her har vi brugt egenskaberne af co-momentet fra Definition 3.27 samt resultatet fra Opgave 5.24. Så

$$\mathcal{V}_L(H) = 0 \quad \Longleftrightarrow \quad \frac{\mathrm{d}\widehat{\mu}(L)}{\mathrm{d}t} = 0$$

hvilket var det der skulle vises. ■

Sætning 4.2 (**Noethers Sætning (Version 2)**). Lad $\mathcal{V}$ være en symmetri af et system $\mathcal{S} = (H,\mathcal{K})$. Hvis man lokalt kan finde en funktion Q således at $\mathcal{V} = \mathcal{V}_Q$, så er $Q\colon \mathcal{K} \to \mathbb{R}$ konstant langs systemets historie. Ligeledes, hvis Q er konstant så er $\mathcal{V}_Q$ en symmetri.

Det er ikke et tilfælde at vi har brugt symbolet Q (hvilket lige så godt kunne være f_μ fra den tidligere beskrivelse). Inspireret af elektromagnetisme i fysik, hvor den totale ladning af et system altid er bevaret, kalder man ligeledes Q for en **bevaret ladning** eller en **Noether ladning**.

Korollar 4.3. $\mathcal{V}_Q$ er en symmetri hvis og kun hvis

$$\{Q,H\} = 0\,.$$

Se Opgave 5.25 for en definition af Poisson parentesen.

Korollar 4.4. $\mathcal{V}_H$ er en symmetri af systemet $\mathcal{S} = (H,\mathcal{K})$ da

$$\{H,H\} = 0\,.$$

Ved brug af resultatet fra Opgave 5.25 betyder det at

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial t}\,.$$

Hvis nu systemet også har en tidssymmetri $t \mapsto t + \delta t$ som svarer til et vektorfelt $\mathcal{V}_t = \frac{\partial}{\partial t}$, betyder det at $\mathcal{V}_t(H) = 0 \iff \frac{\mathrm{d}H}{\mathrm{d}t} = 0$. Ergo er den totale energi af systemet $H = E$ en bevaret ladning!

$$\text{Tidssymmetri} \quad \longleftrightarrow \quad \text{Energibevarelse}$$

Lad os se på et mere avanceret eksempel.

Definition 4.5. Den specielle ortogonale gruppe $\mathrm{SO}(N)$ beskriver gruppen af rotationer i N dimensioner.[4]

$$\mathrm{SO}(N) = \left\{ A = N \times N \text{ matrix} \mid A^T A = \mathrm{Id}_{N\times N}\,, \det(A) = 1 \right\}$$

Et eksempel på et element $A \in \mathrm{SO}(2)$ er

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

da

$$\begin{aligned} A^T A &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \mathrm{Id}_{2\times 2} \end{aligned}$$

Lemma 4.6. Lie algebraen $\mathfrak{so}(N)$ består af alle matricer L som opfylder

$$L^T + L = 0\,.$$

Bevis. Betragt en kurve $\gamma\colon \mathbb{R} \to \mathrm{SO}(N)$ hvor $\gamma(t) = g$, som opfylder at $\gamma(0) = e = \mathrm{Id}_{N\times N}$. Fra definitionen af $\mathrm{SO}(N)$ ved vi også at $g^{-1} = g^T$. Tydeligvis ændrer identiteten sig ikke,

$$0 = \frac{\mathrm{d}e}{\mathrm{d}t} = \frac{\mathrm{d}\gamma(t)^T \gamma(t)}{\mathrm{d}t}\,.$$

Men nu kan vi kigge på små transformationer langs kurver der går gennem identiteten ved brug af eksponentialafbildningen $\gamma(t) = e^{tL}$,

[4]For os er det ikke vigtigt hvad determinanten "$\det(A)$" betyder, så vi vælger ikke at gå i detaljerne med det.

hvor $L \in \mathfrak{so}(N)$.

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\gamma(t)^T\gamma(t)\bigg|_{t=0} \\ &= \left[\left(\frac{\mathrm{d}\gamma(t)}{\mathrm{d}t}\right)^T \gamma(t) + \gamma(t)^T\left(\frac{\mathrm{d}\gamma(t)}{\mathrm{d}t}\right)\right]\bigg|_{t=0} \\ &= \left[L^T \mathrm{Id}_{N\times N} + \mathrm{Id}_{N\times N} L\right] \\ &= L^T + L \end{aligned}$$

■

Vi har vist at $\mathfrak{so}(N)$ består af anti-symmetriske (eller også kaldet skæv-symmetriske) matricer. Et eksempel på $A \in \mathfrak{so}(3)$ er

$$A = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}. \tag{4.51}$$

Eksempel 4.7. Betragt gruppen af rotationer i tre dimensioner $G = \mathrm{SO}(3)$ og dets virkning på $\mathcal{X} = \mathbb{R}^3$. Den lille transformation er $\delta x = L \cdot x$, hvor $L \in \mathrm{so}(3)$ er et element af den tilsvarende Lie algebra. Der gælder at ethvert sådan L kan skrives på samme form som Ligning 4.51. Vi bemærker at hvis vi skriver $x = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}$ på vektorform, så giver

$$\begin{aligned} \delta x &= \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= \begin{pmatrix} bx_3 - cx_2 \\ cx_1 - ax_3 \\ ax_2 - bx_1 \end{pmatrix} \end{aligned}$$

Det er det præcis det samme som hvis vi beregnede et krydsprodukt

mellem to vektorer![5]

$$\delta x = \ell \times x\,, \qquad \ell \coloneqq \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Derefter kan vi bestemme co-momentet $f_\mu = \widehat{\mu}(L)$,

$$f_\mu(k) = \widehat{\mu}(L)(k) = p \cdot (\ell \times x)\,.$$

Man kan gøre brug af følgende identitet som involverer krydsproduktet mellem vektorer,[6]

$$p \cdot (\ell \times x) = \ell \cdot (x \times p) = x \cdot (p \times \ell)\,.$$

Det betyder også at vi kan aflæse hvordan det kanoniske momentum ændrer sig under gruppevirkningen

$$\begin{aligned} \delta p_i &= \mathcal{V}_{f_\mu}(p_i) \\ &= -\frac{\partial f_\mu}{\partial x_i} \\ &= -\frac{\partial}{\partial x_i}\left(x \cdot (p \times \ell)\right) \\ &= -(p \times \ell)_i \\ &= (\ell \times p)_i \end{aligned}$$

Med andre ord er $\delta p = \ell \times p$ som ligeledes svarer til en rotation af $\mathcal{P}$. Nu kan vi bestemme den bevarede Noether ladning. Husk at co-momentet er en afbildning $\widehat{\mu}(L)\colon \mathcal{K} \to \mathbb{R}$ og momentet $\mu\colon \mathcal{K} \to \mathfrak{so}(3)^*$. Så vi kan identificere

$$\begin{aligned} \widehat{\mu}&\colon L \longmapsto \ell \cdot (\bullet) \\ \widehat{\mu}(L)&\colon (x,p) \longmapsto \ell \cdot (x \times p) \end{aligned}$$

[5]Måden man intuitivt kan forstå det er at ℓ er aksen man roterer om. Alle punkter flytter sig derfor i en retning som er vinkelret på både x og ℓ. Dette er samme retning som $\ell \times x$. Det demonstrerer også hvorfor at $\mathfrak{so}(3) \cong \mathbb{R}^3$.

[6]Årsagen til at det virker er fordi $p \cdot (\ell \times x)$ måler volumenet af et parallelepipedum med sidelængder beskrevet af tre vektorer p,ℓ,x. Hvis man roterer figuren svarer det bare til at bytte de tre sider $p \to \ell \to x \to p$, som selvfølgelig ikke ændrer på volumenet.

og ligeledes

$$\mu\colon (x,p) \longmapsto x \times p$$

Her skal $\mu(k)$ forstås som en afbildning som spiser et $L \in \mathfrak{so}(3)$ (eller i samme stil et $\ell \in \mathbb{R}^3$) og giver et tal, $\mu(k)\colon x \times p \longmapsto \ell \cdot (x \times p)$. Den bevarede ladning er dermed $\mathcal{L} = x \times p$, som jo er impulsmomentet!

$$\begin{array}{c}\text{Rotationssymmetri}\\ x \mapsto g \cdot x\,, \quad g \in \mathrm{SO}(3)\end{array} \quad \longleftrightarrow \quad \begin{array}{c}\text{Impulsmomentsbevarelse}\\ \mathcal{L} = x \times p\,, \quad \dfrac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} = 0\end{array}$$

$\circ$

5 Opgaver

- **Opgave 5.1:**
  Vis, at identitetselementet i en gruppe G er unikt. Det vil sige hvis e_1 og e_2 begge er identitets elementer, så har vi $e_1 = e_2$.

- **Opgave 5.2:**
  Vis, at hvis $g^2 = g \cdot g = g$ for alle elementer af en gruppe G så er $G = \{e\}$ den trivielle gruppe.

•• **Opgave 5.3:**
  Vis, at for gruppen $G = (\mathbb{R},+)$ er $\lambda_a(x) = a + x$ en gruppe virkning på $\mathbb{R}$.

•• **Opgave 5.4:**
  Vis, at formlen for inversen i Eksempel 2.7 er korrekt.

•• **Opgave 5.5:**
  Vis, at afbildningen

$$\begin{aligned} \mathcal{V}_x \colon C^\infty(\mathbb{R}) &\to \mathbb{R} \\ f &\mapsto f'(x) \end{aligned}$$

  er en tangentvektor i punktet $x \in \mathbb{R}$.

- **Opgave 5.6:**

  1) Argumenter at $G = (\mathbb{R},+)$ danner en (Lie) gruppe.

  2) Hvad er dens Lie algebra?

- **Opgave 5.7:**
  Er følgende grupper? Er de Lie grupper? Hvis de er, hvad er deres Lie algebra?

  1) $(\mathbb{R},\times)$

  2) $(\mathbb{Z},+)$

3) $(\mathbb{R}\backslash\{0\},\times)$

4) $(\mathbb{R}^n,+)$

- **Opgave 5.8:**
  Hvor mange frihedsgrader har følgende systemer:

  1) Én partikel i $\mathbb{R}$

  2) Tre partikler i $\mathbb{R}^3$

  3) n partikler i $\mathbb{R}^{10}$

  4) Én partikel på overfladen af en klode

  5) Et pendul

- **Opgave 5.9:**
  Som beskrevet er en historie en kurve i et konfigurationsrum.

  1) Må kurven gerne krydse sig selv? Hvis den gør, hvad er den fysiske fortolkning? Forklar.

  2) Hvad betyder det hvis kurven er lukket? Forklar.

- **Opgave 5.10:**
  Argumenter at det eneste potentiale $V(x)$ i et én-dimensionalt system med translationssymmetri er en konstant.

•• **Opgave 5.11:**
  Betragt systemet som beskriver et massivt legeme spændt fast til en væg med en fjeder (med afstand L). Hamiltonianen er

$$H = \frac{p^2}{2m} + \frac{1}{2}k(x-L)^2\,.$$

  1) Beregn kraften F ud fra potentialet $V(x) = \frac{1}{2}k(x-L)^2$ og vis, at det stemmer overens med Hookes lov.

  2) Vis, at Hamiltons ligninger giver

$$\frac{\partial x}{\partial t} = \frac{p}{m}\,, \qquad \frac{\partial p}{\partial t} = -k(x-L)\,.$$

3) Vis, at følgende funktionsudtryk for $x(t)$ og $p(t)$ er løsninger til ligningerne

$$x(t) = L + A\cos(\omega t)\ , \qquad p(t) = -\frac{kA}{\omega}\sin(\omega t)$$

hvor $A \in \mathbb{R}$.

4) Ved at differentiere $x(t)$ to gange, vis ved brug af Newtons 2. lov at

$$\omega = \sqrt{\frac{k}{m}}\,.$$

5) Beregn

$$\left(\frac{\omega}{k}p(t)\right)^2 + (x(t) - L)^2\ .$$

Hvad beskriver denne type ligning? Brug dette til at tegne systemets historie i dets konfigurationsrum.

6) Hvad sker der hvis man gør A større?

•• **Opgave 5.12:**

En bold bliver kastet vertikalt op i luften. Hamiltonianen som beskriver systemet er

$$H = \frac{p^2}{2m} + mgx\,,$$

hvor større x betyder højere op i luften.

1) Beregn Hamiltons ligninger.

2) Beregn kraften på bolden. Stemmer det overens med hvad du ville forvente?

3) Løsningen til Hamiltons ligninger giver

$$x(t) = -\frac{1}{2}gt^2 + \frac{p_0}{m}t + x_0\,, \qquad p(t) = -mgt + p_0\,.$$

Brug de to ligninger til at finde positionen som funktion af kun impuls, $x(p)$.

4) Tegn boldens historie i et konfigurationsrum. Forklar hvad der sker hvis du ændrer g,m,p_0,x_0.

•• **Opgave 5.13:**
I denne opgave skal vi undersøge et eksempel hvor $H = E$ er energien af systemet, men den er ikke bevaret, samt hvor $H \neq E$ men H er bevaret. Betragt en fjeder fastspændt til en vogn som bevæger sig med konstant hastighed v_0 langs fjederens akse (du kan tænke på det som en dynamisk hviletilstand $L = v_0 t$).

$$H(x,p) = \frac{p^2}{2m} + \frac{1}{2}k(x - v_0 t)^2$$

1) Beregn Hamiltons ligninger.

2) Beregn kraften F. Er den konstant? Stemmer det overens med hvad du ville forvente?

3) Er den totale energi i systemet bevaret? Hvorfor?

4) Nu skifter vi til nye koordinater

$$x' = x - v_0 t\,.$$

Er det en kanonisk transformation? Hvad bliver p'?

5) Det viser sig at den korrekte Hamiltonian i de nye koordinater er

$$H(x',p') = \frac{(p' - mv_0)^2}{2m} + \frac{1}{2}kx'^2 - \frac{1}{2}mv_0^2\,.$$

Er H bevaret? Er det den totale energi?

• **Opgave 5.14:**
Vis, at følgende matrix er element af Sp(2).

$$A = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$$

•• **Opgave 5.15:**
Vis, at følgende matricer er elementer af Sp(2).

$$A = \pm \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix}$$

cosh og sinh er de hyperbolske trigonometriske funktioner. Du kan bruge egenskaben $\cosh^2\theta - \sinh^2\theta = 1$, som minder meget om den såkaldte "idiotformel" $\cos^2\theta + \sin^2\theta = 1$.

••• **Opgave 5.16:**
Betragt en lukket kurve (dvs. en periodisk historie) i konfigurationsrummet $\mathcal{K}$ med konstant energi

$$H(x,p) = \frac{1}{2m}p^2 + V(x) = E \in \mathbb{R}\,.$$

1) Skriv p som funktion af x og E.

2) Argumenter at arealet inden for kurven er

$$A(E) = 2\int_{x_{\min}}^{x_{\max}} p\ \mathrm{d}x\,.$$

Du behøver ikke at evaluere integralet! Det kan være en hjælp at lave en tegning.

3) Substituer nu impulsen med hastigheden $p = mv$. Hvis δt er tiden det tager konfigurationen at bevæge sig en afstand δx med hastigheden v, argumenter at

$$\delta t = \frac{\delta x}{v}\,.$$

Vis derefter at perioden (tiden det tager for systemet at komme tilbage til konfigurationen den startede i) er

$$T(E) = 2\int_{x_{\min}}^{x_{\max}} \frac{1}{v}\ \mathrm{d}x$$

4) Vis, at

$$T(E) = \frac{\mathrm{d}A(E)}{\mathrm{d}E}.$$

•• **Opgave 5.17:**
Tjek at resultatet fra del 4) af Opgave 5.16 stemmer overens med Opgave 5.11.

1) Beregn energien af systemet $E = H(x(t),p(t))$ hvor $x(t)$ og $p(t)$ er løsningerne til Hamiltons ligninger.

2) Beregn arealet af historien og omskriv det som en funktion af E. Du kan bruge at en ellipse beskrevet på den gældende form har følgende areal

$$\frac{(x-x_0)^2}{a^2} + \frac{(y-y_0)^2}{b^2} = 1 \quad \Longleftrightarrow \quad \text{Areal} = \pi ab\,.$$

3) Brug de to forrige resultater til at vise at

$$T = \frac{2\pi}{\omega}\,.$$

Er dette hvad du ville forvente?

•• **Opgave 5.18: Elektrisk kraft**
Betragt følgende Hamiltonian på $\mathcal{K} = \mathbb{R}^2$,

$$H = \frac{1}{2m}\left(p - qA(x)\right)^2 - q\varphi(x)\,,$$

hvor $q \in \mathbb{R}$ og $A(x)$,$\varphi(x)$ er funktioner af position. Den beskriver en partikel med position x og impuls p i et elektromagnetisk felt (det elektriske og magnetiske felt er beskrevet af to funktioner $E(x)$ og $B(x)$).

1) Udled Hamiltons ligninger.

2) Beregn $\ddot{x} \coloneqq \frac{\partial^2 x}{\partial t^2}$ og brug Hamiltons lignigner til at vise at kraften

$$F = m\ddot{x} = q\left(\frac{\partial x}{\partial t}\frac{\partial A}{\partial x} - \frac{\partial A}{\partial t} - \frac{\partial \varphi}{\partial x}\right).$$

3) Brug kædereglen og vis, at de to første led i øverste ligning går ud med hinanden.

4) $\varphi(x)$ kaldes også det elektriske potentiale, som definerer det elektriske felt

$$E(x) = -\frac{\partial \varphi}{\partial x}.$$

Vis, at Newtons lov ender med at være det bekendte

$$F = qE,$$

også kaldet Lorentz kraften af et elektrisk felt.

5) (Svær ekstropgave!) Funktionen $A(x)$ er relateret til det magnetiske felt $B(x)$. Men vi så at $A(x)$ i sidste ende ikke bidragede noget til den totale kraft. Vi ved at Lorentz kraften for et elektrisk *og* magnetisk felt i tre dimensioner (som vi er vant til) er

$$\vec{F} = q\left(\vec{E} + \vec{v} \times \vec{B}\right)$$

hvor $\vec{v}$ er hastigheden af partiklen. Ud fra dette resultat, hvorfor giver det mening at vi ikke ser en "magnetisk kraft" i én dimension? Hvad med to dimensioner?

•●● **Opgave 5.19:**
Udled Hamiltons ligninger for følgende Hamiltonian på $\mathcal{K} = \mathbb{R}^6$,

$$H = (\vec{x} \times \vec{p}) \cdot \vec{\omega},$$

hvor $\vec{\omega} \in \mathbb{R}^3 \backslash \{0\}$. Vis, at systemets historie danner cirkler i $\mathcal{K}$ og at vinkelhastigheden er $|\vec{\omega}|$.

•• **Opgave 5.20: Gauge symmetri**
Lad x og x^* være to frihedsgrader. Vi vil gerne have at vores Hamiltonian har en symmetri

$$x(t) \mapsto e^{i\alpha}x(t) \qquad x^*(t) \mapsto e^{-i\alpha}x^*(t)$$

hvor $i^2 = -1$ og $\alpha \in \mathbb{R}$ er en konstant.

1) Vis, at

$$\frac{\partial x}{\partial t} \mapsto e^{i\alpha}\frac{\partial x}{\partial t} \qquad \frac{\partial x^*}{\partial t} \mapsto e^{i\alpha}\frac{\partial x^*}{\partial t}$$

2) Vis, at $H \mapsto H$ når

$$H = \frac{1}{2}\frac{\partial x^*}{\partial t}\frac{\partial x}{\partial t}$$

Ergo har systemet en symmetri.

3) Lad os nu opgradere α så den ikke længere er konstant, men også kan være tidsafhængig, $\alpha(t)$. Vis, at resultatet fra 1) og 2) ikke længere er sandt (husk kædereglen når man differentierer).

4) Vi kan igen gøre H invariant ved at definere en ny type differentiale.

$$\frac{\mathrm{D}}{\mathrm{D}t} \coloneqq \frac{\partial}{\partial t} - i\frac{\partial \alpha}{\partial t} \qquad \frac{\mathrm{D}^*}{\mathrm{D}t} \coloneqq \frac{\partial}{\partial t} + i\frac{\partial \alpha}{\partial t}$$

Vis, at

$$\frac{\mathrm{D}x}{\mathrm{D}t} \mapsto e^{i\alpha(t)}\frac{\mathrm{D}x}{\mathrm{D}t} \qquad \frac{\mathrm{D}^*x^*}{\mathrm{D}t} \mapsto e^{-i\alpha(t)}\frac{\mathrm{D}^*x^*}{\mathrm{D}t}\,.$$

5) Vis dermed at

$$H = \frac{1}{2}\frac{\mathrm{D}^*x^*}{\mathrm{D}t}\frac{\mathrm{D}x}{\mathrm{D}t}$$

er invariant.

Forklaring: Det vi har beskrevet her er, hvad der sker når man forfremmer en *global* symmetri til en *lokal* symmetri – også kaldet en **gauge symmetri**. I denne opgave har vi arbejdet med en U(1) symmetri, som er den der beskriver elektromagnetisme. "Gauge" i dette

tilfælde hentyder til at funktionen $\alpha(t)$ har samme fysiske virkning på systemet hvis man sender $\alpha(t) \mapsto \alpha(t) + \lambda$, hvor $\lambda \in \mathbb{R}$. Det ligner derfor at der er en ekstra frihedsgrad, λ, som man frit kan vælge. Men den er ikke fysisk. Det nye differentiale bevarer gauge symmetrien og hedder et **covariant differentiale**.

- **Opgave 5.21: Kommutator**
  Lad $k = \begin{pmatrix} x \\ p \end{pmatrix}$ og definer et nyt objekt

$$[A,B] \coloneqq AB - BA$$

  kaldet en **kommutator**.

  1) Vis, at $k^T \cdot \Omega \cdot k = [x,p]$.

  2) Vis, at kanoniske transformationer bevarer kommutatoren. Med andre ord, send $k \mapsto k' = Ak$ og vis, at

$$k'^T \cdot \Omega \cdot k' = k^T \cdot \Omega \cdot k \quad \Longleftrightarrow \quad A \in \mathrm{Sp}(2)\,.$$

- **Opgave 5.22:**
  Find Hamilton vektorfelterne svarende til følgende funktioner og vektorer.

  1) $X(x,p) = x^2 + p^2$.

  2) $Y(x,p) = xp$.

  3) $Z(x_1,x_2,p_1,p_2) = x_1x_2^2 + p_1p_2^2$.

  4) $\vec{V}_U = (1,-1) \in \mathbb{R}^2$.

  5) $\vec{V}_W = (-p,x) \in \mathbb{R}^2$.

  6) For $H = \frac{1}{2}(x^2 + p^2)$, beregn $\mathcal{V}_X(H)$, $\mathcal{V}_Y(H)$ og $\mathcal{V}_U(H)$ fra de forrige opgaver.

•• **Opgave 5.23:**
Lad $X\colon \mathcal{K} \to \mathbb{R}$ være en funktion.

1) Vis, at vektoren

$$\vec{V}_X = \Omega \cdot \nabla X = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial X}{\partial x} \\ \frac{\partial X}{\partial p} \end{pmatrix}$$

giver anledning til Hamilton vektorfeltet $\mathcal{V}_X$.

2) Ud fra det, argumenter hvorfor det giver mening at $\mathcal{V}_H(k)$ giver Hamiltons ligninger.

3) For to funktioner X,Y og Hamilton vektorfelter $\mathcal{V}_X$,$\mathcal{V}_Y$, vis, at

$$\mathcal{V}_X(Y) = (\nabla X)^T \cdot \Omega \cdot (\nabla Y) \ .$$

• **Opgave 5.24:**
Lad X,Y være to funktioner i et konfigurationsrum med tilsvarende Hamilton vektorfelter $\mathcal{V}_X, \mathcal{V}_Y$. Vis, at

$$\mathcal{V}_X(Y) = -\mathcal{V}_Y(X)$$

.

• **Opgave 5.25: Poisson parentes**
Lad X,$Y\colon \mathcal{K} \to \mathbb{R}$ være to funktioner med tilsvarende Hamilton vektorfelter $\mathcal{V}_X$,$\mathcal{V}_Y$. Definer **Poisson parentesen** $\{\bullet,\bullet\}$,

$$\{X,Y\} \coloneqq \sum_{i=1}^{N} \left(\frac{\partial X}{\partial x_i} \frac{\partial Y}{\partial p_i} - \frac{\partial Y}{\partial x_i} \frac{\partial X}{\partial p_i} \right) .$$

1) Vis, at Hamilton vektorfelter også kan skrives som en Poisson parentes,

$$\mathcal{V}_Y(X) = \{X,Y\} \, .$$

2) Skriv Hamiltons ligninger som Poisson parenteser.

3) For $f = f(t,x(t),p(t))$, vis, at

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \{f,H\}\,,$$

hvor H er Hamiltonianen.

4) Lad $X,Y,Z\colon \mathcal{K} \to \mathbb{R}$ være funktioner og $a,b \in \mathbb{R}$. Bevis følgende egenskaber:

- Anti-kommutativitet

$$\{X,Y\} = -\{Y,X\}$$

- Biliniaritet

$$\{aX + bY,Z\} = a\{X,Z\} + b\{Y,Z\}$$
$$\{X,aY + bZ\} = a\{X,Y\} + b\{X,Z\}$$

- Leibniz reglen

$$\{XY,Z\} = \{X,Z\}Y + X\{Y,Z\}$$

- **(Valgfri)** Jacobi identiteten

$$\{X,\{Y,Z\}\} + \{Y,\{Z,X\}\} + \{Z,\{X,Y\}\} = 0$$

5) Vis, at

$$[\mathcal{V}_X,\mathcal{V}_Y](Z) = -\mathcal{V}_{\{X,Y\}}(Z)$$

hvor $[\bullet,\bullet]$ er kommutatoren fra Opgave 5.21.

- **Opgave 5.26:**
Lad $f\colon \mathcal{K} \to \mathbb{R}$ være en funktion og $g_1,g_2 \in G$ være to gruppeelementer. Vis ved brug af Ligning (4.50) at

$$(g_1 \cdot (g_2 \cdot f)))\,(k) = (g_1 g_2) \cdot f(k)\,.$$

- **Opgave 5.27:**
  I denne opgave skal vi vise at Lie algebraen for rotationer i to dimensioner er $\mathfrak{so}(2) \cong \mathbb{R}$.

  1) Gennemgå beviset i Lemma 4.6 og vis, at ethvert element $L \in \mathfrak{so}(2)$ kan skrives på formen

$$L = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}, \qquad a \in \mathbb{R}\,.$$

  2) Argumenter derfor at der eksisterer en (kontinuert) bijektiv afbildning $\mathfrak{so}(2) \to \mathbb{R}$.

  3) (Valgfri) Tjek at eksponentialafbildningen rent faktisk giver rotationsmatricen.

  - Beregn L^2,L^3,...,L^n.
  - Brug nu sumformlem for eksponentialfunktionen til at beregne gruppeelementet.

$$g = e^L = \sum_{n=0}^{\infty} \frac{1}{n!} L^n$$

    Du kan bruge sumformlerne for cosinus og sinus.

$$\cos\theta = \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m)!} \theta^{2m}$$
$$\sin\theta = \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m+1)!} \theta^{2m+1}$$

- **Opgave 5.28:**
  Følgende er en opgave som bygger videre på Opgave 5.27. En anden måde at beskrive rotationer i to dimensioner er ved at oversætte det

til komplekse tal. Vi kan afbilde

$$\mathbb{R}^2 \ni \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto x + iy \in \mathbb{C}\,.$$

1) Ved at rotere vektoren, vis, at man på konsistent vis kan afbilde rotationsmatricen som

$$\mathrm{SO}(2) \ni \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \longmapsto \cos\theta + i\sin\theta =: e^{i\theta} \in \mathbb{C}\,.$$

2) Vis, at $\mathrm{U}(1) \coloneqq (\{e^{i\theta} \mid \theta \in \mathbb{R}\},\times)$ danner en gruppe med multiplikation som kompositionsregel. Det vil sige tjek at

- den er associativ.
- der eksisterer en identitet.
- for hvert element er der en invers.

Denne gruppe kaldes for en **unitære gruppe**.

3) Beregn den lille transformation for et $g = e^{i\theta} \in \mathrm{U}(1)$.

4) Argumenter derfra at $\mathfrak{u}(1) \equiv \mathbb{R}$.

Med andre ord har SO(2) og U(1) samme Lie algebra! Altså ser de ens ud lokalt i nærheden af identitetselementet. Man siger derfor at de to grupper er **lokalt isomorfiske** (det viser sig dog at de også er globalt isomorfiske).

••• Opgave 5.29:

I denne opgave skal du bruge eksponentialafbildningen og kommutatoren fra Opgave 5.21.

1) Vis, at hvis $[A,B] = 0$, så er

$$e^A e^B = e^{A+B}.$$

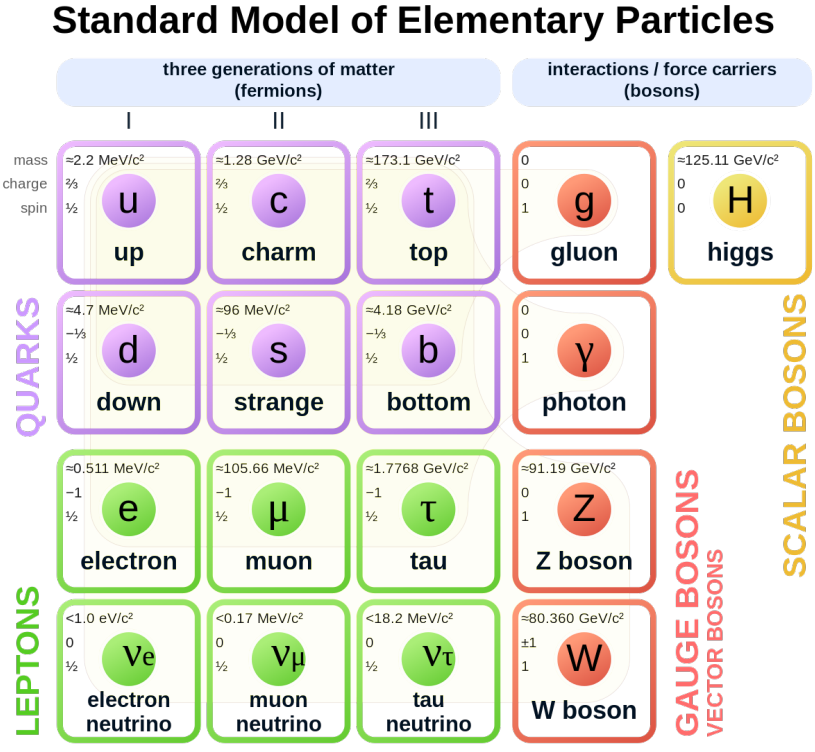
2) Antag at $A,B \in \mathfrak{g}$ nu er elementer af en Lie algebra således at $e^A,e^B \in G$ er elementer af en tilsvarende gruppe (bemærk at G ikke nødvendigvis er unik!). Er det så sandt at $[A,B] = 0$ for alle $A,B \in \mathfrak{g}$ $\Longrightarrow$ $[g_1,g_2] = 0$ for alle $g_1,g_2 \in G$? Forklar.

●●● **Opgave 5.30:**
Lad G være en Lie gruppe der virker på to konfigurationsrum $\mathcal{K}_1$ og $\mathcal{K}_2$, som hver giver anledning til momenter $\mu_1 : \mathcal{K}_1 \mapsto \mathfrak{g}^*$ og $\mu_2 : \mathcal{K}_2 \mapsto \mathfrak{g}^*$. Vis, at momentet $\mu : \mathcal{K}_1 \times \mathcal{K}_2 \mapsto \mathfrak{g}^*$ svarende til den simultane gruppevirkning fra G på $\mathcal{K}_1 \times \mathcal{K}_2$ opfylder

$$\mu(k_1,k_2) = \mu_1(k_1) + \mu_2(k_2)\,.$$

6 Projekt: Antallet af bosoner i Standardmodellen

Moderne fysik består af to primære søjler: Tyngdekraft og partikelfysik. Hver af disse af beskrevet af to teorier – den ene er Einsteins generelle relativitetsteori og den anden er Standardmodellen. Faktisk er det et kæmpe stort problem at vi har to teorier og ikke kun én, som kan beskrive alting (en såkaldt "quantum gravity" teori), men det er en historie til en anden gang. I dette projekt vil vi gerne fokusere på Standardmodellen, som beskriver partikler ved brug af et sprog kaldet kvantefeltteori.



Figur 4.6: Standardmodellen.

Mere specifikt, vil vi gerne diskutere **gauge bosonerne** (også

kaldet vektor bosoner), som er karakteriseret ved at de alle har "spin-1".[7] Den egenskab som de medfører er, at de alle kommunikerer **fundamentalkræfter** mellem partikler. I vores univers har vi fire fundamentale kræfter: Elektromagnetisme, svag kernekraft, stærk kernekraft og tyngdekraft. Men som sagt er tyngdekraft lidt en sjov fætter, så den springer vi over. Tænk nu som eksempel på elektromagnetisme: Hvordan ved én elektron at den skal frastøde og ligeledes frastødes af en anden elektron? Dét der i virkeligheden sker er, at de to elektroner kommunikerer ved hinanden ved at udveksle fotoner, som i sidste ende "skubber" til elektronerne (bid dog mærke at dette er en meget forsimplet forklaring).

I dette projekt vi gerne bruge matematik til at *forudsige* antallet af gauge bosoner i vores univers. Det lyder helt sci-fi, men vi kan rent faktisk gøre det udelukkende ved brug af det vi har lært i løbet af dette forløb. Lad os alligevel være kedelige og give jer svarende på forhånd:

- Der er én partikel som kommunikerer elektromagnetisme, kaldet fotonen, γ. I hverdagstale hedder den også 'lyspartiklen' fordi det er en bølge i det elektromagnetiske felt, som vi opfatter som lys (hvor forskellige farver er forskellige bølgelængder).

- Der findes tre partikler som kommunikerer den svage kernekraft. Disse er W^+, W^- og Z^0 bosonerne. W^+ og W^- har samme masse, men har forskellige elektromagnetisk ladning, på hhv. ± 1. Det betyder også at de interagerer med fotonen! Z^0 bosonen er derimod lidt tungere og har ingen elektromagnetisk ladning.

- Der findes i alt otte partikler som kommunikerer den stærke kernekraft, kaldet gluoner. Men vent, på Standardmodellen står der at der kun er én gluon?? Det er fordi gluoner kommer i

[7] Hvorimod fermionerne (kvarker og leptoner) har spin-$\frac{1}{2}$ og Higgs bosonen har spin-0. Hvis tyngdekraft også havde en beskrivelse som en partikel (kaldet "gravitonen") ville den være spin-2.

forskellige variationer, beskrevet af noget der hedder **farve** (ja, fysikere er opfindsomme med navne). I gluonens verden findes der tre farver (rød, grøn, blå) og tre anti-farver (anti-rød, anti-grøn, anti-blå). I stedet for at farve bogstaverne (det kan være svært at se ordentligt), så skriver vi r for rød, g for grøn og b for blå. Anti-farverne skriver vi med en streg over, så det bliver $\overline{r}$ for anti-rød, $\overline{g}$ for anti-grøn og $\overline{b}$ for anti-blå. Hver gluon kan have én farve og én anti-farve. Med lidt kombinatorik kan du tælle, at der er 9 kombinationer. Så hvordan får vi 8? Det viser sig, at gluoner i virkeligheden er kombinationer af forskellige farvekombinationer:

$$\begin{array}{cc} \frac{1}{\sqrt{2}}\left(r\overline{b}+b\overline{r}\right) & \frac{1}{\sqrt{2}}\left(r\overline{g}+g\overline{r}\right) \\ \frac{1}{\sqrt{2}}\left(b\overline{g}+g\overline{b}\right) & \frac{1}{\sqrt{2}}\left(r\overline{r}-b\overline{b}\right) \\ \frac{i}{\sqrt{2}}\left(b\overline{r}-r\overline{b}\right) & \frac{i}{\sqrt{2}}\left(g\overline{r}-r\overline{g}\right) \\ \frac{i}{\sqrt{2}}\left(g\overline{b}-b\overline{g}\right) & \frac{1}{\sqrt{6}}\left(r\overline{r}+b\overline{b}-2g\overline{g}\right) \end{array}$$

Der findes endnu en kombination som ikke kan skrives som en kombination af de ovenstående! Det er det såkaldte "farveløse" kombination

$$\frac{1}{\sqrt{3}}\left(r\overline{r}+b\overline{b}+g\overline{g}\right) .$$

Den viser sig slet ikke at interagere med nogle af de andre gluoner. Derudover, fordi vi ikke har observeret en gluon som opfører sig på denne måde, er der god grund til at tro at den ikke eksisterer. Derfor er der kun 8 gluoner – og det kan vi rent faktisk også forudsige, hvis vi startede med at tænke på gruppeteori!

Hvordan hænger gauge bosoner sammen med hvad vi har lært om symmetrier? Før du fortsætter med dette projekt kan det være en god idé at lave eller genlæse Opgave 5.20. Idéen er at vi har et system

$\mathcal{S}$ hvor der lever nogle kvantefelter, som er afbildninger $\varphi\colon \mathcal{X} \to \mathbb{C}$ som overholder noget struktur.[8] Denne struktur er, at der findes en transformation

$$\varphi(x) \longmapsto M(x)\varphi(x)$$

som bevarer dynamikken.[9] Husk at $\varphi(x)$ generelt kan beskrives som en vektor. Derfor er $M(x)$ en matrix. Fordi vi vil kigge på transformationer som bevarer en vis struktur, må de forskellige $M(x)$ derfor danne en gruppe – det er en symmetri!

Fra Lie grupper og Lie algebra til kvantefelter

Denne del af projektet giver lidt kontekst til, hvorfor kvantefelter har noget at gøre med Lie grupper og Lie algebraer. Hvis du vil springe direkte til at vise hvordan man får tallene 1, 3 og 8 som beskrevet tidligere, kan du springe disse opgaver over eller endda hoppe direkte til næste del på Side 284.

Som opvarmning, lad os undersøge hvilke forskellige slags udtryk det er som forbliver det samme når vi laver transformationen. Fra nu af vil vi forkorte og blot skrive "$\partial\varphi$" for at beskrive det partiale differentiale af φ med hensyn til et eller andet (vi antager også at det her 'et eller andet' er reelt). Derudover definerer vi en ny operation, †, som er en kombination af at transponere $A \mapsto A^T$ og tage den kompleks konjugerede $i \mapsto -i$ (hvor $i \coloneqq \sqrt{-1}$) af noget. Med andre ord $A^\dagger = (A^T)^* = (A^*)^T$, hvor rækkefølgen ikke betyder noget. Nogle

[8] Fx findes der et Higgsfelt $\mathcal{H}$ som eksisterer i hele universet. Når vi taler om Higgs *partiklen* mener vi i virkeligheden at der er en lokaliseret bølge i feltet. Ligesom hvis du kigger under en trampolin kan du se ud fra hvordan den bølger, hvor der står en person henne.

[9] Typisk i kvantefeltteori bruger vi ikke Hamiltonianen og Hamiltons ligninger til at beskrive dynamikken. I stedet bruger man et lignende objekt, kaldet **Lagrangianen**, $\mathcal{L}$, som kan relateres til Hamiltonianen. Dynamikken deraf kommer via Euler-Lagrange ligningerne, som er tilsvarende Hamiltons ligninger.

eksempler:

$$\begin{aligned}(x+iy)^\dagger &= x-iy\,, \qquad x,y\in\mathbb{R}\\ (AB)^\dagger &= \left[(AB)^T\right]^* = \left[B^TA^T\right]^* = B^\dagger A^\dagger\\ (\partial A)^\dagger &= \partial(A^\dagger)\\ (iA)^\dagger &= -iA^\dagger\end{aligned}$$

- **Opgave 6.1: (Valgfri)**

1) Argumenter at

$$\partial\varphi \longmapsto (\partial M)\varphi + M\partial\varphi\,.$$

Hvad er $(\partial\varphi)^\dagger$?

2) Vis at $(\partial\varphi)^\dagger(\partial\varphi) \not\mapsto (\partial\varphi)^\dagger(\partial\varphi)$.

Ligesom i Opgave 5.20 skal vi introducere et nyt form for differentiale "D", som vi kalder det **covariante differentiale** således at $(\mathrm{D}\varphi)^\dagger(\mathrm{D}\varphi) \mapsto (\mathrm{D}\varphi)^\dagger(\mathrm{D}\varphi)$ er en symmetri.

- **Opgave 6.2: (Valgfri)**

1) Definer nu det covariante differentiale:

$$\mathrm{D}\varphi := \partial\varphi + icA\varphi$$

Vi vil gerne forstå A som at have noget at gøre med en partikel/kvantefelt. Men helt hvordan ved vi ikke endnu! Indtil videre kan du dog forestille dig at produktet "$A\varphi$" betyder at de to partikler taler til hinanden. Tallet $c\in\mathbb{R}\setminus\{0\}$ kan derfor fysisk forstås som en koblingskonstant. Men matematisk skal du umiddelbart tænke på A som en matrix, ligesom vi har set det tidligere.

Vi vil gerne have at

$$\varphi\mapsto M\varphi\,, A\mapsto A' \quad\Longrightarrow\quad \mathrm{D}\varphi\mapsto M(\mathrm{D}\varphi).$$

Vis at ovenstående gælder hvis

$$icA'M\varphi = icMA - \partial M\,. \tag{4.52}$$

Alt hvad vi har gjort indtil videre kan i princippet generaliseres til enhver form for transformation. Men dem som vi er interesserede i her, i konteksten af partikelfysik, er såkaldte **unitære transformationer**. Vi så eksemplet på U(1) i Opgave 5.20 og 5.28, hvor $M \in \mathbb{C}$ bare er et komplekst tal (bemærk at det også betyder at $M^\dagger = M^*$).

Definition 6.1. Den **unitære gruppe** $\mathrm{U}(N)$ består af $N \times N$ matricer med indgange i $\mathbb{C}$, hvor $M^{-1} = M^\dagger$, $\forall M \in \mathrm{U}(N)$.

$$\mathrm{U}(N) \coloneqq \left\{ M = N \times N \text{ matrix af komplekse tal} \mid M^{-1} = M^\dagger \right\}$$

Der findes en undergruppe, kaldet den **specielle unitære gruppe**, $\mathrm{SU}(N)$, hvor elementer også opfylder $\det(M) = 1$, hvor "det" er matricens determinant.

Vi kommer ikke til at bruge determinanten til meget her, så der er ingen grund til at bruge tid på at definere det.

•• **Opgave 6.3: (Valgfri)**

1) Lad $M \in \mathrm{U}(N)$. Ved at gange Ligning (4.52) med $-\frac{i}{c}M^\dagger$ fra højre side, vis at

$$A \mapsto A' = MAM^\dagger + \frac{i}{c}\left(\partial M\right) M^\dagger\,.$$

2) Tænk nu på $M \in G$ som værende et element af en gruppe. Vi vil gerne se hvad der sker hvis vi laver små transformationer. Så vi kan skrive $M = 1 + iL$, hvor $L \in \mathfrak{g}$ er et element af Lie algebraen. Bemærk at vi har en faktor i, hvilken er konvention som man ofte bruger inden for partikelfysik.

Vis at

$$A \mapsto A' = A + i[L,A] - \frac{1}{g}\partial L \tag{4.53}$$

hvor $[L,A] \coloneqq LA - AL$ kaldes for **kommutatoren** af L og A, som vi også brugte i Opgave 5.21. Husk at du kan ignorere termer såsom "L^2" fordi det er små transformationer!

NU kommer punchlinen! Vi introducerede A som en måde at opgradere $\partial \to \mathrm{D}$. Men indtil videre har vi ikke anet hvad A overhovedet er. Nu kan vi endelig stille os selv spørgsmålet:

"Hvad er A? Hvilken mængde tilhører den?"

Kan du huske helt tilbage til Definition 2.11 af vores Lie algebra? Nemlig, at $[\cdot,\cdot]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ er en operation der tager to elementer af en Lie algebra og sender dem til et nye element. Altså er Ligning (4.53) fuldkomment konsistent hvis $A \in \mathfrak{g}$!

Gauge bosonerne, vis interaktioner med andre partikler manifesteres gennem det kovariante differentiale, lever i Lie algebraen af symmetrigrupper! Standardmodellen er symmetrisk under $\mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)$. Det kovariante differentiale er derfor

$$\mathrm{D} = \partial + ic_1 G + ic_2 A + ic_2 Y$$

hvor de forskellige gauge bosoner lever i følgende Lie algebraer:

Gluon: $G \in \mathfrak{su}(3)$

W/Z: $A \in \mathfrak{su}(2)$

Foton: $Y \in \mathfrak{u}(1)$

Bemærkning 6.2. Det faktum at der er tre grupper betyder også at der er tre fundamentalkræfter. Man kan måske spørge sig selv, hvorfor tre og ikke bare én? Det ville jo være meget nemmere og

smukkere! Denne idé kaldes for "grand unification". Idéen om at forene de tre naturlove i én samlet kraft, er derfor et spørgsmål om at finde en Lie gruppe der kan husere $\mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)$. To famøse eksempler er $\mathrm{SU}(5)$ og $\mathrm{SO}(10)$ som begge har Standardmodellen som en undergruppe.

$$\mathrm{SO}(10) \supset \mathrm{SU}(5) \supset \mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)$$

Grand unified theory var rigtig populært tilbage i 70'erne, men da vi endnu ikke har set noget direkte eksperimentel evidens for det endnu, er det ret dødt i dag. Løsningen efterlader vi som en opgave til læseren.

Generatorer af Lie algebraer og antallet af bosoner

Vi har etableret at gauge bosoner faktisk bare danner Lie algebraen af symmetrigrupperne i partikelfysik. Med andre ord er der en én-til-én korrespondence mellem partikler og symmetrier – wow! Men vi har endnu ikke fundet ud af *hvor mange* der er. Så lad os minde os selv om, hvad en Lie algebra er: Det er de "små transformationer" af en Lie gruppe. Men vigtigere for os i det her tilfælde er det faktum, at det er et **tangentrum**. Nemlig, i Definition 2.10 sagde vi at $\mathfrak{g} = T_e G$ er tangentrummet gennem identiteten. Idéen med et tangentrum er, at den fortæller dig alle de retninger man kan bevæge sig i mangfoldigheden. Hvor mange retninger er der?

Definition 6.3. Lad $T_p\mathcal{M}$ være et tangentrum vis elementer er tangentvektorer $\mathcal{V}_p \in T_p\mathcal{M}$. En **basis** er en mængde $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$, hvor $\boldsymbol{v}_i \in T_p\mathcal{M}$, således at alle $\mathcal{V}_p \in T_p\mathcal{M}$ kan skrives på formen

$$\mathcal{V}_p = a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n\,, \quad a_i \in \mathbb{R}\,.$$

Hvis $\mathcal{M} = G$ er en gruppe og $T_p\mathcal{M} = \mathfrak{g}$ er dens Lie algebra, kaldes basiselementerne $\boldsymbol{v}_i = \xi_i$ for **generatorer**.

I gruppeteori kalder vi specifik basiselementerne af Lie algebraen for generatorer fordi de *genererer* små transformationer. Skriv fx

$$g = 1 + i\varepsilon L$$

for en lille transformation, hvor $g \in G$ og $L \in \mathfrak{g}$. Nu kan vi vælge at skrive L ud fra vores basis.

$$g = 1 + i\varepsilon a_1 \xi_1 + i\varepsilon a_2 \xi_2 + \cdots + i\varepsilon a_n \xi_n$$

Så vi kan altid skille små transformationer ad og se dem i flere dele! Der findes derfor n distinkte "retninger" man kan transformere systemet i og alle transformationer kan skrives som en linearkombination af disse retninger.

Definition 6.4. Dimensionen af et tangentrum, $\dim(T_p\mathcal{M})$, er det mindste tal n hvor der findes en basis med n elementer.

En anden måde at se dimensionen er ved at stille sig selv spørgsmålet:

> "*Hvor mange (reelle) tal skal jeg angive for at beskrive et element?*"

Eksempel 6.5. Lad os bestemme dimensionen af $\mathfrak{so}(2)$. Vi ved at elementerne af gruppen SO(2) er rotationsmatricer

$$\mathrm{SO}(2) \ni \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} . \tag{4.54}$$

I Opgave 5.27 fandt vi ud af at alle elementer $L \in \mathfrak{so}(2)$ kan skrives på formen

$$L = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix} , \qquad a \in \mathbb{R} .$$

Nu kan vi nemt finde generatorerne. Den består nemlig af ét element

$$\xi = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

således at alle Lie algebra elementer kan skrives på formen $L = a\xi$. Da det ikke giver mening af en basis er nul elementer og ét er det mindste tal vi derefter kan have, har vi vist at $\dim(\mathfrak{so}(2)) = 1$. ∘

- **Opgave 6.4:**
Betragt Lie gruppen $G = (\mathbb{R}^n,+)$. Hvad er dens Lie algebra? Find en basis. Hvad er dimensionen, $\dim(\mathfrak{g})$?

- **Opgave 6.5:**
Det viser sig at Lie algebraen $\mathfrak{gl}(N,\mathbb{C})$ består af alle $N \times N$ matricer med komplekse tal. Argumenter at $\dim(\mathfrak{gl}(N,\mathbb{C})) = 2N^2$.

•• **Opgave 6.6:**
I denne opgave skal vi bestemme Lie algebraen af U(N) på samme måde som vi gjorde i Lemma 4.6.

1) Gennemgå beviset i Lemma 4.6 og brug samme metode til at vise at

$$L^\dagger + L = 0\,, \qquad \forall L \in \mathfrak{u}(N)\,.$$

2) Hvad er $\mathfrak{u}(1)$? Vis at $\dim(\mathfrak{u}(1)) = 1$.

Du har nu vist at der findes 1 foton.

3) For $N = 2$, vis at ethvert $L \in \mathfrak{u}(2)$ kan skrives på formen

$$L = \begin{pmatrix} ia & -c+id \\ c+id & ib \end{pmatrix}\,. \tag{4.55}$$

4) Vis at $\dim(\mathfrak{u}(2)) = 4$. Hvad er generatorerne?

5) Vis at $\dim(\mathfrak{u}(3)) = 9$.

6) (Valgfri) Vis at $\dim(\mathfrak{u}(N)) = N^2$.

Hvordan springer vi fra $\mathfrak{u}(N)$ til $\mathfrak{su}(N)$? S'et i "SU" eller "SO" står for "speciel" og betyder at elementerne også har determinant $\det(g) = 1$.

Hvad betyder det for Lie algebra elementerne? Desværre er determinanten et koncept som vi ikke har brugt tid på at diskutere. Hvad der følger må I desværre bare tro på eller spørge en af underviserne om. Det er muligt at vise følgende formel for matricer:

$$\det(g) = e^{\mathrm{tr}(g)}\,.$$

Her står $\mathrm{tr}(g)$ for **trace**, hvilket er at man tager summen af alle de diagonale indgange i matricen. For eksempel har matricen i Ligning (4.54) trace $\cos^2\theta$ og de andre matricer i samme eksempel har $\mathrm{tr}(L) = \mathrm{tr}(\xi) = 0$. Så hvis $\det(g) = 1$ for elementer $g \in \mathrm{SO}(N)$ eller $g \in \mathrm{SU}(N)$, må det betyde for Lie algebra elementerne at

$$1 = \det(g) = e^{\mathrm{tr}(L)} \quad \Longrightarrow \quad \mathrm{tr}(L) = 0\,.$$

Ergo har Lie algebra elementerne nul trace! Hvad betyder det for dimensionen af Lie algebraen? Summen af diagonalen kan skrives som en ligning

$$a_1 + a_2 + \cdots + a_N \overset{!}{=} 0\,.$$

Men vi kan altid vælge at skrive fx a_N som funktion af $a_1, \ldots, a_{N-1}$. Det betyder at der altid er ét mindre tal som skal angives! Du kan måske sammenligne det med når du skal løse to ligninger med to ubekendte. Den ene ligning kan du nemlig bruge til at eliminere ét af de variable. Hvis du derimod havde én ligning med to ubekendte, ville du altid skulle angive ét variabel![10] For at opsummere:

$$\dim(\mathfrak{su}(N)) = \dim(\mathfrak{u}(N)) - 1$$

Bemærk dog at elementer af $\mathfrak{o}(N)$ altid har nul trace fordi $L^T + L = 0$ medfører at alle indgange på diagonalen er nul. Så faktisk $\dim(\mathfrak{so}(N)) = \dim(\mathfrak{o}(N))$.

[10] Én ligning med to ubekendte definerer i virkeligheden bare formlen for en linje γ. En linje har dimension $\dim(\gamma) = 1$. Derfra kan du måske forestille dig hvordan ubekendte hænger sammen med dimension.

- **Opgave 6.7:**
  Argumenter at $\dim(\mathfrak{su}(2)) = 3$ og $\dim(\mathfrak{su}(3)) = 8$.

  Du har nu vist at der findes 3 W/Z bosoner og 8 gluoner.

  Generatorerne af $\mathfrak{su}(2)$ kaldes for **Pauli matricerne** og for $\mathfrak{su}(3)$ er det **Gell-Mann matricerne**.

Tillykke! Du har nu brugt dine matematiske superkræfter til at forudsige antallet af forskellige gauge bosoner i universet, uden overhovedet at lave noget som helst fysik! Det viser bare hvor stærkt et værktøj symmetri kan være og hvilke konsekvenser det kan have for systemer. Lad os opsummere hvad vi har lært i dette projekt:

Hvorfor er der 1 foton, 3 W/Z bosoner og 8 gluoner?

- Standardmodellen er symmetrisk under gruppevirkningen af $G = \mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)$.
- Vi har set at hvis dynamik skal være invariant, så skal man opgradere alle partielle differentialer til kovariante differentialer $\partial \to \mathrm{D} = \partial + icA$, hvor A kan forstås som en partikel der interagerer med noget andet og vis styrke bestemmes af koblingskonstanten c.
- Vi har vist at A nødvendigvis må være elementer af Lie algebraen!
- Ethvert Lie algebra element kan skrives som en lineær kombination af basiselementer, kaldet generatorer. Dimensionen er antallet af generatorer.
- Vi har vist at dimensionerne af $\mathfrak{su}(3)$, $\mathfrak{su}(2)$ og $\mathfrak{u}(1)$ er 8, 3 og 1.
- Derfor må der findes 8 SU(3)-partikler (gluoner), 3 SU(2)-partikler (W/Z) og 1 U(1)-partikel (fotonen).

Indeks

Bibliografi

[1] Johannes Ueberberg. *Foundations of Incidence Geometry. Projective and Polar Spaces*. Springer, 2011. ISBN: 978-3-642-20971-0.

[2] Lynn Margaret Batten. *Combinatorics of finite geometries*. 2. udg. Cambridge University Press, 1997. ISBN: 9780521599931.

[3] Kenneth H. Rosen. *Discrete Mathematics and Its Applications*. 8. udg. McGraw-Hill Education. ISBN: 978-1-260-09199-6.

[4] User Kmhkmh at Wikimedia. *Paul Erdos at a student seminar in Budapest (fall 1992)*. Distribueret under Creative Commons Attribution 3.0 Unported licensen. For at se en kopi af licensen, besøg `https://creativecommons.org/licenses/by/3.0/deed.en`. 1992. URL: `https://commons.wikimedia.org/wiki/File:Erdos_budapest_fall_1992.jpg` (hentet 28.04.2024).

[5] Wikipedia. *Paul Erdos*. 2024. URL: `https://en.wikipedia.org/wiki/Paul_Erd%C5%91s` (hentet 28.04.2024).

[6] G. Eric Moorhouse. *Incidence Geometry*. 2007. URL: `https://math.ucr.edu/home/baez/qg-fall2016/incidence_geometry.pdf`.

[7] Albrecht Beutelspracher og Ute Rosenbaum. *Projective Geometry. From Foundations to Applications*. Cambridge University Press, 1998.

[8] David S. Dummit og Richard M. Foote. *Abstract Algebra*. 3. udg. John Wiley & Sons, Inc., 2004. ISBN: 0-471-43334-9.

[9] M. F. Atiyah og I. G. Macdonald. *Introduction to Commutative Algebra*. 2. udg. Addison-Wesley Publishing Company, 1969.

[10] Victor Shoup. *A Computational Introduction to Number Theory and Algebra*. 2. udg. Cambridge University Press, 2008. URL: `https://www.shoup.net/ntb/`.

[11] Keith Conrad. *The Gaussian Integers*. URL: `https://kconrad.math.uconn.edu/blurbs/ugradnumthy/Zinotes.pdf`.