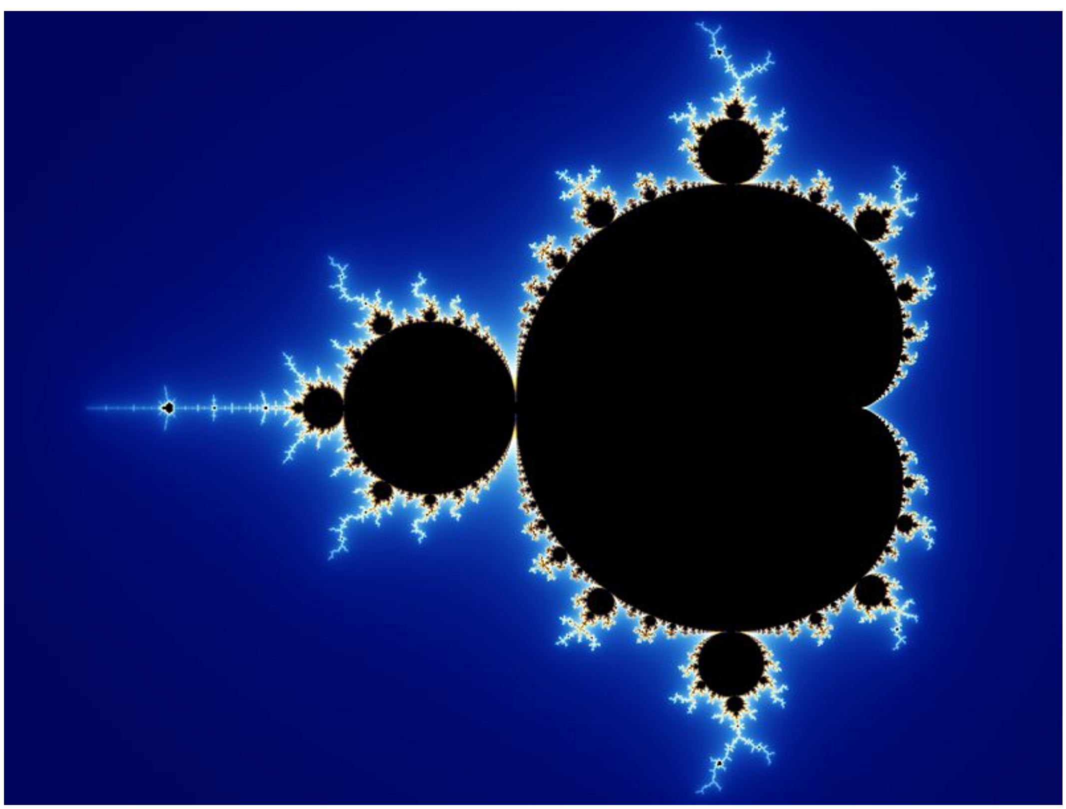
$$\frac{1}{\epsilon}$$

“Nei, så stor du er blitt!”

En innføring i tall, følger og kontinuitet for matematikkstudenter på universitetet

Andreas Leopold Knutsen

Versjon August 20, 2021



Innhold

Forord

Velkommen til et studium i matematikk ved Universitetet i Bergen, enten dere nå er meldt på studieprogrammet *Matematikk* eller *Matematikk for industri og teknologi*. Velkommen også til dere som har valgt å følge kurset *MAT100–Innføringsemne i matematikk* av andre grunner.

Matematikk er et svært viktig fag som står i en særstilling blant alle realfagene, fordi det brukes i alle realfag, i tillegg til andre vitenskaper som medisin og økonomi. Andre fag er derfor helt avhengige av matematikk og matematikkens språk og evne til å modellere problemer fra virkeligheten. Dette kommer spesielt til syne ved at de fleste matematikkurs dere tar i første studieår (og til og med i tredje semester) er felles for alle realfagsstudenter. Dette gjelder kjeden av "kalkuluskurs" MAT111-MAT112-MAT212, der dere lærer teorien om reelle funksjoner av én og flere variable, med fokus på differensial- og integralregning, kurset MAT121, der dere lærer *lineær algebra*, studiet av lineære ligningssystemer og vektorrom, og kurset MAT131 der dere lærer om *differensialligninger*. Til og med fjerdesemesterkurset MAT213, der dere lærer om *komplekse funksjoner* (funksjoner definert på og med verdier i de *komplekse tallene*), er tilrettelagt for at studenter fra alle realfag skal kunne følge det. Dette viser den store betydningen matematikk har i naturvitenskapen.

Samtidig er matematikk et eget fagfelt, som er uavhengig av andre realfag, med sine egne problemstillinger som driver fagfeltet fremover, og sin egen tenkemåte. Hensikten med kurset *MAT100–Innføringsemne i matematikk*, og denne boken som danner grunnlaget for pensum i kurset, er å gå mer i dybden på forskjellige matematiske begreper som gjennomgås på en noe mer overfladisk måte i de ovennevnte kursene, for å gi dere et stødigere grunnlag for videre studier i matematikk. Vi kommer til å fokusere mindre på prosedyrer og drilloppgaver, og mer på bakgrunnen og forutsetningene for resultatene. Vi kommer til å fokusere på *hvorfor* resultater er sanne, og *hvordan* de brukes i den videre logiske oppbyggingen av matematikken. Ett av hovedmålene er å lære dere å *lese* matematikk og å *tenke* matematisk.

Å lære seg den matematiske tenkemåten er ikke helt det samme som å *gjøre* matematikk, ihvertfall ikke på den måten matematikk er presentert i skolen. Skolematematikken fokuserer typisk på innlæring av prosedyrer for å løse stereotypiske problemer. Nøkkelen til suksess er å tenke lineært innenfor rammer gitt av bestemte prosedyrer, ikke å være spesielt fantasifull. Profesjonelle matematikere tenker på en spesiell måte for å løse problemer, enten de oppstår i andre fagfelt eller fra matematikken selv. En nøkkel i matematisk tenking er å tenke utenfor innlærde prosedyrer, en ferdighet som for øvrig er en verdifull egenskap i dagens verden.

Innholdet i boken. Hovedmålene med denne boken er

- å gi en innføring i grunnleggende logikk som brukes i matematiske resonnementer, samt i grunnleggende matematiske begreper;
- å vise konstruksjonen av heltallene, de rasjonale tallene, de reelle tallene og de komplekse tallene – spesielt sentralt står konstruksjonen av de reelle tallene og hvilken rolle tallfølger spiller i denne;
- å gi en innføring i begrepet kontinuitet og dets relasjon til tallfølger;
- å gi en smakebit av generaliseringer av de ovennevnte tallmengdene, som euklidske rom, metriske rom og topologiske rom, for å gjøre dere bedre rustet når dere møter disse senere i studiet.

Boken er ikke ment kun til bruk i kurset MAT100, men ment å kunne følge dere og være til hjelp i de første 3-4 semestrene av studiet.

Vi gir nå en mer presis beskrivelse av innholdet i de forskjellige kapitlene.

Kapittel 1 gjennomgår grunnleggende logikk som trengs i matematiske resonnementer og noen av de vanligste bevistypene i matematiske tekster. Kapitlet avsluttes med en seksjon om noen av de mest kjente problemene i matematikk, løste og uløste, som forhåpentligvis kan virke som motivasjon for den videre lesingen.

Kapittel 2 omhandler begrepene mengder og funksjoner. En god del av dette vil være kjent fra skolen, men stoffet om *injektive, surjektive, bijektive* og *inverse funksjoner* i §2.3, samt om *relasjoner* i §2.4, vil være helt nytt og spille viktige roller senere i teksten og i flere matematikkurs dere tar senere.

Felles for de første to kapitlene er at vi vil ta for gitt eksistensen av de naturlige tallene, heltallene, de rasjonale tallene og de reelle tallene, som vi har gjort på skolen. Dette gjør vi for å være i stand til å gi eksempler fra stoffet vi presenterer. Fra og med Kapittel 3 nullstiller vi oss imidlertid og ser på oppbyggingen av disse tallmengdene helt fra begynnelsen av, uten å anta noe som helst, annet enn den elementære logikken og mengdelæren vi har introdusert i de to første kapitlene.

I *Kapittel 3* starter vi med å bygge opp de naturlige tallene aksiomatisk, samt regneoperasjonene addisjon og multiplikasjon, for så å utvide disse tallene og operasjonene til heltallene og de rasjonale tallene. Disse tallsystemene er spesialtilfeller av mer generelle strukturer i *algebra* som kalles *ringer* og *kropper*, og disse gjennomgås mer generelt i §3.6. Vi tar også med en seksjon om *primtall* (§3.7) og om *binomialteoremet* (§3.8), siden dette er grunnleggende materiale i matematikk.

Kapittel 4 gir noen flere eksempler på ringer og kropper og er således et lite sidesprang inn i *algebra*. Spesielt viktig for senere bruk er stoffet om *polynomringer i én variabel* i §4.1. Et interessant eksempel er også *ringen av heltall modulo m* i §4.3, som har viktige anvendelser innenfor for eksempel kodeteori og det norske personnummersystemet.

Kapittel 5 inneholder en gjennomgang av *velordningsprinsippet* og *induksjonsprinsippet*, to prinsipper og bevisteknikker som henger sammen med ett av aksiomene i definisjonen av de naturlige tallene, nemlig *induksjonsaksiomet*. Siden dette er bevisteknikker som brukes i mange forskjellige deler av matematikk, har vi funnet det naturlig å skille dette stoffet ut i et eget kapittel, etter innføringen av de rasjonale tallene, for å gi flere eksempler på bruk av disse prinsippene.

Kapittel 6 forklarer hva som "mangler" i de rasjonale tallene. Intuitivt har de rasjonale tallene "hull", siden de "mangler" for eksempel $\sqrt{2}$. Matematisk kan dette visualiseres ved hjelp av følger av rasjonale tall som nærmer seg mer og mer et "hull" i den rasjonale tallinjen. Sentrale begreper for å forstå dette skikkelig er begrepene *øvre skranker* og *minste øvre skranker* (§6.1) og *følger* av rasjonale tall (§6.2-6.3).

Kapittel 7 er kanskje høydepunktet i boken: inspirert av manglene i de rasjonale tallene vi studerte i Kapittel 6, vil vi her gjennomgå én av konstruksjonene av de reelle tallene i §7.1. Hovedideen er å tette igjen hullene i de rasjonale tallene ved å legge til grenser av følger av rasjonale tall. Resten av kapitlet utleder alle sentrale egenskaper til de reelle tallene som vi kjenner fra skolen basert på denne konstruksjonen, blant annet desimaltallsrepresentasjonen (§7.8). Vi vil også utlede flere viktige egenskaper og resultater til reelle følger i §7.4-7.7. Dessuten vil vi se, i §7.9, at det finnes forskjellige typer uendeligheter i matematikk: selv om de rasjonale tallene og de irrasjonale tallene (det vil si, de reelle tallene som ikke er rasjonale) er uendelig mange, finnes det mange flere irrasjonale tall enn rasjonale tall. Vi avslutter kapitlet med en seksjon om *algebraiske tall*, som er reelle tall som er løsninger på ligninger med heltallskoeffisienter (som for eksempel $\sqrt{2}$) og et appendiks som viser en alternativ konstruksjon av de reelle tallene.

Kapittel 8 viser hvordan man aksiomatisk definerer de komplekse tallene fra de reelle, samt motivasjonen bak dette, og hvordan man utleder de viktigste egenskapene til de komplekse tallene. Stoffet i §8.1-8.4 vil stort sett være repetisjon av pensum i MAT111. Stoffet i §8.5 om *følger av komplekse tall* er imidlertid nytt, samt beviset for det viktige *fundamentalteoremet i algebra* i §8.6.

Kapittel 9 gjennomgår det sentrale begrepet *kontinuitet*. I motsetning til i kurset MAT111, vil alle resultatene bevises skikkelig, ikke minst de viktige resultatene *skjæringssetningen* (§9.6) og *ekstremalverdisetningen* (§9.7), og flere av deres anvendelser. Nytt i forhold til MAT111 er også sammenhengen mellom kontinuitet og konvergens av følger i §9.4. Vi går gjennom *fikspunktiterasjon* i §9.8, som er en av de sentrale numeriske metodene for å finne tilnærmede verdier på løsninger på ligninger. Her spiller teorien om følger vi har utviklet tidligere en viktig rolle. Vi tar også med en seksjon om *uniform kontinuitet* (§9.9), som er en litt sterkere egenskap enn vanlig kontinuitet, samt en seksjon om *grenseverdier* (§9.10), som er sterkt relatert til kontinuitet og som stort sett er repetisjon av stoff fra MAT111.

Kapittel 10 viser hvordan begreper og resultater i de reelle og komplekse tallene som omhandler avstander aksiomatisk kan overføres til mengder hvor vi kan definere et slikt avstandsbegrep. Slike mengder kalles *metriske rom*, og et viktig spesialtilfelle er *euklidske rom*. Sistnevnte spiller en sentral rolle i kursene *MAT112–Grunnkurs i matematikk II* og *MAT212–Funksjoner av flere variable*. Vi vil se på konvergens av følger og kontinuitet i disse rommene (§10.2–10.4), med spesielt fokus på euklidske rom, samt generaliseringer av *skjæringssetningen* og *ekstremalverdisetningen* (§10.5–10.6). Vi gir flere eksempler på metriske rom foruten euklidske rom i §10.7–10.9. Deretter avslutter vi med en liten seksjon om topologiske rom (§10.10), som er en ytterligere generalisering av metriske rom.

Om seksjoner merket med *. Stoffet i disse seksjonene blir (for det meste) ikke brukt i resten av boken og kan derfor utelates dersom en ønsker en kortere variant av pensum. Ønsker man en enda kortere variant, kan for eksempel hele Kapittel 8 om komplekse tall eller resten av Kapittel 10 om euklidske og metriske rom utelates.

Stoffet i noen av de *-merkede seksjonene vil likevel kunne kaste nytt lys over resten av stoffet og virke motiverende for lesingen. Det gjelder spesielt stoffet om *ringen av heltall modulo m* (§4.3) og *fikspunktiterasjon* (§9.8). Beviset for *fundamentalteoremet i algebra* (§8.6) viser fine anvendelser av det man har lært tidligere og de *-merkede seksjonene i §10 gir fine smakebiter på dypere matematikk.

Om oppgavene. Oppgavene er samlet mot slutten av hvert kapittel. I motsetning til i skolematematikken, hvor oppgavene er laget for å drille inn teknikker, fokuserer oppgavene i denne boken på bakgrunnen for forskjellige matematiske resultater og resonnementer og har som mål å gi økt forståelse for den matematiske teksten. De fleste oppgavene går derfor ut på å fylle ut detaljer i bevis og eksempler i teksten, eller å studere eksempler på spesielt uventede fenomener. I motsetning til det dere er vant til fra skolen vil dere derfor komme til å oppleve å ikke umiddelbart vite hva dere skal gjøre og å ikke bare kunne “anvende en regel” for å komme til svaret. Dere kommer dessuten til å føle at oppgavene er mye vanskeligere enn dere har vært vant til fra skolen. Dette er midlertid bare sunt, for det er slik den virkelige verdenen er, enten i matematikk eller andre fag: det er ytterst sjelden at problemer løses med en rett frem prosedyre.

I tråd med ånden i boken og i kurset, som er å lære seg matematisk tenkemåte og å lese og forstå matematikk, finnes det ingen fasit eller løsningsforslag til oppgavene. Dette er også i tråd med den virkelige verdenen. Siden hensikten med oppgavene ikke er å komme frem til et svar, men øke forståelsen og få dere til å tenke, er ikke en fasit nødvendig.

Tilbakemeldinger. Et første utkast av denne boken har blitt først brukt høsten 2020. Dagens versjon er en forhåpentligvis mye bedre versjon, hvor flere feil og uklarheter har blitt rettet opp og en del forklaringer og oppgaver har blitt lagt til. Jeg har gjort mitt beste for å korrekturlese og luke bort trykkfeil og andre feil, men det finnes sikkert fremdeles en del av dem. Jeg setter pris på at disse signaliseres til meg og mottar også alle former for tilbakemeldinger med stor takk, på `andreas.knutsen@math.uib.no`.

Takk. Takk til undervisningsassistent René Langøen for mange rettelser og nyttige tilbakemeldinger. Takk til Sigurd Yrke Bleie, Morten Brun, Herman Mostein, Sebastian Merkesdal Oen, Marthe Ulmo Rønneseth, Edvard Raddum Skaug og Sigmund Selberg for å ha meldt inn feil. Takk også til Silke Lekaus for mange gode råd og for å ha laget en del av figurene.

Jeg takker også alle deltagerne på forelesningene høsten 2020: Aleksander, Edvard, Erjona, Ingrid, Johanne, Julie, Karen, Marthe, Sigurd og Sebastian, for å ha gjort kurset så trivelig å undervise!

Bergen, 16. august 2021,
Andreas Leopold Knutsen

"ET VANSKELIG MATEMATIKKPROBLEM" (1948), AV DEN RUSSISKE MALEREN MIKHAIL LIKHACHYOV (1919-1997)

KAPITTEL 1

Matematikkens språk og resonnementer

Som de fleste andre fagfelt har matematikk sitt eget språk og sin egen notasjon. Selv om matematikk brukes i et mangfold av andre fagfelt, spesielt innenfor naturvitenskap, medisin og økonomi, er matematikk et fagfelt i seg selv, hvor en viktig del består av å utlede konklusjoner ut fra hypoteser og antagelser gjennom logiske slutninger.

Vi skal i dette kapitlet gå gjennom en del grunnleggende logikk som brukes i matematikk (§1.1), så vel som de grunnleggende bevistypene som brukes i matematiske resonnementer (§1.3). Vi avslutter kapitlet med en uformell gjennomgang av berømte matematiske problemer som har spilt en viktig rolle i utviklingen av matematikken (§1.4).

1.1. Implikasjoner

En god del matematiske resonnementer inneholder utsagn av typen,

(1) "Hvis A, så B",

hvor A og B er påstander eller utsagn. Disse kan i utgangspunktet hver for seg være sanne eller ikke, det tar vi ikke stilling til når vi hevder (1). Det vi hevder (i matematikk) er at hver gang A er sann, så er også B sann. Utsagnet A kalles *hypotesen*[1], mens utsagnet B kalles *konklusjonen*.

Et eksempel er utsagnet om reelle tall:

(2) "Hvis $x > 0$, da er $x^3 > 0$."

Her er hypotesen A utsagnet "$x > 0$", mens konklusjonen B er utsagnet "$x^3 > 0$". Utsagnet (2) hevder at hver gang et reelt tall x oppfyller $x > 0$, da oppfyller det også $x^3 > 0$. Det er viktig å være klar over at et slikt utsagn i matematikk ikke betyr at tredjepotensen av *ett bestemt* positivt tall er et positivt tall, men at trediepotensen av *et hvilket som helst* positivt tall er et positivt tall.

Et annet eksempel er utsagnet om heltall:

(3) "Hvis n er delelig på 6, da er n et partall."

Her er hypotesen "n er delelig på 6" og konklusjonen er "n er partall". Utsagnet hevder altså at et hvilket som helst heltall delelig på 6 er et partall.

[1] *Hypothesis* på engelsk.

Merk at utsagn av typen (1) i matematikk gjerne har en litt annen betydning enn i dagligtalen. Når man i dagligtalen sier:

(4) "Hvis du ikke er snill, får du ikke is",

vil et barn oppfatte dette som det samme som utsagnet

(5) "Hvis du er snill, får du is".

I matematikk og logikk sier imidlertid (4) kun at hvis barnet *ikke* er snilt, vil det som konsekvens ikke få is, og sier absolutt ingenting om hva som skjer hvis det motsatte av hypotesen holder, nemlig at barnet er snilt. Dette på samme måte som at utsagnet (2) ikke sier noe om hva konklusjonen er dersom $x < 0$.

Et annet eksempel på at logikk og matematikk er litt annerledes enn dagligdags språk er følgende utsagn om reelle tall:

(6) "Hvis $x^2 < 0$, da er $x = 21$."

Dette er igjen et utsagn av typen (1). Det kommer kanskje som en overraskelse at utsagnet (6) er helt logisk og matematisk riktig: hver gang et reelt tall x oppfyller *hypotesen* "$x^2 < 0$", da holder *konklusjonen* "$x = 21$". Det er imidlertid slik at hypotesen aldri er oppfylt, men det tar vi ikke stilling til. Utsagnet (6) er et eksempel på et utsagn som er "tomt sant"[2], fordi hypotesen aldri er sann.

Andre ekvivalente uttrykksmåter for (1) er

(7) "B, hvis A"

(8) "A medfører/impliserer[3] B"

(9) "A er en tilstrekkelig betingelse[4] for B"

(10) "B er en nødvendig betingelse[5] for A"

(11) "A bare hvis B"

(12) "B følger fra/av A"

(13) "$A \Longrightarrow B$".

Symbolet "$\Longrightarrow$" kalles en *implikasjonspil*. Et slikt symbol, som flere andre, er vanligere å bruke i egne notater og på tavlen under en forelesning, der man gjerne av tidsmessige og plassmessige årsaker forsøker å komprimere teksten, mens det er ikke fullt så vanlig i en lærebok eller en matematisk artikkel. Siden denne boken inneholder en del tankegang av den mer utforskende typen i forhold til hva som er vanlig i en lærebok, kommer vi imidlertid til å bruke slike symboler en del ganger.

[2] *Vacuously true* på engelsk.
[3] *Implies* på engelsk.
[4] *Sufficient condition* på engelsk.
[5] *Necessary condition* på engelsk.

Et ord bør kanskje sies om utsagnene (10)-(11). Siden (1) hevder at B alltid er sann dersom A er sann, betyr det at B må nødvendigvis være sann for at A skal være sann, hvilket ligger til grunn for uttrykksmåten i (10), eller at A bare kan være sann dersom B er sann, som ligger til grunn for uttrykksmåten i (11). Med denne uttrykksmåten kan (2) uttrykkes ekvivalent som

(14) "En nødvendig betingelse for at $x^3 > 0$ er at $x > 0$",

og (3) som

(15) "En nødvendig betingelse for at n er delelig på 6 er at n er partall".

Kontrapositive utsagn. Vi ser at (10) betyr at dersom B ikke er sann, så kan vi konkludere at A ikke er sann. Med andre ord har vi at det opprinnelige utsagnet (1) er ekvivalent med utsagnet

(16) "Hvis ikke-B, så ikke-A",

som kalles *det kontrapositive*[6] *utsagnet til* (1). Her kalles gjerne "ikke-A" for *negasjonen* av A, og tilsvarende for B. For de spesielt interesserte kan nevnes at symbolet man bruker for dette i matematikk og logikk gjerne er "$\neg$".

For å skrive ned det kontrapositive utsagnet til (2) må vi altså skrive hva negasjonen av hypotesen og konklusjonen er. Negasjonen av hypotesen "$x > 0$" er "$x \leq 0$", mens negasjonen av konklusjonen "$x^3 > 0$" er "$x^3 \leq 0$". Det kontrapositive utsagnet til (2) er da

(17) "Hvis $x^3 \leq 0$, da er $x \leq 0$".

Tilsvarende er det kontrapositive utsagnet til (3) lik

(18) "Hvis n er et oddetall, da er n ikke delelig på 6".

Det viktige med kontrapositive utsagn er at siden de er ekvivalent med det opprinnelige utsagnet kan vi likegodt vise at et utsagn er sant ved å vise at det kontrapositive utsagnet er sant, noe som en del ganger er lettere. Dette er bakgrunnen for såkalte *kontrapositive bevis*, som vi skal se på i §1.3.

Motsatte utsagn. Implikasjonen i motsatt retning av (13):

(19) $$B \Longrightarrow A,$$

som vi også kan uttrykke som

(20) $$A \Longleftarrow B,$$

betyr det samme som "Hvis B, så A". Dette sier selvsagt noe annet enn "$A \Rightarrow B$", og har faktisk veldig lite med det opprinnelige utsagnet (1) å gjøre. Utsagnet (19) kalles gjerne *det motsatte*[7] *av* $A \Rightarrow B$. For eksempel er det motsatte utsagnet av (2) lik "Hvis $x^3 > 0$, er $x > 0$" (som er sant!),

[6] *Contrapositive* på engelsk.
[7] *Converse* på engelsk.

mens det motsatte av (3) er lik "Hvis n er partall, så er n delelig på 6" (som generelt er usant!).

Ekvivalente utsagn. I de tilfellene hvor begge implikasjonspilene "$A \Rightarrow B$" og "$B \Rightarrow A$" er sanne (som i tilfellet med utsagnet (2)), sier vi at *A og B er ekvivalente*[8] *(utsagn)* og betegner dette med en dobbel implikasjonspil

$$A \iff B; \tag{21}$$

dette uttrykkes også som

$$\text{"}A \text{ hvis og bare hvis } B\text{"}, \tag{22}$$

som ikke har noe motstykke i det dagligdagse språk. Noen ganger presiserer vi at "$A \Rightarrow B$" er "bare hvis-delen" av utsagnet, mens "$A \Leftarrow B$" er "hvis-delen" av utsagnet. Uttrykksmåter av denne typen er svært vanlige i matematikk, slik at det er viktig å bite seg merke i disse. Eksempler vi kjenner til fra skolen er

$$\sin x = 0 \iff x = n\pi \text{ for et heltall } n$$

og

"En trekant er likebent hvis og bare hvis to av vinklene er like store."

1.2. Resultater og antagelser

Matematiske tekster, både artikler publisert i tidsskrifter og lærebøker, inneholder, for å si det litt enkelt, resultater som er utledet ved logiske resonnementer ut fra antagelser. Disse har forskjellige typer navn.

Et *teorem*[9] er et matematisk resultat som blir regnet som et av de viktigste i teksten, et hovedresultat i teksten, som gjerne blir brukt om og om igjen.

En *setning*[10], noen ganger også kalt *sats*[11] er et viktig resultat som godt kan stå på egne ben, men som ikke regnes som like sentralt, ihvertfall av forfatteren. Her er det viktig å være klar over at en del svært viktige teoremer i matematikk, som har fått betegnelsen "teorem" i andre språk, har fått betegnelsen "setning" på norsk[12], slik som "skjæringssetningen" (Teorem 9.6.1) og "ekstremalverdisetningen" (Teorem 9.7.2), som vi skal møte på senere.

---

[8] *Equivalent* på engelsk.

[9] *Theorem* på engelsk. Ordet kommer fra gammelgresk *theórema*, som betyr "læresetning".

[10] *Proposition* på engelsk.

[11] Dette ordet kommer fra det tyske ordet *Satz* eller *Lehrsatz*, som betyr "læresetning".

[12] Dette skyldes at matematisk terminologi på norsk stort sett kommer fra tysk, hvor begrepene *Satz* og *Lehrsatz* i utgangspunktet ble brukt for alle viktige resultater, uavhengige om de var hovedresultater eller ikke.

En *hjelpesetning*[13] (eller *lemma*) er et resultat som ifølge forfatteren ikke har egeninteresse, men er et resultat på vei mot noe større.

Et *korollar*[14] (eller *følgesetning*) er en mer eller mindre umiddelbar konsekvens av et tildigere resultat. Her kan man i noen tilfeller stusse på hva "umiddelbar" betyr; for en ikke så erfaren leser kan nok en del utregninger ikke fremstå som umiddelbare, men for en matematiker blir ofte et resultat som kan utledes fra et annet resultat ved en lang, men standard utregning betraktet som en umiddelbar konsekvens.

Felles for alle disse resultat-typene er at de må begrunnes, eller som vi kaller det i matematikk (og logikk), *bevises*. Bevisene bygger på tidligere resultater og såkalte *aksiomer* eller *postulater*, som er utsagn vi aksepterer uten bevis som grunnlag for teorien. Eksempler er de såkalte *kroppsaksiomene*[15] vi vil møte på i Definisjon 3.6.1, som rasjonale og reelle tall oppfyller, og Euklids postulater for geometri. Matematiske tekster vil også inneholde *definisjoner*, som man bruker til å slå fast viktig notasjon og viktige begreper og uttrykk som vil brukes senere. Ulike forfattere kan operere med forskjellige definisjoner, hvilket gjør at man som leser må være ekstra oppmerksom, mens i mange tilfeller er imidlertid definisjonene såpass sentrale at alle forfattere bruker samme definisjon.

Merk også at i en definisjon vil "hvis" eller "dersom" faktisk uttrykke "hvis og bare hvis" eller en ekvivalens av utsagn. Som et eksempel la oss se på definisjonen av et partall:

Definisjon 1.2.1: Partall

Et heltall m er et *partall*[a] dersom det finnes et heltall k slik at $m = 2k$.

[a] *Even number* på engelsk.

Dette uttrykker at *per definisjon* gjelder:

$$m \text{ er et partall} \iff \text{det finnes et heltall } k \text{ slik at } m = 2k.$$

I egne notater, eller på tavlen i en forelesning, vil man ofte uttrykke et slikt ekvivalenstegn som definerer den ene siden med en

$$\overset{\text{def}}{\iff}.$$

La oss for helhetens skyld også skrive ned definisjonen av et oddetall:

Definisjon 1.2.2: Oddetall

Et heltall m er et *oddetall*[a] dersom det finnes et heltall k slik at $m = 2k + 1$.

[a] *Odd number* på engelsk.

[13] *Lemma* på engelsk. Ordet kommer fra gammelgresk.

[14] *Corollary* på engelsk. Ordet kommer fra latin *corollarium*, som betyr "konsekvens" eller "deduksjon".

[15] *Field axioms* på engelsk.

Dette uttrykker at *per definisjon* gjelder:

$$m \text{ er et oddetall} \iff \text{det finnes et heltall } k \text{ slik at } m = 2k + 1.$$

MERKNAD 1.2.3. Ethvert heltall er enten partall eller oddetall. Dette er kjent fra skolematematikken og virker intuitivt opplagt. Ser vi på Definisjonene 1.2.1 og 1.2.2 betyr dette at alle heltall kan skrives på formen $2k$ eller $2k + 1$. For øyeblikket vil vi bare godta dette. Når vi har gått gjennom den rigorøse definisjonen av heltall, vil vi imidlertid kunne gi et skikkelig bevis for dette i Oppgave 5.12.

1.3. Bevistyper

Et veldig viktig kjennetegn ved matematikk er bevis[16] for alle påstander. Her skiller matematikk seg ut fra de andre realfag (bortsett fra teoretiske deler av informatikk, som har mye til felles med matematikk), hvor man kan bruke eksperimentelle observasjoner til å "begrunne" påstander. I matematikk godtas ikke en påstand før den er bevist. Det holder ikke å sannsynliggjøre påstanden eller argumentere at ingen har klart å finne et moteksempel, man må vise at ingen moteksempler finnes.

Hvorfor må man lære seg å lese bevis, ihvertfall som matematikkstudent, men helst også som realfagsstudent generelt? Holder det ikke å bare godta alle resultater, siden mange andre har sjekket bevisene før oss? Her er svaret nei, av den enkle grunn at bevis forteller oss ikke bare *at* et resultat er sant, men også *hvorfor* et resultat er sant. Skjønner man ihvertfall hovedtrekkene i hvorfor noe er sant, så kan man mye enklere organisere kunnskapen slik at man i etterkant benytter den riktig. Alternativet er å pugge utenat et sammensurium av løsrevne påstander, noe som ikke lenger fungerer når fagstoffet blir omfattende og krevende, hvilket det blir på universitetsnivå.

Hva er så et bevis? Egentlig er det ikke annet enn en kjede av logiske (og fornuftige) slutninger trukket fra noe man allerede har bevist til det man ønsker å vise. Man kunne også rett og slett kalle et bevis for en begrunnelse, men ofte av en mye lengre sort enn dem man møter på i dagliglivet.

Vi skal nå se på en del av de viktigste typene bevis.

Direkte bevis. De fleste resultater i matematikk dreier seg om en *konklusjon* (som gjerne kan være en liste av flere konklusjoner), basert på *hypoteser*, det vil si utsagn man tar utgangspunkt i er sanne. I den forstand er de fleste resultater av typen "$A \Rightarrow B$", hvor A og B gjerne kan bestå

[16] *Proof* på engelsk. Dette er også tittelen på et veldig interessant skuespill om en matematiker, hans datter og et matematisk bevis av den amerikanske dramatikeren David Auburn (1969–), som vant både *Pulitzer Prize for Drama* og *Tony Award for Best Play* i 2001, og har blitt filmatisert under samme tittel i 2005 av John Madden, med Gwyneth Paltrow, Anthony Hopkins, Jake Gyllenhaal og Hope Davis i hovedrollene. Trailer kan ses på https://www.youtube.com/watch?v=DI1vRch4omw.

av flere utsagn. En vanlig bevistype er det såkalte direkte bevis, der man tar utgangspunkt i utsagnene i hypotesen, og deduserer konklusjonen ved hjelp av flere steg, som hvert skal være matematisk korrekt. La oss se på et eksempel.

Setning 1.3.1

Hvis n er et partall, da er n^2 et partall.

Bevis
Hypotesen i setningen er at "n er et partall". Definisjon 1.2.1 gir da at $n = 2k$ for et heltall k. Siden vi ønsker å konkludere at n^2 er et partall, bruker vi opplysningen vi har om n til å beregne n^2. Vi har

$$n^2 = (2k)^2 = 4k^2 = 2(2k^2).$$

Dette betyr at vi kan skrive n^2 som $n^2 = 2l$ med $l = 2k^2$ et heltall. Per definisjon av partall, betyr dette at n^2 er et partall, som er det vi skulle bevise. $\square$

Utsagnet i siste setning vil også kunne uttrykkes uten matematiske symboler som: "Kvadratet av et partall er et partall". Som nevnt i §1.1, betyr ikke dette at kvadratet av *ett bestemt* oddetall er et oddetall, men at kvadratet av *et hvilket som helst* oddetall er et oddetall. Skal man bevise utsagnet, er det altså ikke tilstrekkelig å velge et oddetall, si 1, og vise at $1^2 = 1$, et oddetall!

Merk at vi avsluttet beviset med tegnet "$\square$", noe som er vanlig i matematiske tekster. En annen vanlig måte å avslutte bevis på er med uttrykket "Q.E.D.", forkortelse for *quod erat demonstrandum*, latin[17] for "hvilket skulle bevises".

La oss se på ett eksempel til på et direkte bevis. Vi minner om at et *rasjonalt tall* er et tall r som kan skrives som $r = \frac{m}{n}$, hvor m og n er heltall, og $n \neq 0$. (Vi skal kommer nærmere inn på definisjonen av rasjonale tall i §3.4.)

Setning 1.3.2

Summen av to rasjonale tall er et rasjonalt tall.

Bevis
Som nevnt ovenfor, sier ikke utsagnet i setningen at summen av to *spesifikke* rasjonale tall igjen er rasjonalt, men at summen to *hvilke som helst* rasjonale tall er et rasjonalt tall. Vi kan uttrykke dette som: *Hvis r_1 og r_2 er rasjonale tall, da er $r_1 + r_2$ et rasjonalt tall.* Hypotesen er altså at "r_1 og r_2 er rasjonale tall", og konklusjonen (som er det vi skal vise) er at "$r_1 + r_2$ er et rasjonalt tall".

[17] Uttrykket kommer imidlertid fra de gamle greske matematikerne og ble oversatt til latin og brukt av mange matematikere fra 1600-tallet.

Som i forrige bevis starter vi med hypotesen: per definisjon av rasjonale tall, vet vi at $r_1 = \frac{m_1}{n_1}$ og $r_2 = \frac{m_2}{n_2}$, hvor m_1, m_2, n_1, n_2 er heltall, og $n_1 \neq 0$ og $n_2 \neq 0$. Da er

$$r_1 + r_2 = \frac{m_1}{n_1} + \frac{m_2}{n_2} = \frac{m_1 n_2 + m_2 n_1}{n_1 n_2}.$$

Siden både $m_1 n_2 + m_2 n_1$ og $n_1 n_2$ er heltall, og $n_1 n_2 \neq 0$ (siden $n_1 \neq 0$ og $n_2 \neq 0$), kan vi konkludere at også $r_1 + r_2$ er et rasjonalt tall, som er det vi skulle vise. □

Bevis som ikke er direkte bevis, kalles gjerne *indirekte bevis*. Vi skal se nærmere på et par av disse, nemlig *kontrapositive bevis* og *motsigelsesbevis* (eller *bevis ved (selv)motsigelse*).

Kontrapositivt bevis. Det er ikke alltid at direkte bevis fungerer så bra for å vise resultater av typen "$A \Rightarrow B$". Vi kan gjerne starte med utsagnet eller utsagnene i hypotesen A, men finne ut at det vi klarer å dedusere ut fra hypotesen, ikke virker å føre oss nærmere konklusjonen.

La oss se på et eksempel:

Setning 1.3.3

Hvis n^2 er et partall, da er n et partall.

Bevis (Forsøk med direkte bevis)

Hypotesen i setningen er at "n^2 er et partall". Definisjon 1.2.1 gir da at $n^2 = 2k$ for et heltall k. Vi ønsker å konkludere at n er et partall. Men nå sitter vi litt fast, for hvordan kan vi bruke at $n^2 = 2k$ til å vise at n er et partall, det vil si at n kan skrives som $n = 2l$ for et heltall l? Den umiddelbare konsekvensen at $n = \pm\sqrt{2k}$ er ikke så nyttig.......... □

I slike tilfeller er det lurt å huske at utsagnet "$A \Rightarrow B$" er logisk ekvivalent med det *kontrapositive utsagnet* "ikke-$B \Rightarrow$ ikke-A", hvor "ikke-A" og "ikke-B" er negasjonene av A og B, henholdsvis.

Å utføre et *kontrapositivt bevis*[18] betyr at vi istedenfor å bruke et direkte bevis på utsagnet "$A \Rightarrow B$", heller bruker et direkte bevis på utsagnet "ikke-$B \Rightarrow$ ikke-A".

La oss se på hvordan dette fungerer på forrige eksempel.

Bevis (kontrapositivt bevis) for Setning 1.3.3

Hypotesen i setningen er at "n^2 er et partall", som har negasjon "n^2 er et oddetall" (jf. Merknad 1.2.3). Konklusjonen i setningen er "n er et partall", som har negasjon "n er et oddetall". Det kontrapositive utsagnet er altså:

$$n \text{ er et oddetall} \implies n^2 \text{ er et oddetall.}$$

[18] *Proof by contraposition* på engelsk.

For å vise dette, antar vi at n er et oddetall, det vil si $n = 2k + 1$ for et heltall k ved Definisjon 1.2.2. Da er

$$n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1.$$

Altså kan vi skrive $n^2 = 2l + 1$, med $l = 2k^2 + 2k$, som viser at n^2 er et oddetall (igjen ved Definisjon 1.2.2). Vi har altså vist det kontrapositive utsagnet, og dermed setningen. □

La oss se på et annet eksempel:

Setning 1.3.4

La a, b og n være positive heltall. Hvis $n \geq ab$, da er $a \leq \sqrt{n}$ eller $b \leq \sqrt{n}$.

Igjen ser vi at det ikke finnes noen opplagt måte å konkludere at $a \leq \sqrt{n}$ eller $b \leq \sqrt{n}$ ut fra ulikheten $n \geq ab$, så vi forsøker et kontrapositivt bevis.

Bevis (kontrapositivt bevis) for Setning 1.3.4

Hypotesen er "$n \geq ab$", som har negasjon "$n < ab$". Konklusjonen er "$a \leq \sqrt{n}$ *eller* $b \leq \sqrt{n}$", som har negasjon "$a > \sqrt{n}$ *og* $b > \sqrt{n}$". Vi vil vise at

$$a > \sqrt{n} \text{ og } b > \sqrt{n} \implies ab > n.$$

Dette er enkelt: Hvis både $a > \sqrt{n}$ og $b > \sqrt{n}$, da er $ab > \sqrt{n} \cdot \sqrt{n} = n$, og beviset er fullført. □

I matematiske tekster vil man ofte ikke nevne hva slags type bevis man utfører, fordi man regner med at (den erfarne) leseren vil forstå dette. Et bevis for siste setning vil da gjerne bare bestå av den siste setningen.

Bevis ved selvmotsigelse. Anta at vi ønsker å vise at en påstand B er sann (gjerne ved å anta en del hypoteser A som ovenfor). En veldig nyttig fremgangsmåte er å anta at B er usann, og vise at dette medfører en umulighet, en *selvmotsigelse*. Denne umuligheten er noe som enten er opplagt feil, som at $1 = 0$, eller en negasjon av en av hypotesene. Denne teknikken kalles *motsigelsesbevis* eller *bevis ved (selv)motsigelse*[19]. Den har også beholdt det latinske navnet *reductio ad absurdum* (som kan oversettes direkte med "reduksjon til det absurde") i logikk og matematikk, som man kan vifte med for å virke intellektuell! Metoden er ofte effektiv, fordi vi får en ekstra hypotese å jobbe med, nemlig at B er usann (med andre ord *negasjonen* til B), i tillegg til alle andre hypoteser og resultater vi vet fra før.

Et berømt eksempel på et slikt bevis er det som viser at $\sqrt{2}$ *ikke* er et rasjonalt tall, noe vi vil få bruk for flere ganger senere i denne teksten. Reelle tall som ikke er rasjonale, kalles *irrasjonale tall*.

[19] *Proof by contradiction* på engelsk.

Setning 1.3.5

Det finnes ikke noe rasjonalt tall x som oppfyller $x^2 = 2$. Med andre ord: $\sqrt{2}$ er et irrasjonalt tall.

Bevis

Vi argumenterer ved selvmotsigelse og antar at påstanden vi skal vise er usann, nemlig at det finnes et rasjonalt tall x slik at $x^2 = 2$. Vi kan skrive $x = \frac{m}{n}$, der m og n er heltall, og $n \neq 0$. Vi kan videre anta at brøken $\frac{m}{n}$ er forkortet mest mulig, slik at m og n ikke har felles heltallsfaktorer. (Vi skal senere, i Teorem 3.7.7, se at alle heltall kan faktoriseres på en entydig måte i produkt av primtall.) Vi har

$$2 = x^2 = \frac{m^2}{n^2},$$

og dermed

$$2n^2 = m^2. \tag{23}$$

Ved definisjonen av et partall, ser vi at m^2 må være et partall. Da må også m være et partall, ved Setning 1.3.3. Per definisjon betyr det at $m = 2k$ for et heltall k. Dermed kan vi skrive (23) om til

$$2n^2 = 4k^2.$$

Deler vi begge sider på 2, får vi

$$n^2 = 2k^2.$$

Igjen ved definisjonen av partall, betyr dette at n^2 er et partall, som igjen medfører at n er et partall ved Setning 1.3.3.

Vi har altså vist at både m og n er partall. Dette er en selvmotsigelse, fordi vi startet med to heltall m og n som *ikke* har felles faktorer, og nå har vi vist at 2 er en felles faktor. Antagelsen om at x er rasjonalt kan følgelig ikke være sann. Dermed har vi bevist at et rasjonalt tall x slik at $x^2 = 2$ ikke kan eksistere. □

Vi tar med ett eksempel til på et bevis ved selvmotsigelse:

Setning 1.3.6

Av et vilkårlig utvalg av 50 personer, må minst 5 ha bursdag i samme måned.

Bevis

Anta at konklusjonen er usann. Da kan høyst 4 personer ha bursdag i samme måned. Siden det finnes 12 måneder, kan vi derfor høyst ha $4 \cdot 12 = 48$ personer, som er en motsigelse, siden vi vet at antallet personer er 50. □

Dette siste beviset kan egentlig like godt ses på som et kontrapositivt bevis. Setningen kan nemlig omskrives til utsagnet

$$\begin{aligned}&\text{En mengde består av 50 personer}\\ \Longrightarrow\quad &\text{minst 5 personer i mengden har bursdag i samme måned,}\end{aligned}$$

og det kontrapositive utsagnet er da

$$\begin{aligned}&\text{I en mengde personer har høyst 4 bursdag i samme måned}\\ &\Longrightarrow \text{mengden består ikke av 50 personer}\end{aligned}$$

Ser vi på beviset på forrige setning, så har vi nettopp bevist det kontrapositive utsagnet.

1.4. Feil, åpne problemer og formodninger

Selv om matematikk som fagfelt kjennetegnes av kravet om rigorøse bevis for alle påstander, betyr dette ikke at faget ikke inneholder rom for fantasifulle spekulasjoner, feil og misforståelser. Bevis man leser i artikler og bøker er gjerne skrevet på den kortest mulige måten og forteller ikke noe om den lange, kronglete veien matematikerne har brukt for å komme seg dit. Langs denne veien vil selv den erfarne matematikeren gjøre feil og i mange tilfeller hoppe over detaljer som han eller hun tror kan bevises for å se hva neste fjelltopp er. Men til slutt må alle detaljene bevises for at resultatet skal bli akseptert av alle matematiske kolleger og bli publisert. Likevel er historien full av eksempler der resultater blir publisert og akseptert som riktige, før de mange år etterpå blir falsifisert ved at noen peker ut et moteksempel eller en feil i beviset. Det siste betyr ikke at resultatet ikke er riktig, men at man må betrakte det som et "åpent problem", siden beviset ikke er helt riktig. Funnet av et moteksempel er imidlertid mye mer dramatisk: det betyr at resultatet er feil og kan i beste fall kun "reddes" ved å endre på hypotesene.

I andre tilfeller kan matematikere, etter å ha undersøkt mange spesialtilfeller, være temmelig sikre på at et resultat holder, uten å være i stand til å bevise resultatet. Da vil de gjerne fremme en *formodning*[20] som sier at "jeg er temmelig sikker på at resultatet holder, men kan ikke bevise det", i håp om at andre matematikere vil interessere seg for problemet og klare å bevise det, gjerne i en eller annen modifisert form.

Mye av fremgangen innenfor matematikk er knyttet til åpne problemer og formodninger. La oss se nærmere på noen her. Andre formodninger vi vil møte på i teksten er *formodningen om Fermat-tall* (Eksempel 5.2.13), *Goldbachs formodning* (Formodning 3.7.11) og *tvillingsprimtallsformodningen* (Formodning 3.7.12).

Fermats siste teorem. Dette er kanskje det aller mest berømte problemet i matematikk, ihvertfall blant ikke-matematikere:

[20] *Conjecture* på engelsk.

Teorem 1.4.1: Fermats siste teorem

Ligningen

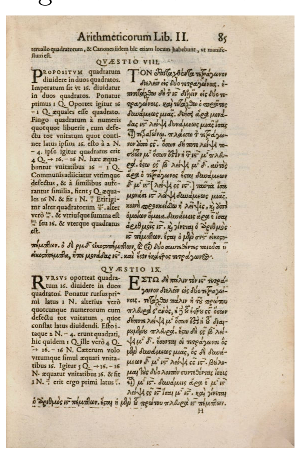
$$x^n + y^n = z^n$$

har ingen løsninger for positive heltall x, y, z, dersom n er et heltall større enn 2.

Merk at teoremet ikke gjelder for $n = 2$, siden $x^2 + y^2 = z^2$ har uendelig mange løsninger for positive heltall x, y, z; disse løsningene kalles *Pythagoreiske tripler*[21] og tilsvarer lengdene av siden på rettvinklede trekanter med sider av heltallslengder, for eksempel er $3^2 + 4^2 = 5^2$ (se Oppgave 1.7).

Pierre de Fermat. (Kilde: Wikipedia, public domain)

Teoremet har en fascinerende historie. Den franske matematikeren (og juristen!) Pierre de Fermat (1607–1665) skrev i margen på sin kopi av boken *Arithmetika* (skrevet av den greske matematikeren Diofantus i det tredje århundre f.Kr.) at han hadde funnet "et virkelig bemerkelsesverdig bevis" av dette teoremet, "som margen er for smal til å inneholde".



Arithmeticorum Lib. II. 85

teruallo quadratorum, & Canones iidem hic etiam locum habebunt, vt manifestum est.

QVÆSTIO VIII.

PROPOSITVM quadratum diuidere in duos quadratos. Imperatum sit vt 16. diuidatur in duos quadratos. Ponatur primus 1 Q. Oportet igitur 16 − 1 Q. æquales esse quadrato. Fingo quadratum à numeris quotquot libuerit, cum defectu tot vnitatum quot continet latus ipsius 16. esto à 2 N. − 4. ipse igitur quadratus erit 4 Q. + 16. − 16 N. hæc æquabuntur vnitatibus 16 − 1 Q. Communis adiiciatur vtrimque defectus, & à similibus auferantur similia, fient 5 Q. æquales 16 N. & fit 1 N. [illegible] Erit igitur alter quadratorum [illegible]. alter verò [illegible]. & vtriusque summa est [illegible] seu 16. & vterque quadratus est.

QVÆSTIO IX.

RVRSVS oporteat quadratum 16. diuidere in duos quadratos. Ponatur rursus primi latus 1 N. alterius verò quotcunque numerorum cum defectu tot vnitatum, quot constat latus diuidendi. Esto itaque 2 N. − 4. erunt quadrati, hic quidem 1 Q. ille verò 4 Q. + 16. − 16 N. Cæterum volo vtrumque simul æquari vnitatibus 16. Igitur 5 Q. + 16. − 16 N. æquatur vnitatibus 16. & fit 1 N. [illegible] erit ergo primi latus [illegible].

Den berømte siden i *Arithmetika*, med margen til høyre. (Kilde: Wikipedia, public domain)

[21] *Pythagorean triples* på engelsk.

I 1670, etter Fermats død, publiserte sønnen en ny utgave av *Arithmetika* med farens kommentar, som etter hvert ble kjent som *Fermats siste teorem*, selv om det ikke var bevist. Det er uklart, og faktisk høyst tvilsomt, om Fermat virkelig hadde et korrekt bevis for teoremet. Det faktum at han aldri publiserte beviset kan tyde på det motsatte.

Problemet fascinerte matematikere i over tre hundre år, og mange ma–tematikere hadde flere mislykkede forsøk på å bevise teoremet. Tilfellet $n = 3$ ble riktignok bevist av den sveitsiske matematikeren (og fysikeren, astronomen, geografen, logikeren og ingeniøren!) Leonhard Euler (1707–1783) og $n = 4$ av Fermat selv. Problemet ble også kjent vidt utenfor matematiske kretser: I episoden "The Royale" av science fiction serien "Star Trek: The Next Generation" fra 1989, se

https://www.youtube.com/watch?v=xNk0QYP0w5Q,

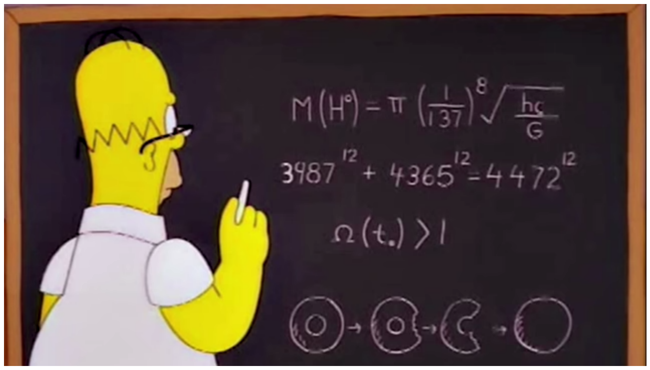
forteller Kaptein Jean-Luc Picard om sine forsøk på å bevise Fermats siste teorem, og han hevder blant annet at "mennesker har forsøkt å vise dette i mer enn 800 år". Dette er en morsom anakronisme, siden teoremet ble endelig bevist i 1994 av den britiske matematikeren Andrew John Wiles (1953–). Wiles brukte nesten 10 år på dette arbeidet, som resulterte i et bevis på flere hundre sider, som bygger på mye avansert matematikk fra 1900-tallet, deriblant *algebraisk tallteori* og *elliptiske kurver*. Wiles annonserte beviset sommeren 1993, men det ble funnet et alvorlig hull i beviset noen måneder senere. Høsten 1994 leverte Wiles inn to artikler, den ene i samarbeid med sin tidligere student Richard Taylor (1962–), til fagfellevurdering[22] i det prestisjefylte tidsskriftet *Annals of Mathematics*, hvor hullet i beviset ble tettet igjen. Artiklene ble publisert våren 1995.

ANDREW WILES (COPYRIGHT C. J. MOZZOCHI, PRINCETON N.J)

[22] *Peer review* på engelsk. Dette er prosessen der et tidsskrift ber fageksperter om å uttale seg om arbeidets betydning og riktighet, før et arbeid blir publisert.

Wiles har mottatt flere priser for sitt bevis, som *Abelprisen*[23] i 2016.

For øvrig har Fermats siste teorem funnet sin vei inn i TV-serien *The Simpsons*. I episoden "The Wizard of Evergreen Terrace" fra 1998 skriver Homer Simpson likheten $3987^{12} + 4365^{12} = 4472^{12}$, som (hvis riktig) ville motsi Fermats siste teorem:

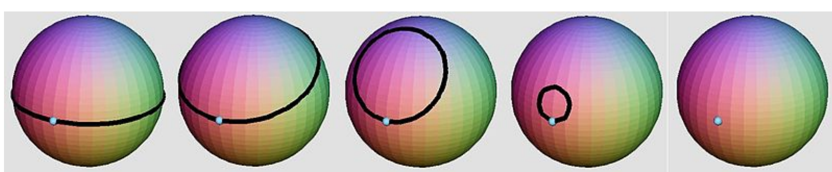


Homer Simpson finner tilsynelatende et moteksempel til Fermats siste teorem

Det er til og med blitt laget en musikal kalt *Fermat's last tango* inspirert av Wiles' søken etter beviset for Fermats siste teorem. Musikalen er produsert av *Clay Mathematical Institute*, en organisasjon som arbeider med utdanning i matematikk. Hele musikalen er tilgjengelig på

https://www.youtube.com/watch?v=RNqQPcMYcG8.

Poincarés formodning. Denne formodningen ble fremsatt av den franske matematikeren (og fysikeren, ingeniøren og vitenskapsfilosofen) Henri Poincaré (1854–1912) i 1904. Noe upresist sier den *todimensjonale* analogien til formodningen at enhver lukket, *enkeltsammenhengende*[24] flate kan deformeres på en kontinuerlig måte (det vi si uten at man river den i stykker og setter den sammen igjen) til en kuleflate. At flaten er enkeltsammenhengende, betyr at den er sammenhengende[25] og at enhver løkke på flaten kan kontrakteres til et punkt.



Enhver løkke på kuleflaten kan kontrakteres til et punkt
(Kilde: Wikipedia, GNU Free Documentation License.)

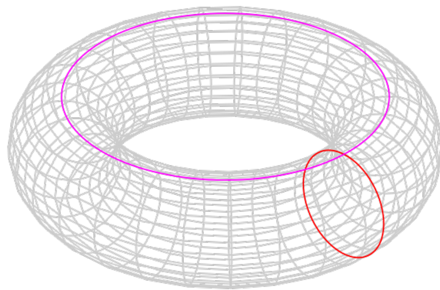
Et eksempel på en sammenhengende, men ikke enkeltsammenhengende flate, er overflaten til en smultring, kalt *torus* i matematisk språk: ingen av de to

[23]Denne prisen utdeles årlig siden 2003 av *Det norske Videnskaps-Akademi* til forskere som har utmerket seg i matematikk og er oppkalt etter den norske matematikeren Niels Henrik Abel (1802–1829).

[24]*Simply connected* på engelsk.

[25]*Connected* på engelsk.

fargede løkkene på figuren nedenfor kan kontrakteres til et punkt. (Vi skal se nærmere på torusen i Eksempel 10.10.12.)



To løkker på torusen som ikke kan kontrakteres til et punkt. (Kilde: Wikipedia, public domain)

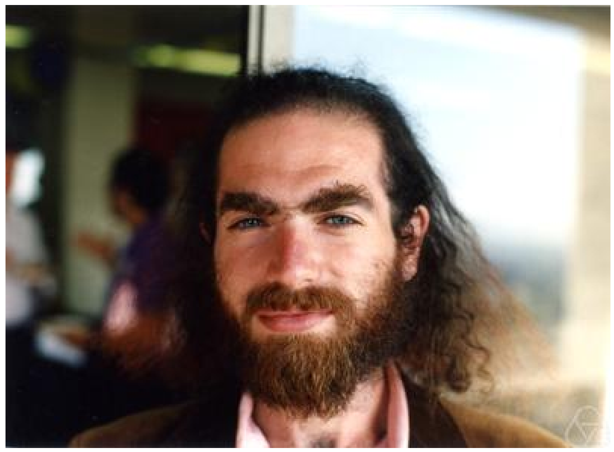
Poincarés formodning sier at det tilsvarende utsagnet også stemmer for tredimensjonale objekter: Et lukket, enkeltsammenhengende tredimensjonalt geometrisk objekt[26] kan deformeres på en kontinuerlig måte til en tredimensjonal sfære.

Henri Poincaré (Kilde: Wikipedia, public domain)

Formodningen ble viet stor interesse blant matematikere, og flere feilaktige bevis ble fremsatt. Den ble løst etter nesten hundre år av den russiske matematikeren Grigorij Perelman (1966–), som imidlertid ikke sendte sitt arbeid inn til et matematisk tidsskrift for fagfellevurdering, men bare la det ut på databasen arXiv på internett, som er den vanlige databasen der matematikere (og andre vitenskapsfolk) legger ut sitt arbeid i påvente av at fagfellevurderingen fullføres. Arbeidet skapte selvsagt stor interesse og

[26]Her er vi med vilje litt upresise: det riktige matematiske uttrykket i dette tilfellet er en *mangfoldighet* ("*manifold*" på engelsk), som, grovt sagt, er et topologisk rom (jf. Definisjon 10.10.1) som har lokale koordinater på seg. Dimensjonen til en mangfoldighet er, igjen grovt sagt, det minste antallet koordinater som må til for å beskrive objektet lokalt. For eksempel er dimensjonen til kulen og torusen lik 2.

ble grundig undersøkt av flere matematikere, og Perelman selv holdt gjesteforelesninger om arbeidet ved flere prestisjefylte universiteter. Beviset for formodningen, langt og komplisert, ble etter noen år godkjent av det internasjonale matematikersamfunnet. Perelman viste imidlertid ikke interesse for å publisere arbeidet i noe tidsskrift, og takket nei til flere priser, blant annet den prestisjefylte *Fieldsmedaljen*[27] i 2006, som regnes som den gjeveste prisen i matematikk, og *the Millennium Prize* i 2010 fra *Clay Mathematical Institute*, som ville ha gitt ham 1 million dollar!



Grigorij Perelman (Kilde: Wikipedia, public domain)

Riemannhypotesen. *Riemannhypotesen* regnes kanskje som det viktigste uløste problemet i matematikken, ihvertfall etter de nylige løsningene på *Fermats siste teorem* og *Poincarés formodning* beskrevet ovenfor.

Den italienske matematikeren Pietro Mengoli (1626–1686) fremla i 1650 et problem som ble veldig kjent og fikk navnet *Basel-problemet*: det dreide seg om å finne summen på rekken

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \cdots$$

(Her er tegnet Σ den store greske bokstaven "sigma". Vi vil se senere i teksten, jf. slutten av §7.4, at vi med denne summen mener grensen av følgen $\{1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \cdots + \frac{1}{n^2}\}$.) Den sveitsiske matematikeren Leonhard Euler (1707–1783) løste problemet i 1735 ved å vise at summen er $\frac{\pi^2}{6}$. Dette gjorde Euler umiddelbart berømt i en alder av 28. Euler generaliserte problemet betraktelig og studerte funksjonen

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{3^s} + \cdots ,$$

[27]Denne prisen deles ut hvert fjerde år siden 1924 av *International Mathematical Union* under IMUs internasjonale matematikkongress til fremragende matematikere under 40 år.

for et vilkårlig reellt tall $s > 1$ (hvor ζ er den greske bokstaven "zeta"), som han viste var konvergent (altså at summen har en veldefinert grense), og han fant en forbindelse mellom denne funksjonen og primtallene (se §3.7 for definisjon av primtall). Blant annet viste Euler at $\zeta(s)$ er lik det uendelige produktet

$$\zeta(s) = \frac{1}{1-2^{-s}} \cdot \frac{1}{1-3^{-s}} \cdot \frac{1}{1-5^{-s}} \cdot \frac{1}{1-7^{-s}} \cdot \frac{1}{1-11^{-s}} \cdots, \tag{24}$$

hvor altså alle faktorene er på formen $\frac{1}{1-p^{-s}}$ med p primtall; vi skriver et slikt produkt gjerne som

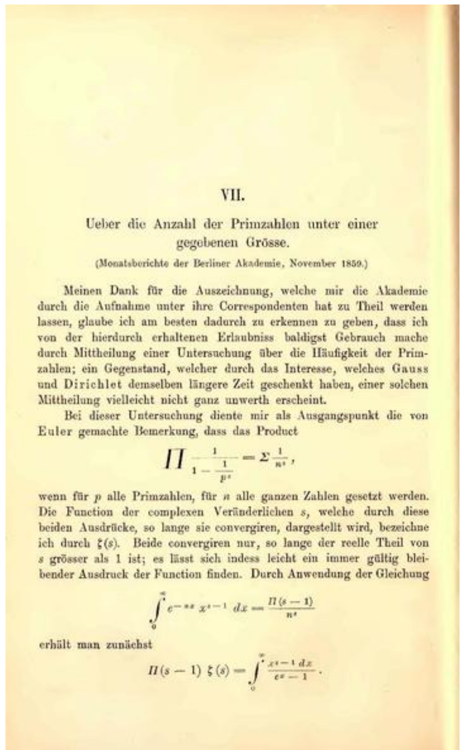
$$\zeta(s) = \prod_{p \text{ primtall}} \frac{1}{1-p^{-s}}. \tag{25}$$

(Her er Π den store greske bokstaven "*pi*" og brukes for å uttrykke et produkt.) Ved hjelp av denne formelen kunne Euler blant annet gi et nytt bevis for at det finnes uendelig mange primtall (dette er Teorem 3.7.10, som vi skal gi et annet bevis for). I 1837 kunne den tyske matematikeren Peter Gustav Lejeune Dirichlet (1805–1859) forfine Eulers bevis og vise at for hvert par av positive heltall a og b uten felles faktorer, finnes uendelig mange primtall på formen $an + b$, med n et heltall.

BERNHARD RIEMANN (KILDE: WIKIPEDIA, PUBLIC DOMAIN)

I 1859 publiserte den tyske matematikeren Bernhard Riemann (1826–1866) en 9–siders artikkel med tittelen *Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse* ("Om antallet primtall mindre enn en gitt størrelse"), som inneholder ideer som har påvirket uttallige matematikere til den dag i dag. I denne artikkelen brukte Riemann *zeta*-funksjonen til å finne antallet primtall $\pi(x)$ mindre eller lik et gitt tall x.

I sitt studium viste Riemann at man kunne utvide *zeta*-funksjonen på en slik måte at den ble definert for alle *komplekse tall* $s \neq 1$. (Vi skal definere komplekse tall i 8.1, jf. også [**AE**, App. I] eller [**Li**, Kap. 3]. For øyeblikket nevner vi kun at et slikt tall har formen $a + ib$, hvor a og b er reelle tall og i er et symbol som oppfyller $i^2 = -1$.) Riemann kunne lett vise at $\zeta(s) = 0$ når s er et negativt partall. I artikkelen fremsatte han så en formodning om at alle andre nullpunkter til ζ er på formen $\frac{1}{2} + ib$, med b et reellt tall.



VII.

Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse.

(Monatsberichte der Berliner Akademie, November 1859.)

Meinen Dank für die Auszeichnung, welche mir die Akademie durch die Aufnahme unter ihre Correspondenten hat zu Theil werden lassen, glaube ich am besten dadurch zu erkennen zu geben, dass ich von der hierdurch erhaltenen Erlaubniss baldigst Gebrauch mache durch Mittheilung einer Untersuchung über die Häufigkeit der Primzahlen; ein Gegenstand, welcher durch das Interesse, welches Gauss und Dirichlet demselben längere Zeit geschenkt haben, einer solchen Mittheilung vielleicht nicht ganz unwerth erscheint.

Bei dieser Untersuchung diente mir als Ausgangspunkt die von Euler gemachte Bemerkung, dass das Product

$$\Pi \frac{1}{1 - \frac{1}{p^s}} = \Sigma \frac{1}{n^s},$$

wenn für p alle Primzahlen, für n alle ganzen Zahlen gesetzt werden. Die Function der complexen Veränderlichen s, welche durch diese beiden Ausdrücke, so lange sie convergiren, dargestellt wird, bezeichne ich durch $\zeta(s)$. Beide convergiren nur, so lange der reelle Theil von s grösser als 1 ist; es lässt sich indess leicht ein immer gültig bleibender Ausdruck der Function finden. Durch Anwendung der Gleichung

$$\int_0^\infty e^{-nx} x^{s-1}\, dx = \frac{\Pi(s-1)}{n^s}$$

erhält man zunächst

$$\Pi(s-1)\,\zeta(s) = \int_0^\infty \frac{x^{s-1}\, dx}{e^x - 1}.$$

FØRSTE SIDE AV RIEMANNS ARTIKKEL *Ueber die Anzahl der Primzahlen...* (KILDE: WIKIPEDIA, PUBLIC DOMAIN)

Problemet er fremdeles uløst og er ett av *the Millennium Prize problems* etablert i 2010 av *Clay Mathematical Institute*, som har lovet den eller de som løser problemet 1 million dollar.

Artikkelen *Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse* er den eneste artikkelen Riemann skrev om tallteori. Blant studenter er Riemann nok mest kjent for å ha definert (Riemann)integrasjon rigorøst. Ellers skrev Riemann innflytelsesrike arbeider innenfor *kompleks analyse* (grovt sagt studiet av funksjoner som involverer komplekse tall) og flere grener av *geometri*. Hans arbeider kombinerte analyse og geometri på nye måter og dannet grunnlaget for store deler av *differensialgeometri*, *algebraisk geometri* og *kompleks geometri*. Riemanns geometri og teori om *krumning*[28] av rom er selve grunnlaget for *den generelle relativitetsteorien* publisert av Albert Einstein (1879–1955) i 1915.

[28] *Curvature* på engelsk.

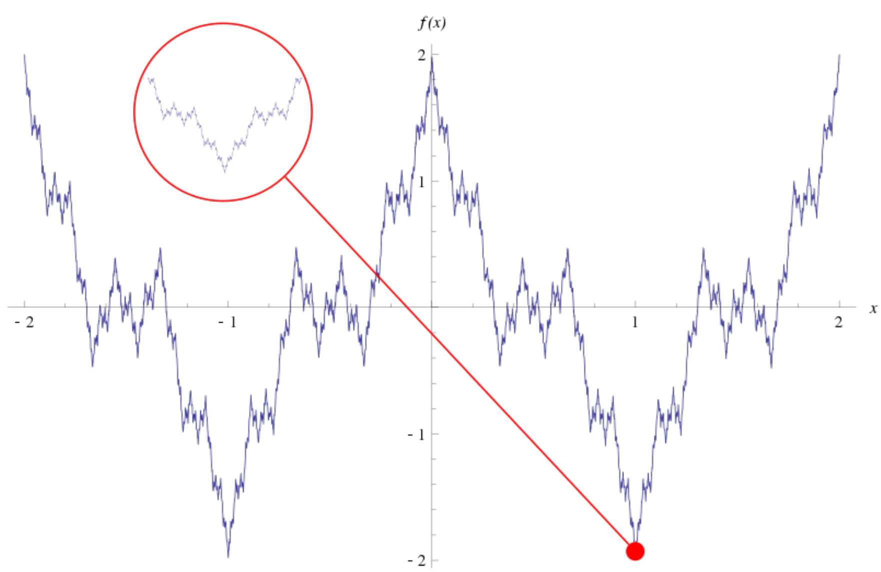
Trist nok led Riemann av tuberkolose. Selv lange opphold i det varmere klimaet i Italia i de siste leveårene hjalp ikke. Han døde ved Lago Maggiore i Nord-Italia kort tid før han fylte 40.

Ingensteds deriverbare funksjoner. Langt ut på 1800-tallet trodde de fleste matematikere at kontinuerlige funksjoner er deriverbare i "de fleste punkter" og mange forsøkte sogar å gi bevis for dette, uten å lykkes. Dette stemmer med den intuisjonen vi har opparbeidet oss innenfor skolematematikken: her er de fleste funksjonene bygd opp av pene og kjente funksjoner ved hjelp av de aritmetiske operasjonene addisjon, subtraksjon, multiplikasjon og divisjon, og noen ganger ved å legge inn en delt forskrift. Slike funksjoner har pene egenskaper: de er som regel kontinuerlige og deriverbare, bortsett fra i isolerte punkter. Geometrisk betyr dette at grafen til en kontinuerlig funksjon stort sett er glatt, bortsett fra i noen få isolerte punkter. Slik var også oppfatningen blant erfarne matematikere på 1800-tallet.

Det skapte derfor stor furore da den tyske matematikeren Karl Weierstrass (1815–1897) i 1872 publiserte et eksempel på en funksjon som er kontinuerlig overalt, men ikke er deriverbar i noe punkt! Vi skal ikke gå inn i detaljer: *Weierstrass-funksjonen* er definert ved hjelp av en uendelig sum som

$$f(x) = \cos(\pi x) + \frac{3}{4}\cos(9\pi x) + \frac{9}{16}\cos(81\pi x) + \frac{27}{64}\cos(729\pi x) + \cdots, \tag{26}$$

og grafen ser slik ut:



Grafen til Weierstrass-funksjonen (Kilde: Wikipedia, public domain)

Det innringede fenomenet viser at mønsteret gjentar seg uansett hvor mye vi zoomer inn (og det er nettopp dette som gjør at funksjonen ikke er deriverbar i noe punkt). Matematisk kalles dette fenomenet *fraktaler*[29] og fenomenet dukker opp i mange mønstre i naturen.

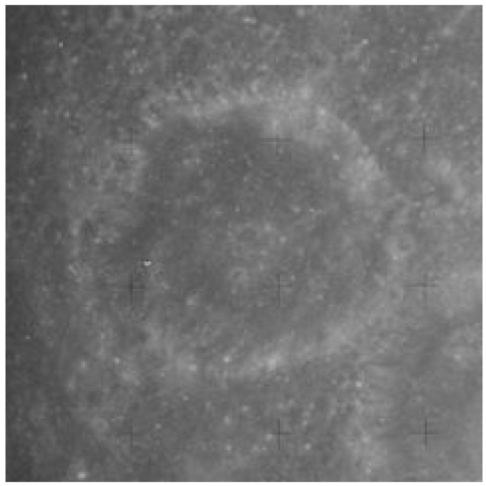
Karl Theodor Wilhelm Weierstrass (Kilde: Wikipedia, public domain)

Funksjonen og arbeidet til Weierstrass rokket så mye ved den allminnelige oppfatningen blant matematikere flest at resultatet ikke ble så godt mottatt, ikke minst fordi den slo benene under mange tidligere bevis som hadde basert seg på feilaktig intuisjon og vag forståelse av begrepet glatthet. Henri Poincaré kalte funksjoner som Weierstrass-funksjonen for "monstre". Den franske matematikeren Charles Hermite (1822–1901) skrev i et brev til den nederlanske matematikeren Thomas Joannes Stieltjes (1856–1894): "Jeg snur meg med frykt og avsky fra denne beklagelige epidemien av kontinuerlige funksjoner som ikke har noen derivert". Det er verdt å merke seg at grafen til Weierstrass-funksjonen ikke var mulig å visualisere før man fikk datamaskiner hundre år senere, slik at beviset kun bestod av teknisk krevende teoretiske steg. Weierstrass' resultat ble imidlertid sakte, men sikkert allmennt akseptert, spesielt siden man fant ut at mange fenomener fra virkeligheten faktisk kun kan beskrives nettopp gjennom kontinuerlige ingensteds deriverbare funksjoner: det er tilfellet for eksempel innenfor teorien av *Brownske bevegelser*, som har store anvendelser i fysikk, økonomi og elektronikk. Disse er oppkalt etter den skotske botanikeren

[29] *Fractals* på engelsk.

Robert Brown (1773–1858), som først beskrev fenomenet i 1827 mens han studerte pollen under vann i mikroskop.

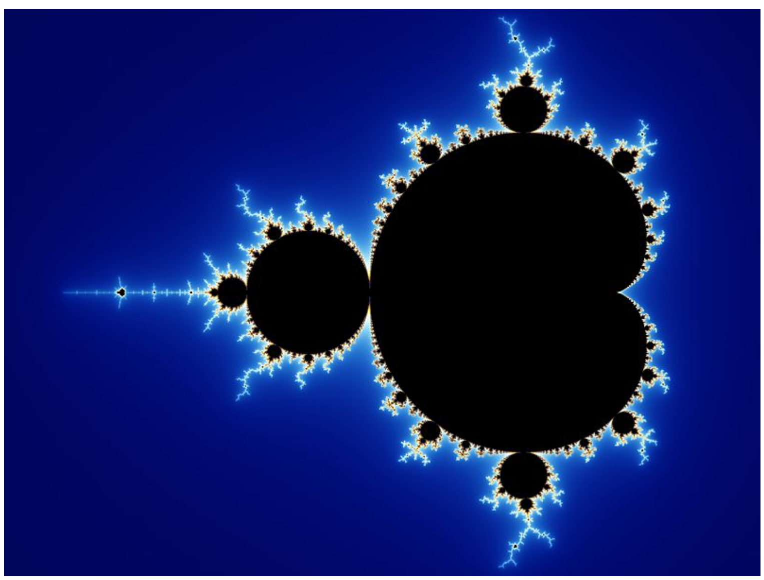
Vi skal møte på Weierstrass mange ganger i løpet av denne teksten, siden han har spilt en veldig viktig rolle i utviklingen av matematikken slik vi kjenner den i dag. For øvrig har Weierstrass fått oppkalt etter seg både asteroiden 14100 Weierstrass og Weierstrass-krateret på månen.



Weierstrass-krateret på månen (Kilde: Wikipedia, public domain)

Som en kuriositet nevner vi at en av de mest kjente fraktalmengdene, *Mandelbrot-mengden* vist i figuren nedenunder, har funnet veien inn i åpningssekvensen til James Bond filmen *Casino Royale* fra 2006, se

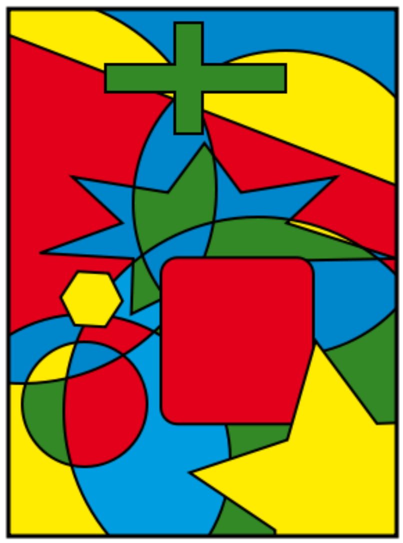
https://www.youtube.com/watch?v=A1AMUmkj-ck.



Mandelbrot-mengden i planet (i svart) (Kilde: Wikipedia, public domain)

For definisjonen av Mandelbrot-mengden, se Eksempel 9.8.2.

Firfargeproblemet. Dette problemet ble først satt fram i 1852 av den sørafrikanske matematikeren (og botanikeren!) Francis Guthrie (1831–1899), da han merket at han kun trengte fire farger for å fargelegge alle fylkene[30] i England, på en slik måte at ingen av fylkene med felles grense hadde samme farge. Han var, sammen med sin bror, student av den britiske matematikeren og logikeren Augustus De Morgan (1806–1871), og stilte sistnevnte spørsmålet via et brev om hvorvidt *et landkart i planet eller på kuleflaten kan farges med fire farger slik at land med felles grenselinje alltid får forskjellig farge.*



Eksempel på kart med fire farger
(Kilde: Wikipedia, GNU Free Documentation License)

Mange matematikere forsøkte å løse firfargeproblemet, og flere bevis ble publisert, som i etterkant viste seg å være feil. Problemet ble endelig løst (med positivt svar) av matematikerne Kenneth Appel (1932–2013) og Wolfgang Haken (1928–) ved University of Illinois at Urbana-Champaign i 1976 ved omfattande bruk av datamaskiner. En feil ble funnet av den tyske Masterstudenten(!) Ulrich Schmidt ved RWTH Aachen noen år senere, men denne (og andre) ble rettet i en bok av Appel og Haken fra 1989. Senere har flere alternative, og enklere bevis, blitt funnet, som ikke bygger så mye på bruk av datamaskin.

Erdős–Straus formodningen. Denne formodningen er ikke blant de mest kjente, men vi tar den med fordi den er svært enkel å formulere og er interessant historisk sett, foruten å være fremdeles uløst!

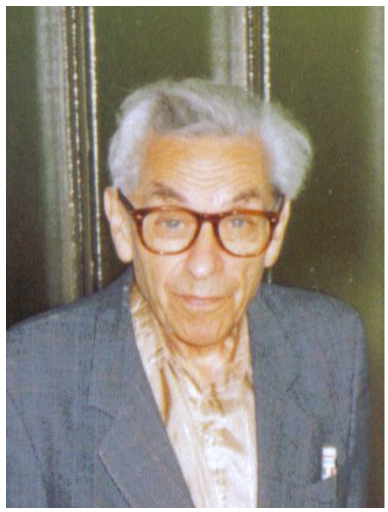
Egypterne regnet med brøker allerede et par tusen år før Kristus, men uttrykte alle brøker som sum av brøker på formen $\frac{1}{n}$, såkalte *stammbrøker*. Fremdeles i dag kalles en sum av stammbrøker for en *Egyptisk brøk*.

[30]Rettere sagt *counties*.

Egyptiske brøker er av interesse innen tallteori i dagens matematikk. I 1948 fremsatte den ungarske matematikeren Paul Erdős (1913–1996) og den tysk-amerikanske matematikeren Ernst Gabor Straus (1922–1983) følgende formodning: *for ethvert heltall $n \geq 2$, kan $\frac{4}{n}$ skrives som en sum av tre positive stammbrøker*, med andre ord finnes positive heltall x, y, z slik at

$$\frac{4}{n} = \frac{1}{x} + \frac{1}{y} + \frac{1}{z}.$$

Ved hjelp av datamaskiner er formodningen verifisert for alle $n \leq 10^7$.



PAUL ERDŐS (KILDE: WIKIPEDIA, CREATIVE COMMONS ATTRIBUTION 3.0 UNPORTED LICENSE)

Paul Erdős er en av de mest produktive matematikere gjennom tidene, med mer enn 1500 artikler skrevet med mer enn 500 medforfattere. Han er også kjent for å ha fremsatt mange formodninger, for det meste innenfor tallteori og kombinatorikk. Han viet hele sitt liv til matematikk, hadde ikke fast bosted og reiste rundt i verden for å arbeide med andre matematikere, delta på konferanser og besøke universiteter og forskningsinstitusjoner. Han hadde svært lite bagasje og svært få eiendeler. Mesteparten av penger han fikk fra priser donerte han til veldedige formål eller til stipender for yngre forskere.

Idag regnes det såkalte *Erdős-tallet* til en matematiker å være 1 hvis vedkommende har skrevet en artikkel med Erdős, 2 hvis vedkommende ikke har skrevet en artikkel med Erdős, men skrevet en artikkel med noen som har, osv. (bortsett fra at Erdős-tallet til Erdős selv er 0).

$3x+1$-formodningen. Et annet problem som er enkelt å beskrive, men som har vist seg vanskelig å løse er følgende:

FORMODNING 1.4.2 ("$3x + 1$-formodningen" eller "Collatz-formodningen"). *Definér en funksjon T på heltallene ved at*

$$T(n) = \begin{cases} \frac{n}{2}, & \textit{hvis } n \textit{ er partall,} \\ 3n+1, & \textit{hvis } n \textit{ er oddetall} \end{cases}$$

og la $a_0, a_1, a_2, a_3, \ldots$ *være følgen av heltall man får ved å starte med et vilkårlig heltall og anvende funksjonen* T *gjentagende ganger; det vil si*

$$a_0 = \textit{vilkårlig heltall}, \quad a_i = T(a_{i-1}) \quad \textit{for} \quad i = 1, 2, 3, 4, \ldots$$

Da vil følgen inneholde tallet 1.

Som et eksempel, starter vi med $a_0 = 12$, får vi følgen

$$12, 6, 3, 10, 5, 16, 8, 4, 2, 1, \ldots$$

Starter vi med $a_0 = 19$, får vi følgen

$$19, 58, 29, 88, 44, 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1, \ldots$$

Formodningen går tilbake til den tyske matematikeren Lothar Collatz (1910–1990) i 1937 og har siden 50–tallet fascinert uttallige matematikere, se [**La**] for mer bakgrunn. Siden mange matematikere har blitt drevet bort fra sine opprinnelige forskningsprosjekter pga. denne formodningen, har dette ført til en spøk om at problemet er del av en konspirasjon for å svekke amerikansk forskning i matematikk! Hittil har formodningen blitt verifisert for alle startverdier $a_0 \leq 87 \cdot 2^{60}$, og er ellers fullstendig åpen.

Også dette problemet har blitt kjent utenfor matematiske kretser: Den kanadiske filmen *Incendies* fra 2010, regissert av Denis Villeneuve (1967–), åpner med en scene hvor en kvinnelig matematikkstudent forklarer formodningen for en gruppe yngre studenter. Dette før hun må avbryte studiene for å ta seg av uløste problemer fra familiens fortid. Det viser seg at $3x+1$-formodningen forutser en dyster familiehemmelighet (hvorfor skulle man ellers ha åpnet filmen med den?).

Oppgaver

Oppgave 1.1. Vis, ved kun å bruke Definisjonene 1.2.1 og 1.2.2, at et heltall ikke kan være samtidig et partall og et oddetall.

Oppgave 1.2. Finn feilen i følgende berømte "bevis" for at $1 = 2$:

"Bevis". Vi starter med to like positive heltall a og b og utfører følgende steg:

$$\begin{aligned}
a &= b && \text{(Gitt)}\\
a^2 &= ab && \text{(Multiplisert begge sidene med } a\text{)}\\
a^2 - b^2 &= ab - b^2 && \text{(Subtrahert } b^2 \text{ på begge sider)}\\
(a-b)(a+b) &= b(a-b) && \text{(Faktorisert begge sider)}\\
a + b &= b && \text{(Delt begge sider med } a - b\text{)}\\
2b &= b && \text{(Erstattet } a \text{ med } b\text{, siden } a = b\text{)}\\
2 &= 1 && \text{(Delt begge sider med } b\text{)}
\end{aligned}$$

□

Oppgave 1.3. La n være et heltall.
Vis at dersom $n^3 + 5$ er et oddetall, da er n et partall.

Oppgave 1.4. La m og n være heltall.
Vis at dersom mn er et partall, da må minst én av m og n være et partall.

Oppgave 1.5. La m og n være heltall.
Vis at dersom både mn og $m + n$ er partall, da er både m og n partall.

Oppgave 1.6. Vis at det ikke finnes noe rasjonalt tall x som oppfyller $x^3 + x + 1 = 0$.

Oppgave 1.7. Vis at det finnes uendelig mange *Pythagoreiske tripler*, det vil si tripler av positive heltall x, y, z som oppfyller ligningen $x^2 + y^2 = z^2$, og som ikke er multipler av hverandre. (To tripler (x, y, z) og (x', y', z') er multipler av hverandre dersom det finnes et positivt heltall n slik at $(x', y', z') = (nx, ny, nz)$.)

Hint: La $x = m^2 - n^2$, $y = 2mn$ og $z = m^2 + n^2$, hvor m og n er heltall. Denne metoden for å generere Pythagoreiske tripler virker å ha vært kjent av babylonerne (1900-1650 f.Kr.), men den første fullstendige beskrivelsen av metoden finner vi i *Elementer*, av den greske matematikeren Euklid (ca. 300 f.Kr.).

Oppgave 1.8. Avgjør om følgende resonnement er logisk gyldig:

- "Regnfulle dager gjør at hager gror"
- "Hager gror ikke dersom det ikke er varmt"
- "Det regner alltid når det ikke er varmt"
- "Derfor, om det ikke er varmt, så er det varmt"

Oppgave 1.9. Følgende logiske resonnementer er hentet fra *Erasmus Montanus*, skrevet av den bergenskfødte forfatteren Ludvig Holberg (1684–1754) i 1723. Avgjør hvilke som er (logisk sett) gyldige:

Første eksempel: "Morlille er en sten"

```
MONTANUS: Morlille, jeg vil gjøre jer til en sten.
NILLE: Ja, snakk; det er enda mer kunstig.
MONTANUS: Nu skal I få dette å høre:  En sten kan ikke fly.
NILLE: Nei; det er visst nok, unntatt man kaster den.
MONTANUS: I kan ikke fly.
NILLE: Det er og sant.
MONTANUS: Ergo er Morlille en sten.  (Nille gråter.)  Hvorfor gråter Morlille?
NILLE: Akk, jeg er så bange at jeg blir til sten.  Mine ben begynner alt å bli
kalde.
```

Andre eksempel: "Morlille er ingen sten"

```
MONTANUS:  Gi jer tilfreds, Morlille!  Jeg skal straks gjøre jer til menneske
igjen.  En sten kan ikke tenke eller tale.
NILLE: Det er sant.  Jeg vet ikke om den kan tenke, men tale kan den ikke.
MONTANUS: Morlille kan tale.
NILLE: Ja, gudskjelov, som en stakkars bondekone kan jeg tale.
MONTANUS: Godt.  Ergo er Morlille ingen sten.
```

NILLE: Akk, det gjorde godt, nu kommer jeg meg igjen. Det må sannelig sterke hoder til å studere, jeg vet ikke hvordan deres hjerne kan holde det ut.

Tredje eksempel: "Den som drikker er lykksalig"

MONTANUS: (...) Hør, farlille, vil I tro at den som drikker vel, er lykksalig?
JEPPE: Jeg tror snarere at han er ulykksalig, ti man kan drikke både forstand og penger bort.
MONTANUS: Jeg vil bevise at han er lykksalig. Qvicunque bene bibit, bene dormit. Nei, det er sant, I forstår ikke latin, jeg må si det på dansk. Den som drikker vel, sover gjerne vel, er det ikke sant?
JEPPE: Det er sant nok. Når jeg har en halv rus, sover jeg som en hest.
MONTANUS: Den som sover vel, synder ikke, er ikke det også sant?
JEPPE: Jo, det er sant nok. Så lenge man sover, synder man ikke.
MONTANUS: Den som ikke synder, er lykksalig.
JEPPE: Det er og sant.
MONTANUS: Ergo, den som drikker vel, er lykksalig.

Statue av Ludvig Holberg i Bergen
(Kilde: Wikipedia, GNU Free Documentation License).

Oppgave 1.10. På en øy er alle innbyggere enten riddere, som alltid snakker sant, eller kjeltringer, som alltid lyver.

Du ankommer et veiskille og ved skillet finnes to innbyggere som hver vokter sin vei. Du vet at den ene er en ridder og den andre er en kjeltring, men du vet ikke hvem som er hvem. Du vet også at den ene veien leder til sikker død, den andre til frihet.

Kan du finne ut hvilken vei du skal ta ved å stille ett eneste ja/nei-spørsmål til én av de to innbyggerne?

(Denne oppgaven, og neste, er basert på idéen av "Knights and knaves" av den amerikanske matematikeren, logikeren og filosofen Raymond Smullyan (1919–2017), kjent for sine mange logiske gåter.)

Oppgave 1.11. På en øy er alle innbyggere enten riddere, som alltid snakker sant, kjeltringer, som alltid lyver, eller spioner, som kan både lyve og snakke sant.

Du treffer tre personer, A, B og C, og du vet at én er ridder, én er kjeltring og én er spion.

A sier: "Jeg er ridderen". B sier: "Jeg er kjeltringen". C sier "B er ridderen". Hvem er hva?

Oppgave 1.12. Herr X er blitt myrdet. Du er politiinspektør og har tre mistenkte: Herr A, Herr B og Herr C. De sier følgende:

- Herr A: "B og X var gode venner og C mislikte X".
- Herr B: "Jeg kjente ikke X og var utenombys da mordet skjedde".
- Herr C: "A og B var sammen med X mordnatten og A skyldte X penger".

Anta at det ikke finnes mer enn én morder og at de uskyldige snakker sant. Hvem, om noen, er morderen?

Oppgave 1.13. En mann har tre døtre. En intelligent kvinne spør ham om alderen til døtrene. Mannen svarer at produktet av aldrene er 36. Etter å ha tenkt et øyeblikk, ber kvinnen om et annet hint. Mannen svarer at summen av aldrene er lik hans eget dørnummer. Fremdeles er kvinnen ikke i stand til å finne ut alderen til døtrene og spør om enda et hint. Mannen sier da at den yngste datteren har brune øyne.

Hva er alderen til de tre døtrene?

Oppgave 1.14. Avgjør om følgende resonnement, tatt fra [**KM**], er logisk gyldig: *If God exists, then He is omnipotent. If God exists, then He is omniscient. If God exists, then He is benevolent. If God can prevent evil, then if He knows that evil exists, then He is not benevolent if he does not prevent it. If God is omnipotent, then He can prevent evil. If God is omniscient, then He knows that evil exists if it does indeed exist. Evil does not exist if God prevents it. Evil exists. Therefore, God does not exist.*

Oppgave 1.15. Vis at dersom Erdős–Straus formodningen er usann, da finnes et moteksempel n som er et primtall. (Se Definisjon 3.7.6 for definisjonen av et primtall.)

Oppgave 1.16. Bevis en av de åpne formodningene i §1.4 og oppnå heder og ære (og 1 million dollar dersom det dreier seg om *Riemannhypotesen*).

KAPITTEL 2

Mengder og funksjoner

Mange viktige deler av matematikken dreier seg om funksjoner mellom mengder. I skolematematikken, og i førsteårskursene i matematikk på universitetet, består disse mengdene som oftest av reelle tall.

Vi vil i dette kapitlet oppsummere de viktigste begrepene som omhandler mengder og funksjoner. Mye vil være kjent fra skolen, men stoffet om *injektive, surjektive, bijektive* og *inverse funksjoner* i §2.3, samt *relasjoner* i §2.4, vil være helt nytt og spille viktige roller for oss senere. Dessuten er notasjonen $f : X \to Y$ for en funksjon (jf. Definisjon 2.2.1) ny i forhold til skolen og første års kalkuluskurs på universitetet.

2.1. Mengder

En *mengde*[1] er en samling av elementer eller objekter, som ofte er tall, men som også kan være punkter, vektorer, bokstaver, Bjørn Eidsvåg låter eller andre ting. At objektet x er et element i mengden A, skriver vi som

$$x \in A,$$

og vi bruker også uttrykk som "x ligger i A" eller "x er med i A". At objektet x *ikke* er et element i A, skriver vi som

$$x \notin A.$$

Vi lister gjerne opp elementene i en mengde omgitt av klammeparenteser $\{\ldots\}$, som vist nedenfor.

Eksempler på mengder som vi kjenner fra skolen er:

- $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$, mengden av alle *heltall*[2].
- $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$, de *naturlige tallene*[3] (definisjoner i noen lærebøker tar ikke med 0).
- $\mathbb{Q}$, de *rasjonale tallene*[4], som er alle reelle tall som kan uttrykkes som en brøk av heltall, med andre ord

$$\mathbb{Q} = \left\{ \frac{a}{b} \;\middle|\; a, b \in \mathbb{Z}, b \neq 0 \right\},$$

[1] *Set* på engelsk.
[2] *Integers* på engelsk.
[3] *Natural numbers* på engelsk.
[4] *Rational numbers* på engelsk

hvor tegnet "|" betyr "slik at". (Noen ganger bruker også ":" for å betegne dette.)

- $\mathbb{R}$, de *reelle tallene*[5]; for øyeblikket vil vi nøye oss med å definere disse som alle tall som kan uttrykkes som desimaltall. (Vi skal se nærmere på konstruksjonen av de reelle tallene i §7.1.)
- *de irrasjonale tallene*[6], det vil si alle reelle tall som *ikke* er rasjonale, som for eksempel $\sqrt{2}$, jf. Setning 1.3.5; andre eksempler er π og eulertallet e (jf. Oppgave 7.41). Mengden av alle irrasjonale tall har ikke fått et eget symbol.
- $\emptyset$, den *tomme mengden*[7]. Dette er mengden som ikke inneholder noen elementer i det hele tatt. Det kan virke litt sært å ha en egen betegnelse på denne mengden, men den er nyttig når man for eksempel skal beskrive ligningssystemer som ikke har noen løsning.

Vi visualiserer gjerne de reelle tallene som alle punkter på en tenkt linje, *den reelle tallinjen*, som dermed inneholder også alle de rasjonale og hele tallene. På grunn av dette betegnes tallene også som *punkter*.

En annen mengde som vi skal bli kjent med i Kapittel 8, og som dere også blir kjent med i kurset MAT111 (se [**AE**, App. I]), er:

- $\mathbb{C}$, de *komplekse tallene*[8], som er alle uttrykk på formen $a+ib$, med $a, b \in \mathbb{R}$ og i er den *imaginære enheten*, som oppfyller $i^2 = -1$, med andre ord
$$\mathbb{C} = \left\{a + ib \mid a, b \in \mathbb{R}, i^2 = -1\right\}.$$

Et hovedmål med kurset MAT100 er å sette seg inn i de aksiomatiske definisjonene for alle disse mengdene, mens dere i de andre førsteårskursene i matematikk bare aksepterer at disse finnes. I dette kapitlet skal vi imidlertid uproblematisk akseptere at disse mengdene finnes og oppfyller alt vi er vant til at de gjør fra skolen. Det skal vi gjøre for å være i stand til å gi noen flere eksempler på begrepene vi skal innføre.

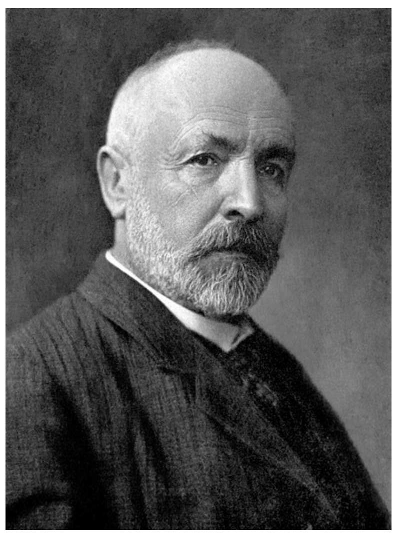
EKSEMPEL 2.1.1. Mengden av alle vokaler i det norske språket er {a,e,i,o,u,y,æ,ø,å}.

EKSEMPEL 2.1.2. En mengde kan også bestå av mengder. For eksempel er $S = \{\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$ en mengde som inneholder fire elementer, nemlig mengdene $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$. Vi har $1 \notin S$, selv om 1 er et element i hver av mengdene $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ og $\mathbb{R}$.

[5] *Real numbers* på engelsk
[6] *Irrational numbers* på engelsk
[7] *The empty set* på engelsk
[8] *Complex numbers* på engelsk

MERKNAD 2.1.3. Definisjonen av en mengde som en samling av objekter går tilbake til den tyske matematikeren Georg Cantor (1845–1918) i 1895. Denne intuitive måten å definere mengder på kalles i dag *naiv mengdeteori*[9]. Den fungerer utmerket i mange sammenhenger, men det er viktig å være klar over at denne måten å definere mengder på ved hjelp av egenskaper som er oppfylt av objektene i mengden kan føre til selvmotsigelser. Et berømt eksempel er *Russels paradoks*, som vi presenterer i Oppgave 2.11, oppdaget i 1901 av den britiske filosofen, logikeren, matematikeren og forfatteren Bertrand Russell (1872–1970). Disse logiske selvmotsigelsene kan unngås ved å bygge opp en mengdeteori ved hjelp av aksiomer, kalt *aksiomatisk mengdeteori*[10].

Vi vil ikke gå nærmere innpå dette og holde oss til Cantors naive mengdeteori, siden alle mengdene vi vil betrakte kan behandles på denne måten uten motsigelser.



GEORG CANTOR (KILDE: WIKIPEDIA, PUBLIC DOMAIN)

MERKNAD 2.1.4. I matematiske tekster forsøker man gjerne å unngå for mye bruk av symboler. I egne notater eller på tavlen under en forelesning bruker man imidlertid langt flere symboler for å komprimere teksten og spare tid. Som et eksempel vil man med symbolene introdusert ovenfor og tegnet $\exists$ som betyr "eksisterer", og dens negasjon $\nexists$, kunne uttrykke utsagnet "Det finnes ikke noe rasjonalt tall x som oppfyller $x^2 = 2$" fra Setning 1.3.5 som

$$\nexists x \in \mathbb{Q} \mid x^2 = 2.$$

[9] *Naïve set theory* på engelsk.
[10] *Axiomatic set theory* på engelsk.

Symboler og operasjoner knyttet til mengder. Når A og B er mengder, og alle elementene/tallene i mengden A også er elementer i B, sier vi at A *er en delmengde*[11] *av* B, og vi skriver $A \subset B$ eller $B \supset A$ og kaller disse også for en *inklusjon* av mengder. Vi leser disse også som, henholdsvis, "A er inneholdt i B" og "B inneholder A". Merk at vi automatisk har at $A \subset A$, det vil si at en mengde er en delmengde av seg selv. Per konvensjon er dessuten den tomme mengden $\emptyset$ inneholdt i alle mengder.

EKSEMPEL 2.1.5. Mengdene vi nevnte ovenfor oppfyller

$$\emptyset \subset \mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

Den siste inklusjonen skyldes at ethvert reelt tall a kan skrives som $a + 0 \cdot i$.

MERKNAD 2.1.6. Merk forskjellen på symbolbruken $x \in X$ og $\{x\} \subset X$. Med "$x \in X$" mener vi som sagt at objektet x er med i mengden X. Med "$\{x\} \subset X$" mener vi at "mengden bestående kun av x er en delmengde av X". I praksis gir begge uttrykkene samme informasjon, men avhengig av sammenhengen vil vi bruke den ene eller andre uttrykksmåten.

Vi sier at to mengder er *like*, og skriver $A = B$, dersom A og B inneholder nøyaktig de samme elementene. Merk at utsagnet $A = B$ er ekvivalent med at både $A \subset B$ og $B \subset A$.

Å si at $A \subset B$ utelukker ikke muligheten at $A = B$. Dersom $A \subset B$ og i tillegg $A \neq B$, sier vi at A er en *ekte delmengde*[12] av B. Et symbol for dette er $A \subsetneq B$. I Eksempel 2.1.5 er det altså like riktig å erstatte $\subset$ med $\subsetneq$ alle steder.

Vi sier at en mengde A er *ikke-tom*[13] dersom $A \neq \emptyset$, det vil si A inneholder minst ett element.

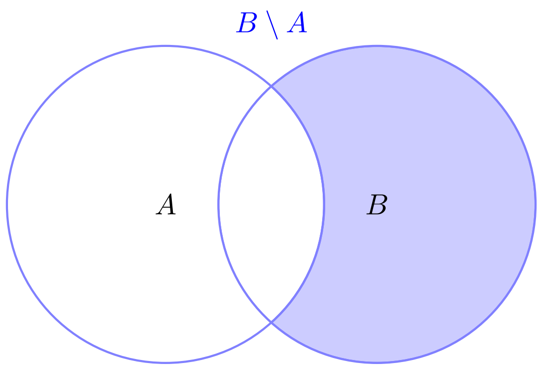
EKSEMPEL 2.1.7. Mengdene $\{1, 3, 5\}$ og $\{3, 5, 1\}$ er like, siden rekkefølgen av elementene ikke spiller noen rolle.

Hvis A og B er mengder, defineres *mengdedifferansen*

$$B \setminus A = \{x \mid x \in B,\ x \notin A\}.$$

Tegnet "$\setminus$" kan vi lese som "bortsett fra", men det leses gjerne også som "minus".

[11] *Subset* på engelsk.
[12] *Proper subset* på engelsk.
[13] *Nonempty* på engelsk.



Mengdedifferansen $B \setminus A$ i lyseblått.

Dersom $A \subset B$ kalles mengdedifferansen $B \setminus A$ for *komplemenentet til A i B*[14].

Eksempel 2.1.8. Vi har $i \in \mathbb{C} \setminus \mathbb{R}$, $\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$, $\frac{1}{2} \in \mathbb{Q} \setminus \mathbb{Z}$.

Eksempel 2.1.9. Mengden $\mathbb{R} \setminus \mathbb{Q}$ består av alle reelle tall som ikke er rasjonale, med andre ord er $\mathbb{R} \setminus \mathbb{Q}$ lik mengden av irrasjonale tall og er komplementet til mengden av alle rasjonale tall.

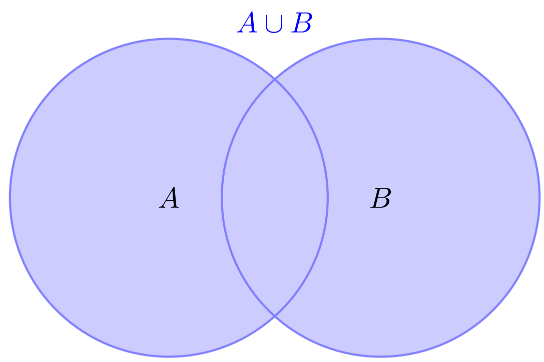
Eksempel 2.1.10. Hvis $B = \{1, 3, 5, 7\}$ og $A = \{1, 2, 3, 4\}$, da er $B \setminus A = \{5, 7\}$.

Eksempel 2.1.11. Komplementet til mengden av alle oddetall i $\mathbb{N}$ er mengden av alle partall.

Dersom A og B er mengder er *unionen*[15] av mengdene A og B de elementene som er med i *minst én av mengdene*; formelt er unionen av to mengder A og B definert som

$$A \cup B = \{x \mid x \in A \text{ eller } x \in B\}. \tag{27}$$

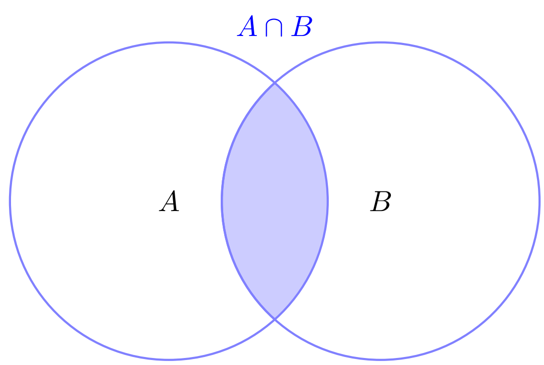
[14] *Complement* på engelsk.
[15] *Union* på engelsk



Mengdeunionen $A \cup B$ i lyseblått.

Snittet[16] av mengdene A og B er de elementene som er med i *begge mengdene*; formelt er snittet av to mengder A og B definert som

$$A \cap B = \{x \mid x \in A \text{ og } x \in B\}. \tag{28}$$



Mengdesnittet $A \cap B$ i lyseblått.

Mengdene A og B kalles *disjunkte*[17] dersom $A \cap B = \emptyset$, det vil si de har ingen elementer til felles.

Disse begrepene har sine opplagte generaliseringer til union og snitt av vilkårlig mange mengder. Vi vil komme tilbake til dette i §10.3.

Eksempel 2.1.12. Hvis $A = \{1, 2, 3, 4\}$ og $B = \{1, 3, 5, 7\}$, da er $A \cup B = \{1, 2, 3, 4, 5, 7\}$ og $A \cap B = \{1, 3\}$.

Eksempel 2.1.13. Vi har $\mathbb{R} = \mathbb{Q} \cup (\mathbb{R} \setminus \mathbb{Q})$ og unionen er disjunkt.

Vi oppsummerer en del mengdeidentiteter i tabellen nedenunder, hvor A, B og C er mengder, og U er en mengde som inneholder A, B og C.

[16] *Intersection* på engelsk
[17] *Disjoint* på engelsk.

Identitet	Navn
$A \cup \emptyset = A$ $A \cap U = A$	Identitet
$A \cap \emptyset = \emptyset$ $A \cup U = U$	Dominering
$A \cap A = A$ $A \cup A = A$	Idempotens
$A \setminus (B \setminus A) = A$	Komplementering
$A \cap B = B \cap A$ $A \cup B = B \cup A$	Kommutativitet
$(A \cap B) \cap C = A \cap (B \cap C)$ $(A \cup B) \cup C = A \cup (B \cup C)$	Assosiativitet
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	Distributivitet
$C \setminus (A \cup B) = (C \setminus A) \cap (C \setminus B)$ $C \setminus (A \cap B) = (C \setminus A) \cup (C \setminus B)$	De Morgans lover[18]
$A \cap (A \cup B) = A$ $A \cup (A \cap B) = A$	Absorbering
$A \cap (B \setminus A) = \emptyset$ $A \cup (B \setminus A) = A \cup B$	Komplement

Vi viser bare én av disse identitetene og overlater resten til Oppgave 2.4.

Setning 2.1.14

Hvis A, B, C er mengder, er $C \setminus (A \cap B) = (C \setminus A) \cup (C \setminus B)$.

BEVIS (ALTERNATIV I)

Vi vil vise at mengdene ovenfor er like ved å vise at hver er en delmengde av den andre, det vil si vi viser mengdelikheten $=$ ved å vise begge inklusjonene $\subset$ og $\supset$ hver for seg.

Vi viser først at $C \setminus (A \cap B) \subset (C \setminus A) \cup (C \setminus B)$. For å gjøre det, viser vi at hvis $x \in C \setminus (A \cap B)$, da er også $x \in (C \setminus A) \cup (C \setminus B)$.

Anta derfor at $x \in C \setminus (A \cap B)$. Per definisjon av mengdedifferanse, betyr dette at $x \in C$ og $x \notin A \cap B$. Det siste betyr at det *ikke* er sant at $x \in A \cap B$, med andre ord, per definisjon av snitt, at det *ikke* er sant at $x \in A$ og $x \in B$. Dette betyr at vi må ha $x \notin A$ eller $x \notin B$, med andre ord $x \in C \setminus A$ eller $x \in C \setminus B$. Per definisjon av union, er dette ekvivalent med $x \in (C \setminus A) \cup (C \setminus B)$, som er det vi skulle vise.

Vi viser dernest at $C \setminus (A \cap B) \supset (C \setminus A) \cup (C \setminus B)$. For å gjøre det, viser vi at hvis $x \in (C \setminus A) \cup (C \setminus B)$, da er også $x \in C \setminus (A \cap B)$.

Anta derfor at $x \in (C \setminus A) \cup (C \setminus B)$. Per definisjon av union betyr dette at $x \in (C \setminus A)$ eller $x \in (C \setminus B)$. Dette betyr at $x \in C$, og at $x \notin A$

[18] Etter den britiske matematikeren og logikeren Augustus De Morgan (1806–1871).

eller $x \notin B$. Dette vil si at x *ikke* kan være med i både A og B, med andre ord $x \notin (A \cap B)$. Altså er $x \in C \setminus (A \cap B)$, som vi skulle vise. □

Bevis (alternativ II)

Vi vil vise at mengdene ovenfor er like ved å eksplisitt skrive ned elementene i mengden og utlede en kjede av likheter:

$$\begin{aligned}
C \setminus (A \cap B) &= \{x \in C \mid x \notin A \cap B\} && \text{per def. av "}\setminus\text{"} \\
&= \{x \in C \mid \text{usant at } x \in A \cap B\} && \text{per def. av "}\notin\text{"} \\
&= \{x \in C \mid \text{usant at } [x \in A \text{ og } x \in B]\} && \text{per def. av "}\cap\text{"} \\
&= \{x \in C \mid x \notin A \text{ eller } x \notin B\} \\
&= \{x \mid x \in C \setminus A \text{ eller } x \in C \setminus B\} && \text{per def. av "}\notin\text{"} \\
&= \{x \mid x \in (C \setminus A) \cup (C \setminus B)\} && \text{per def. av "}\cup\text{"} \\
&= (C \setminus A) \cup (C \setminus B).
\end{aligned}$$

□

Det kartesiske produktet. Følgende er en definisjon som spiller en viktig rolle innenfor matematikk og mengdelære og som vi vil komme til å bruke mye:

Definisjon 2.1.15: Kartesisk produkt

La X and Y være mengder. Det *kartesiske produktet*[a] av X og Y, betegnet som $X \times Y$, er mengden av ordnede par[b] (x, y) slik at $x \in X$ og $y \in Y$. (At parene (x, y) er *ordnet* betyr at rekkefølgen spiller en rolle, det vi si, (x, y) er ikke det samme som (y, x).) Med andre ord,

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}.$$

Likeledes, om $X_1, \ldots, X_n$ er mengder, er det *kartesiske produktet* av mengdene, betegnet som $X_1 \times \cdots \times X_n$, mengden av alle ordnede n-tupler $(x_1, \ldots, x_n)$ slik at $x_i \in X_i$ for alle $i \in \{1, \ldots, n\}$.

Vi sier at $x_1, \ldots, x_n$ er *koordinatene* til punktet $(x_1, \ldots, x_n)$.

[a] *Cartesian product* på engelsk.

[b] *Ordered pairs* på engelsk.

Vi pleier ofte å skrive

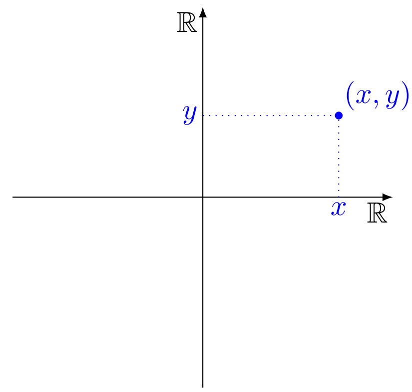
$$X^n = \underbrace{X \times \cdots \times X}_{n \text{ ganger}}.$$

Eksempel 2.1.16 (Planet). Mengden $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ er vi vant til å kalle *(det reelle) planet*[19]. Vi pleier å visualisere punkter (x, y) i et såkalt *koordinatsystem* hvor vi tenker oss to akser, én horisontal (gjerne kalt *abscisseaksen*[20]) og

[19] *Plane* på engelsk.

[20] Fra latin *abscisse*=avskåret linje.

én vertikal (gjerne kalt *ordinataksen*[21]), som hver representerer $\mathbb{R}$, og punktene (x, y) er tegnet inn slik at den vertikale projeksjonen på den horisontale aksen gir oss x og den horisontale projeksjonen på den vertikale aksen gir oss y:



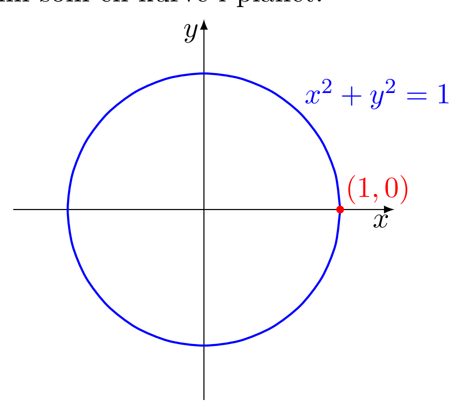
VISUALISERING AV ET PUNKT $(x, y) \in \mathbb{R}^2$ VED HJELP AV ET KOORDINATSYSTEM

Skjæringspunktet mellom aksene er punktet $(0, 0)$, som kalles *origo*[22].

Noe av det aller fineste med å visualisere $\mathbb{R}^2$ med et slikt koordinatsystem er at man kan beskrive delmengder av $\mathbb{R}^2$ av punkter som oppfyller en viss ligning som en kurve i planet. Det gjelder for eksempel mengden av alle punkter $(x, y) \in \mathbb{R}^2$ som oppfyller ligningen

$$x^2 + y^2 = 1.$$

"Geometrisk" er dette alle punkter i planet som er i avstand 1 fra origo, og disse kan vi tegne inn som en kurve i planet:



ENHETSSIRKELEN I $\mathbb{R}^2$

Denne kurven kalles *enhetssirkelen*[23]. Andre ligninger beskriver andre kurver. For eksempel beskriver ligningen

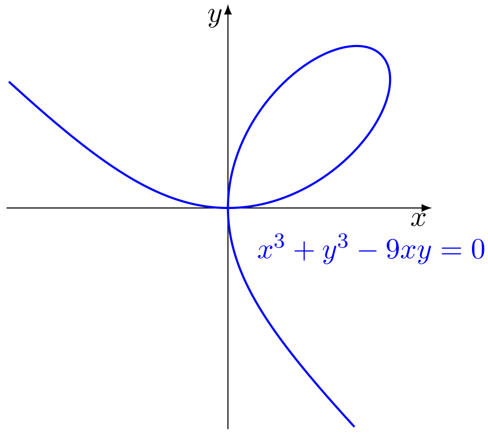
$$x^3 + y^3 - 9xy = 0$$

[21]Fra latin *linea ordinata*=ordnet linje.

[22]*Origin* på engelsk; ordet kommer fra latin "opprinnelse".

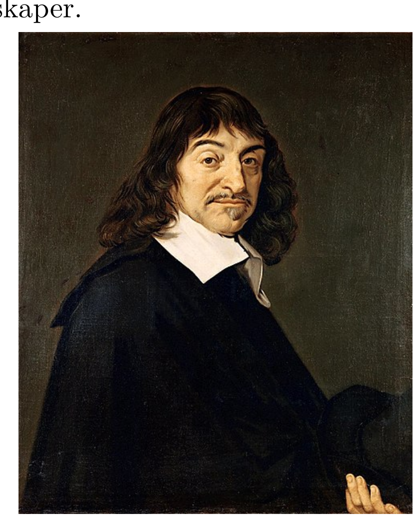
[23]*Unit circle* på engelsk.

følgende kurve i planet, som kalles *Descartes' blad*[24]:



Descartes' blad i $\mathbb{R}^2$

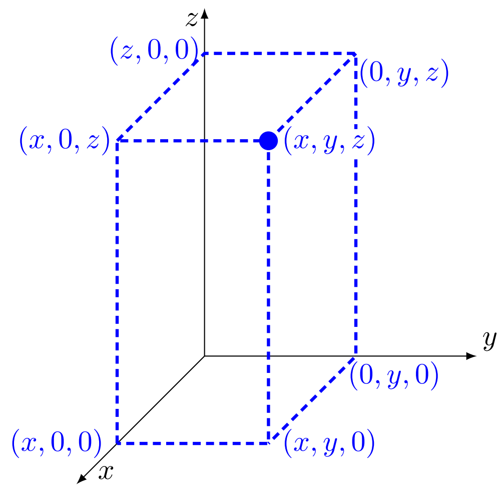
Samspillet mellom algebra, eller "analytiske uttrykk", og geometri går helt tilbake til de gamle grekerne, som ikke hadde et skarpt skille mellom algebra og geometri og oppfattet de aller fleste ligningene som geometrioppgaver. Grunnlaget for den *analytiske geometrien*, den delen av geometrien som undersøker geometriske egenskaper ved å bruke koordinater, ble lagt i 1637 gjennom to arbeider av Pierre de Fermat (1601–1665) og av den franske filosofen og matematikeren René Descartes (1596–1650). Planet kalles derfor også det *kartesiske planet*[25]. etter Descartes, på samme måte som *kartesiske koordinater* brukes for tallparet (x, y) som beskriver et punkt i planet. Analytisk geometri danner grunnlaget for moderne geometri, inkludert fagfelt som *algebraisk geometri* og *differensialgeometri*, og er også brukt som verktøy i alle naturvitenskaper.



René Descartes (Kilde: Wikipedia, public domain)

[24] *Folium of Descartes* på engelsk.
[25] *Cartesian plane* på engelsk.

EKSEMPEL 2.1.17 (Rommet). Mengden $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ er det vi er vant til å kalle *(det kartesiske) rommet*. Også i dette tilfellet pleier vi å visualisere punktene (x, y, z) i et *koordinatsystem*:



VISUALISERING AV ET PUNKT $(x, y, z) \in \mathbb{R}^3$ VED HJELP AV ET KOORDINATSYSTEM

Som i planet vil alle punkter (x, y, z) som oppfyller en bestemt ligning beskrive en delmengde av $\mathbb{R}^3$. For eksempel vil alle punktene som oppfyller ligningen

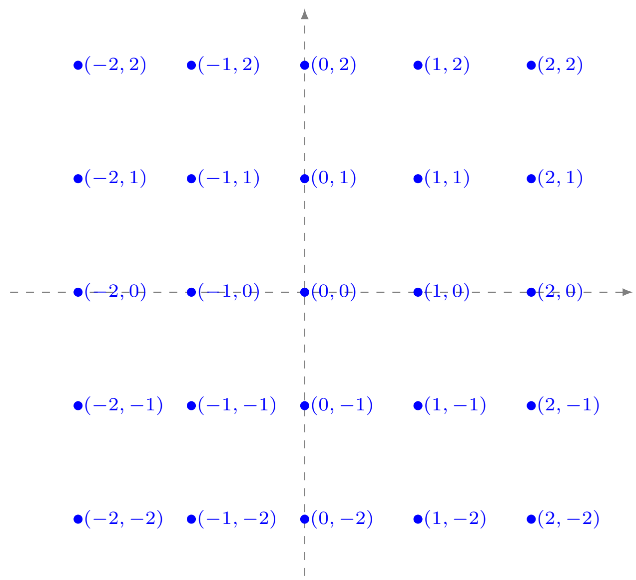
$$x^2 + y^2 + z^2 = 1$$

beskrive delmengden av alle punkter i $\mathbb{R}^3$ i avstand 1 fra origo, som kalles *enhetskulen*[26] eller *enhetssfæren*[27].

EKSEMPEL 2.1.18. $\mathbb{R} \times [0, \infty)$ er delmengden av $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ bestående av alle par (x, y) slik at $y \geq 0$. Dette er altså *øvre halvplan* (inkludert den horisontale aksen).

EKSEMPEL 2.1.19. $\mathbb{Z} \times \mathbb{Z} = \mathbb{Z}^2$ er delmengden av $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ bestående av alle par (x, y) av heltall. Dette kan visualiseres som et "gitter" i planet:

[26] *Unit ball* på engelsk.
[27] *Unit sphere* på engelsk.



Visualisering av $\mathbb{Z}^2$ i $\mathbb{R}^2$

Intervaller og omegner. Viktige delmengder av $\mathbb{R}$ som vi kjenner fra skolen er *intervaller*. Noen kalles *endelige*[28] eller *begrenset*[29] og er av følgende typer, der a og b er reelle tall slik at $a < b$:

$$\begin{aligned}
(a, b) &= \{x \in \mathbb{R} \mid a < x < b\} \\
[a, b] &= \{x \in \mathbb{R} \mid a \leq x \leq b\} \\
[a, b) &= \{x \in \mathbb{R} \mid a \leq x < b\} \\
(a, b] &= \{x \in \mathbb{R} \mid a < x \leq b\}.
\end{aligned}$$

I disse tilfellene kalles gjerne $b - a$ for *lengden* av intervallet og a og b for *endepunktene*[30] til intervallet.

Andre intervaller kalles *uendelige*[31], og betegnes ved hjelp av symbolet ∞ som leses som "uendelig"[32]:

$$\begin{aligned}
(-\infty, a) &= \{x \in \mathbb{R} \mid x < a\} \\
(-\infty, a] &= \{x \in \mathbb{R} \mid x \leq a\} \\
(a, \infty) &= \{x \in \mathbb{R} \mid a < x\} \\
[a, \infty) &= \{x \in \mathbb{R} \mid a \leq x\}.
\end{aligned}$$

Merk at notasjonen ovenfor overhodet ikke innebærer at $-\infty$ og ∞ betraktes som reelle tall. Spesielt har de fire siste intervallene kun a som endepunkt.

[28] *Finite* på engelsk.
[29] *Bounded* på engelsk.
[30] *Endpoints* på engelsk.
[31] *Infinite* på engelsk.
[32] *Infinity* på engelsk.

I alle tilfeller kan man lett overbevise seg om at et intervall er en delmengde av $\mathbb{R}$ som har egenskapen at *alle tall mellom to vilkårlige tall i mengden også er med i mengden.* Skrevet litt mer formelt:

Definisjon 2.1.20: Intervall

Et intervall I er en delmengde av $\mathbb{R}$ med følgende egenskap: dersom $x_1, x_2 \in I$, $x_1 < x_2$, og $x_1 < x < x_2$, da er $x \in I$.

Definisjonen inkluderer spesialtilfellet av *ettpunktsmengder* $\{a\} \subset \mathbb{R}$, for $a \in \mathbb{R}$, siden betingelsen er tom i dette tilfellet. (For øvrig kan vi uttrykke $\{a\}$ som $[a, a]$, selv om det er litt uvanlig.) Legg også merke til at $\mathbb{R}$ selv er et intervall, som også kan skrives som:

$$\mathbb{R} = (-\infty, \infty).$$

Et intervall kan inneholde sine endepunkter eller ikke. Dersom alle endepunktene til et intervall er med i intervallet, kalles intervallet *lukket*[33]; dette er oppfylt for intervallene $[a, b]$, $(-\infty, a]$ og $[a, \infty)$ ovenfor. Dersom alle endepunktene til et intervall ligger *utenfor* intervallet, kalles intervallet *åpent*[34]; dette er oppfylt for intervallene (a, b), $(-\infty, a)$ og (a, ∞). De to typene $[a, b)$ og $(a, b]$ er verken åpne eller lukkede, siden de har ett endepunkt som ligger i intervallet og ett som ligger utenfor; disse intervallene kalles gjerne for *halvåpne*. Intervallet $\mathbb{R} = (-\infty, \infty)$ har ikke noen endepunkter og er derfor både åpent og lukket samtidig: det er nemlig både riktig å si at alle endepunktene ligger i intervallet og at alle endepunktene ligger utenfor intervallet, siden endepunkter ikke finnes i dette tilfellet.

MERKNAD 2.1.21. Noen lærebøker bruker symbolene

- $\langle$" eller "]" istedenfor "(",
- '$\rangle$" eller "[" istedenfor ")",
- "$\leftarrow$" og "$\rightarrow$" istedenfor $-\infty$ og ∞, henholdsvis.

Dette er ikke en veldig utbredt notasjon ellers, så vi kommer til å holde oss til notasjonen nevnt tidligere. Eneste ulempen med denne notasjonen er at man kan forveksle *intervallet* (a, b) med *punktet* (a, b) i planet, slik at vi må være klar over hva vi snakker om.

Vi minner om at *absoluttverdien* $|x|$ til et reelt tall x er (den ikke-negative) *avstanden mellom tallet x og origo* på tallinjen $\mathbb{R}$, med andre ord

$$|x| = \begin{cases} x, & x \geq 0 \\ -x & x < 0. \end{cases} \tag{29}$$

Avstanden mellom to punkter på tallinjen a og b kan ved hjelp av absoluttverdi uttrykkes som $|a - b| = |b - a|$. Avstandsbegrepet brukes når vi

[33] *Closed* på engelsk.

[34] *Open* på engelsk.

er interessert i hva som skjer i *nærheten* av et tall/punkt $a \in \mathbb{R}$. (Det kan for eksempel dreie seg om oppførselen til en funksjon (jf. §2.2).) Da holder det å konsentrere seg om små, *symmetriske* åpne intervaller om punktet a. Dersom vi går ut en avstand $\delta > 0$ på begge sider av a (hvor δ står for den greske bokstaven "delta") får vi intervaller på formen $(a-\delta, a+\delta)$, som har fått et eget navn:

Definisjon 2.1.22: Omegn

La $a \in \mathbb{R}$. Et intervall på formen

$$(a-\delta, a+\delta),$$

der $\delta \in \mathbb{R}$ og $\delta > 0$, kalles *en omegn*[a] *om* a *(av radius* δ*)* eller *en* δ*-omegn om* a.

[a] *Neighborhood* på engelsk.

Merk at en δ-omegn om a består av alle punkter som ligger i avstand mindre enn δ fra a. Vi kan altså skrive

$$(a-\delta, a+\delta) = \{x \in \mathbb{R} \mid |x-a| < \delta\}.$$

Dersom vi "tar bort" punktet a (som vi gjør dersom vi er interessert i å studere oppførselen til en funksjon *rundt*, men ikke *i*, punktet a), får vi mengden

$$(a-\delta, a+\delta) \setminus \{a\} = \{x \in \mathbb{R} \mid 0 < |x-a| < \delta\},$$

hvor vi minner om at notasjonen "$\setminus$" betyr "bortsett fra". Dersom vi bruker notasjonen "$\cup$" for "union av mengder", definert i (27), kan vi også skrive

$$(a-\delta, a+\delta) \setminus \{a\} = (a-\delta, a) \cup (a, a+\delta).$$

En slik mengde har også fått et eget navn:

Definisjon 2.1.23: Punktert omegn

La $a \in \mathbb{R}$. En mengde på formen

$$(a-\delta, a+\delta) \setminus \{a\} = (a-\delta, a) \cup (a, a+\delta) = \{x \in \mathbb{R} \mid 0 < |x-a| < \delta\}$$

der $\delta \in \mathbb{R}$ og $\delta > 0$, kalles *en punktert omegn*[a] *om* a *(av radius* δ*)* eller *en punktert* δ*-omegn om* a. Noen ganger brukes også notasjonen

$$(a-\delta, a+\delta) \setminus \{a\} = (a-\delta, a+\delta)^*.$$

[a] *Punctured neighborhood* på engelsk.

MERKNAD 2.1.24. Noen lærebøker og matematikere bruker betegnelsen "omegn om a" for ethvert åpent intervall som innholder a, det vil si et intervall på formen $(a-\delta_1, a+\delta_2)$ der $\delta_1 > 0$ og $\delta_2 > 0$ ikke nødvendigvis er like. Det er i praksis uten betydning hvilken definisjon vi bruker.

2.2. Funksjoner

Begrepet av en *funksjon* mellom to mengder er svært viktig. Vi skriver ned definisjonen og bruker en notasjon som er vanlig i mange lærebøker og i matematikk generelt, men som for eksempel [**AE**] dessverre ikke innfører:

Definisjon 2.2.1: Funksjon/avbildning

En *funksjon*[a] eller *avbildning*[b] f fra en mengde X til en mengde Y, betegnet med

$$f : X \longrightarrow Y,$$

er en regel som tilordner til ethvert element x i X et *entydig* element i Y, som vi kaller $f(x)$.

Vi bruker også notasjonen

$$X \xrightarrow{f} Y$$

og skriver gjerne

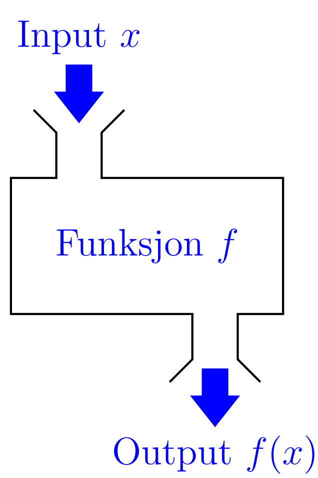
$$\begin{aligned} f : X &\longrightarrow Y \\ x &\mapsto f(x) \end{aligned}$$

for å spesifisere hva regelen er.

Mengden X kalles *definisjonsmengden*[c] eller *domenet* til funksjonen, og betegnes ofte med $D(f)$ eller D_f. Mengden Y kalles *kodomenet*[d] til funksjonen.

[a] *Function* på engelsk.
[b] *Map* eller *mapping* på engelsk.
[c] *Domain* på engelsk.
[d] *Codomain* på engelsk.

En funksjon forteller oss altså hvilket tall/element i Y vi får ut ("Output") når vi setter inn et tall/element i X ("Input"):



EN FUNKSJON SIER OSS HVA "OUTPUT" $f(x)$ ER FOR HVER "INPUT" x

Faren til funksjonsbegrepet som i Definisjon 2.2.1 regnes for å være den tyske matematikeren Peter Gustav Lejeune Dirichlet (1805–1859). Interessant nok er denne definisjonen fremdeles litt unøyaktig sett med dagens øyne, siden det ikke er klart hva som menes med en "regel" eller en "tilordning": vi skal se på en mer formell definisjon i Definisjon 2.4.5.

Peter Gustav Lejeune Dirichlet (Kilde: Wikipedia, public domain)

Eksempel 2.2.2. La S være mengden av studenter som tar en bestemt eksamen ved Universitetet i Bergen, hvor det gis bokstavkarakter. Da finnes en naturlig funksjon

$$f : S \to \{A, B, C, D, E, F\}$$

som assosierer til en student den karakteren han eller hun får på eksamen.

Det er et krav til en funksjon at ethvert element i X skal tilordnes et element i Y, og – veldig viktig – et *entydig* element i Y. Men det er ikke noe krav at ethvert element $y \in Y$ skal ha tilordnet et element $x \in X$ slik at $f(x) = y$. Mengden bestående av de elementene i Y som har minst ett slikt element i X tilordnet, har fått et eget navn:

Definisjon 2.2.3: Verdimengde

Verdimengden[a] til en funksjon $f : X = D(f) \longrightarrow Y$ er

$$V(f) = \{f(x) \mid x \in X = D(f)\}.$$

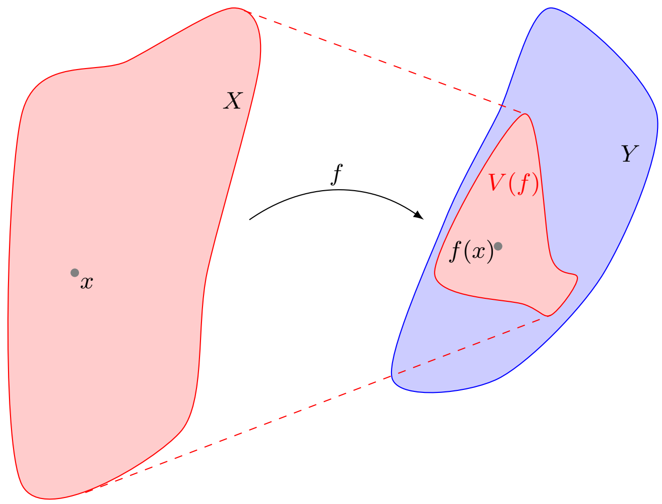
(Noen ganger brukes betegnelsene $R(f)$ og $\mathrm{im}(f)$.)

[a] *Range* på engelsk.

Merk at $V(f) \subset Y$. For en funksjon $f : D(f) \to Y$ kan vi altså også skrive

$$f : D(f) \longrightarrow V(f),$$

men dette er strengt tatt en annen funksjon, siden kodomenet er en del av dataene som inngår i definisjonen til en funksjon.



VERDIMENGDEN $V(f)$ TIL $f : X \to Y$ MARKERT I RØDT.

Mer generelt har vi:

Definisjon 2.2.4: Bilde og inversbilde

La $f : X \longrightarrow Y$ være en funksjon.

- Hvis $A \subset X$, definerer vi mengden

$$f(A) = \{y \in Y \mid y = f(x) \text{ for en } x \in A\}$$

og kaller den *bildet*[a] *av* A *under/ved* f.

- Hvis $B \subset Y$, definerer vi mengden

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\}$$

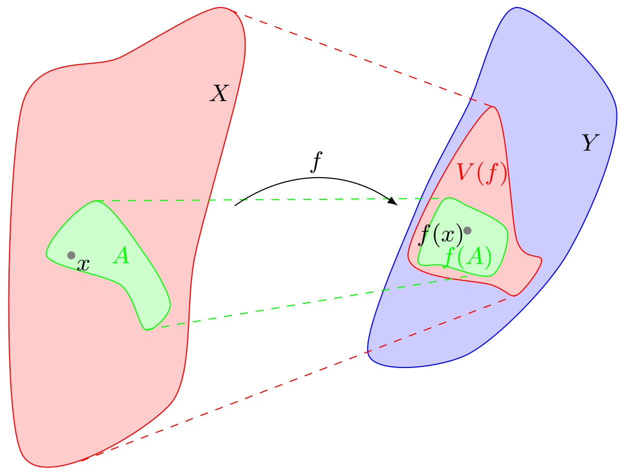
og kaller den *det inverse bildet*[b] *av* B *under/ved* f.

[a] *Image* på engelsk.
[b] *Inverse image* på engelsk.

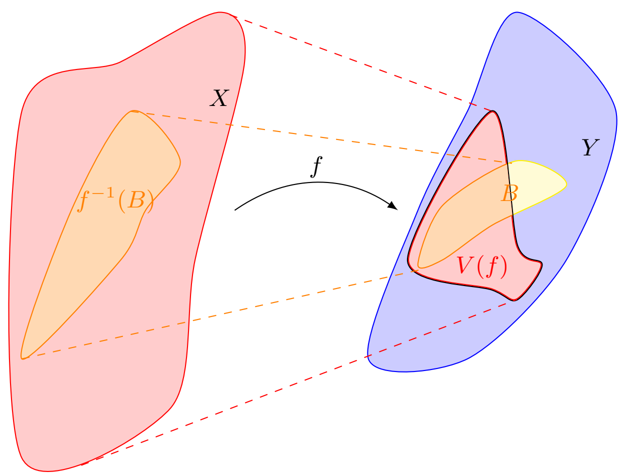
Med notasjonen ovenfor er med andre ord $V(f) = f(X) = f(D(f))$ og $X = f^{-1}(Y)$. Dessuten er $V(f) = \{y \in Y \mid f^{-1}(\{y\}) \neq \emptyset\}$.

MERKNAD 2.2.5. Notasjonen f^{-1} i forbindelse med mengder som i forrige definisjon må ikke forveksles med notasjonen f^{-1} for den inverse funksjonen til f som i Definisjon 2.3.5 nedenunder.

Følgende figurer visualiserer mengdene vi har definert:



Delmengden $A \subset X$ og dens bilde $f(A) \subset Y$ ved f markert i grønt.



Delmengden $B \subset Y$ og dens inversbilde $f^{-1}(B) \subset X$ markert i gult og oransje.

I den nederste figuren til venstre er den gule mengden lik den delen av B som faller utenfor bildet $V(f)$, mens den resterende delen av B er tegnet i oransje og er følgelig lik $B \cap V(f)$. Det er ikke vanskelig å se at $f^{-1}(B) = f^{-1}(B \cap V(f))$ (jf. Oppgave 2.15), som antydet på figuren.

Eksempel 2.2.6. La S være funksjonen fra eksempel 2.2.2. At ingen stryker på eksamen er ekvivalent med utsagnet $F \notin V(S)$. Mengden $S^{-1}(A)$ er mengden av studenter som får karakteren A.

Eksempel 2.2.7. La $f : \mathbb{R} \to \mathbb{R}$ være *absoluttverdifunksjonen*, nemlig funksjonen gitt ved

$$f(x) = |x|.$$

Her er verdimengden

$$V(f) = [0, \infty)$$

og vi kunne like gjerne skrevet $f : \mathbb{R} \to [0, \infty)$, dersom vi ønsket å være mer presise. Merk også at bortsett fra 0, så er det alltid *to* elementer fra definisjonsmengden som sendes på samme element y i verdimengden, nemlig $-y$ og y. med andre ord har vi for alle $y > 0$ at $f^{-1}(\{y\}) = \{-y, y\}$.

Som ytterligere eksempler på notasjonene ovenfor har vi at $f((-2, \frac{1}{2})) = [0, 2)$ og $f^{-1}([0, 1)) = (-1, 1)$.

EKSEMPEL 2.2.8. La $f : \mathbb{R} \to \mathbb{R}$ være sinusfunksjonen definert ved $f(x) = \sin x$ (jf. [**AE**, §P.7]). Da er $V(f) = [-1, 1]$, slik at vi også kan skrive $f : \mathbb{R} \to [-1, 1]$. Vi har for eksempel $f^{-1}(\{0\}) = \{n\pi \mid n \in \mathbb{Z}\}$, $f([0, \frac{\pi}{2})) = [0, 1)$ og $f([0, \pi)) = [0, 1]$.

De funksjonene vi er vant til fra skolen, og de funksjonene som studeres i MAT111, har både definisjonsmengde og verdimengde i $\mathbb{R}$ (med andre ord plugger vi inn en "input"-verdi som er et reelt tall, og får ut en "output"-verdi som er et reelt tall). Vanligvis er definisjonsmengden sogar et intervall eller en union av intervaller. De fleste er nok vant til å tenke på slike funksjoner som en formel av typen

$$f(x) = x \cos x - \sqrt[3]{\sin x}, \tag{30}$$

eller som i Eksempel 2.2.7, det vil si at en funksjon er gitt ved en formel for "output"-verdien $f(x)$ for en vilkårlig "input"-verdi x, som gjerne også kalles *variabelen* i funksjonsuttrykket. Det finnes imidlertid også funksjoner som er definert med to (eller flere) formler, som for eksempel

$$f(x) = \begin{cases} -\frac{1}{4}x^2 - 1, & x < 2 \\ 2x - 8, & x \geq 2, \end{cases} \tag{31}$$

eller

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & x \neq 0 \\ 0 & x = 0, \end{cases} \tag{32}$$

eller ved hjelp av ord, som

$$f(x) = \begin{cases} 1, & x \text{ rasjonal;} \\ 0, & x \text{ irrasjonal} \end{cases} \tag{33}$$

(dette er den berømte *Dirichlet-funksjonen*, se Eksempel 9.3.2). Når f er definert ved (ett eller flere) formeluttrykk som ovenfor, og det er klart at vi jobber innenfor mengden av reelle tall, vil vi ofte la være å spesifisere definisjonsmengden. Underforstått er da at funksjonen er definert for alle $x \in \mathbb{R}$ hvor formelen gir mening, jf. [**AE**, "The domain convention" i §P.4].

for eksempel er alle funksjonene f i (30)-(33) definert for alle $x \in \mathbb{R}$, slik at $D(f) = \mathbb{R}$ og vi skriver $f : \mathbb{R} \to \mathbb{R}$. Når vi imidlertid sier at vi betrakter funksjonen

$$g(x) = \frac{\cos x}{x-2} - \ln x,$$

så gir dette mening kun når $x \neq 2$ (siden vi ikke kan dele på null) og $x > 0$ (siden logaritmen kun er definert for positive tall). Altså mener vi underforstått at definisjonsmengden er

$$D(g) = (0, 2) \cup (2, \infty).$$

MERKNAD 2.2.9. Funksjoner som er gitt ved to eller flere uttrykk som (31)-(33) sies noen ganger å ha "delt forskrift". Det er interessant å merke seg at matematikeren som regnes for å ha innført funksjonsbegrepet, nemlig den sveitsiske matematikeren Leonhard Euler (1707–1783), i utgangspunktet *ikke* godtok disse som funksjoner, siden han krevde at funksjonen skulle gis ved ett analytisk uttrykk. Han godtok uendelige summer (som for eksempel Weierstrass-funksjonen (26)) og uendelige produkter (som for eksempel *zeta*-funksjonen (24)-(25)), og til og med uttrykk som brøt med entydigheten i Definisjon 2.2.1; for eksempel ville han betrakte uttrykket $y^2 = x^2$ som en funksjon med to forskjellige verdier $y = \pm x$. Men han godtok imidlertid i utgangspunktet ikke funksjoner med delt forskrift. Senere i karrieren godtok Euler funksjoner med delt forskrift, blant annet fordi en del interessante funksjoner som beskriver amplituden til en svingende streng kun kan beskrives ved å starte med ett uttrykk på et bestemt intervall og så "speile" dette uttrykket videre langs hele den reelle tallinjen.

LEONHARD EULER (KILDE: WIKIPEDIA, PUBLIC DOMAIN)

Legg også merke til at hvis vi har en funksjon $f : X \to Y$ og $A \subset X$ er en delmengde, så har vi også automatisk definert en funksjon $A \to Y$ ved samme regel. Denne betegnes gjerne med et eget symbol og navn:

Definisjon 2.2.10: Restriksjon av funksjon

Hvis $f : X \to Y$ er en funksjon og $A \subset X$ er en delmengde, definerer vi *restriksjonen*[a] *av f til A* til å være funksjonen

$$\begin{aligned} f|_A : A &\longrightarrow Y \\ x &\mapsto f(x). \end{aligned}$$

[a] *Restriction* på engelsk.

Merknad 2.2.11. Strengt tatt er det dårlig matematisk språkbruk å si for eksempel at "$\cos x$ er en funksjon". Det riktige er å si at "regelen eller tilordningen $x \mapsto \cos x$ er en funksjon". For å unngå unødvendig tungt og langt språk er det likevel ikke uvanlig å ty til den ikke fullt korrekte språkbruken selv i matematiske tekster.

Fra skolen er vi også vant til å jobbe med *sammensatte funksjoner*, selv om vi kanskje ikke har sett den matematiske definisjonen før:

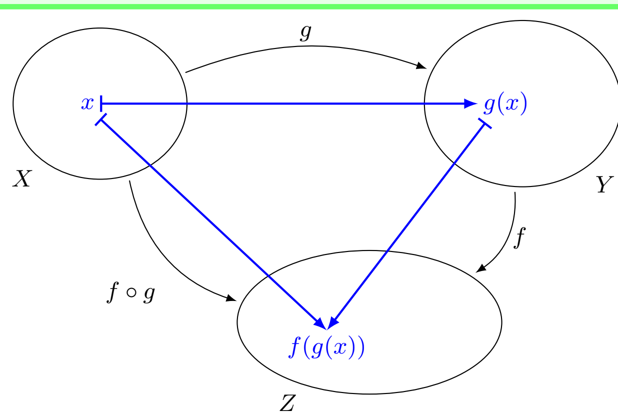
Definisjon 2.2.12: Sammensetning av funksjoner

Dersom $g : X \to Y$ og $f : Y \to Z$ er funksjoner mellom mengder, definerer vi *sammensetningen*[a] av f og g til å være funksjonen

$$\begin{aligned} f \circ g : X &\longrightarrow Z \\ x &\mapsto f(g(x)). \end{aligned}$$

Funksjonen $f \circ g$ kalles en *sammensatt funksjon*[b].

[a] *Composition* på engelsk.
[b] *Composite function* på engelsk.

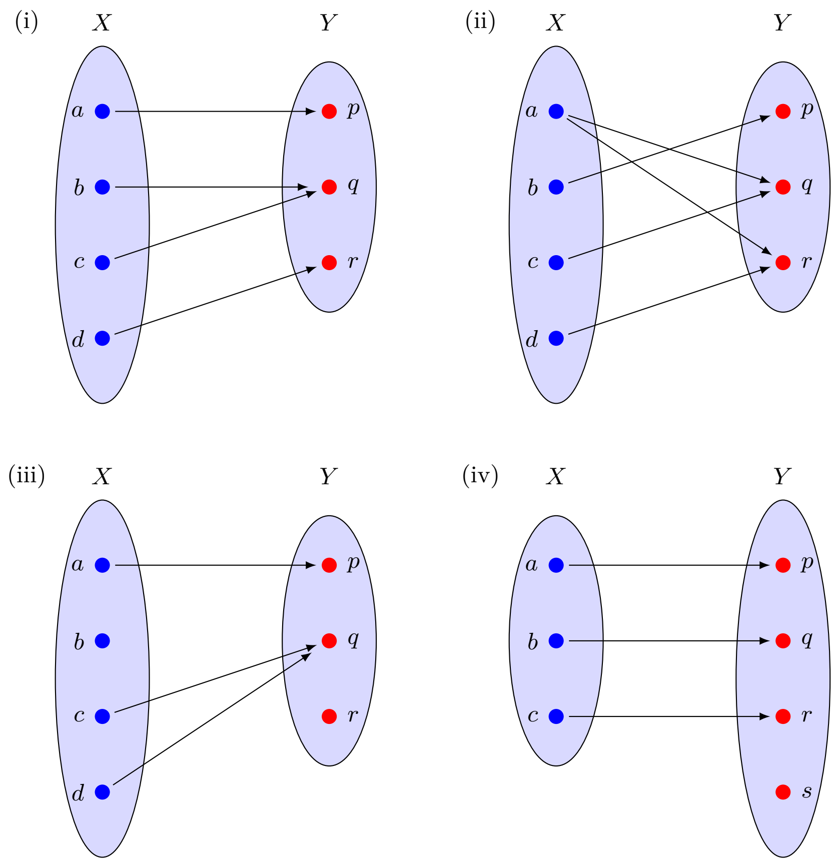


Funksjonen $f \circ g$ tar x til $f(g(x))$

Eksempel 2.2.13. Hvis $f : [0, \infty) \to \mathbb{R}$ er definert ved $f(x) = \sqrt{x}$ og $g : \mathbb{R} \to [0, 2]$ er definert ved $g(x) = \cos x + 1$, er $f \circ g : \mathbb{R} \to \mathbb{R}$ definert ved $f \circ g(x) = f(g(x)) = \sqrt{\cos x + 1}$, mens $g \circ f : [0, \infty) \to [0, 2]$ er definert ved $g \circ f(x) = g(f(x)) = \cos \sqrt{x} + 1$.

La oss se på et eksempel på funksjoner mellom *endelige mengder*, det vil si mengder som inneholder kun endelig mange elementer. Slike funksjoner er viktige spesielt innenfor informatikk og tallteori.

Eksempel 2.2.14. Figurene (i)-(iv) nedenunder angir regler som tilordner elementer i en mengde Y til elementer i en mengde X, men kun to av dem er funksjoner fra X til Y:



La oss se på hvilke to som er funksjoner fra X til Y.

(i) Her er $X = \{a, b, c, d\}$ og $Y = \{p, q, r\}$ og figuren definerer en funksjon $f : X \to Y$ ved at

$$f(a) = p, \; f(b) = f(c) = q, \; f(d) = r.$$

Verdimengden til funksjonen er hele Y, siden alle elementene i Y blir "truffet" av pilene; altså $V(f) = \{p, q, r\} = Y$. Vi merker oss at det er *to forskjellige* elementer fra X, nemlig b og c som "sendes på" samme element $q \in Y$, det vil si slik at $f(b) = f(c)$, eller ekvivalent $f^{-1}(\{q\}) = \{b, c\}$. Dette er fullt tillatt og bryter ikke med definisjonen til en funksjon. Definisjonen krever nemlig kun at det ikke finnes to elementer i Y som "kommer fra" samme element i X.

(ii) Her er igjen $X = \{a, b, c, d\}$ og $Y = \{p, q, r\}$, men figuren definerer *ikke* en funksjon, siden regelen tilordner *to* elementer $q \in Y$ og $r \in Y$ til det *ene* elementet $a \in X$. Det er *entydigheten* i Definisjon 2.2.1 som ikke er oppfylt.

(iii) Her er igjen $X = \{a, b, c, d\}$ og $Y = \{p, q, r\}$, men figuren definerer *ikke* en funksjon $f : X \to Y$ siden den ikke forteller oss hvilket element i Y vi skal tilordne $b \in X$. Figuren angir en funksjon $f : \{a, c, d\} = X \setminus \{b\} \to Y$, men denne funksjonen er *ikke* definert i b.

(iv) Her er $X = \{a, b, c\}$ og $Y = \{p, q, r, s\}$ og figuren definerer en funksjon $f : X \to Y$ ved at

$$f(a) = p,\ f(b) = q,\ f(c) = r.$$

Verdimengden til funksjonen er de elementene i Y som faktisk blir "truffet" av pilene, nemlig $V(f) = \{p, q, r\} = Y \setminus \{s\}$.

Grafen til en funksjon. Vi bruker det kartesiske produktet til å definere grafen til en funksjon:

Definisjon 2.2.15: Grafen til en funksjon

Grafen[a] $\Gamma(f)$ til en funksjon $f : X \to Y$ er mengden av punkter $(x, y) \in X \times Y$ som oppfyller at $x \in X$ og $y = f(x)$, med andre ord

$$\Gamma(f) = \{(x, y) \in X \times Y \mid x \in X,\ y = f(x)\},$$

eller, skrevet på en ekvivalent måte,

$$\Gamma(f) = \{(x, f(x)) \in X \times Y \mid x \in X\}.$$

[a] *Graph* på engelsk.

Bokstaven Γ brukt for grafen i definisjonen er den greske store bokstaven "gamma". (Den tilsvarende lille bokstaven er γ.)

Eksempel 2.2.16. (i) Grafen til funksjonen i Eksempel 2.2.14(i) er

$$\{(a, p), (b, q), (c, q), (d, r)\} \subset \{a, b, c, d\} \times \{p, q, r\}.$$

(ii) Grafen til funksjonen i Eksempel 2.2.14(iv) er

$$\{(a,p),(b,q),(c,r)\} \subset \{a,b,c\} \times \{p,q,r,s\}.$$

(iii) Grafen til funksjonen i Eksempel 2.2.7 er

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

(iv) Grafen til funksjonen f i Eksempel 2.2.8 er

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

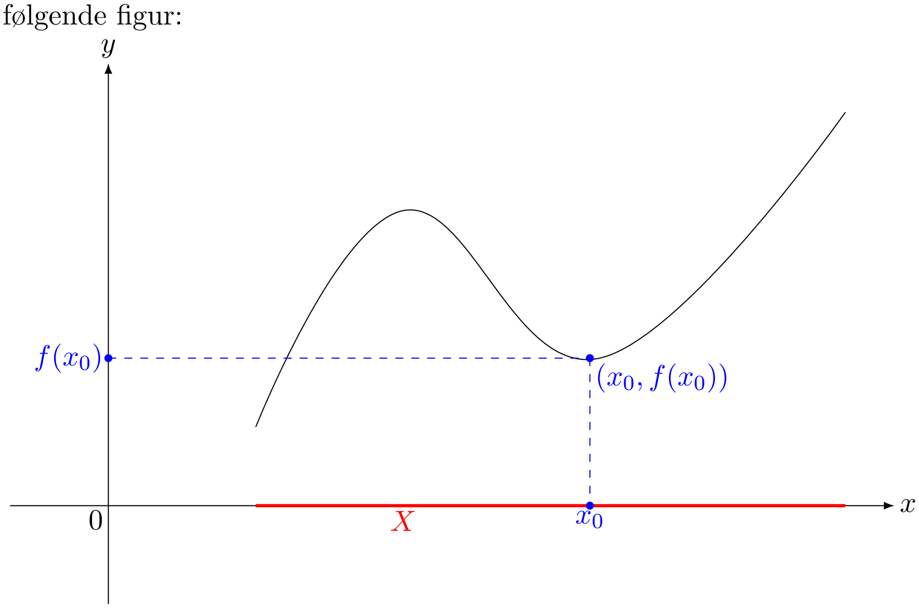
mens grafen til funksjonen g i Eksempel 2.2.8, som er restriksjonen av f til intervallet $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, er

$$\left\{(x,\sin x) \mid x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]\right\} \subset \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \times \mathbb{R}.$$

Fra skolen, der vi jobbet med funksjoner som hadde definisjonsmengde og verdimengde i $\mathbb{R}$, er vi vant til å tegne/skissere grafen til f eller la en kalkulator gjøre det, som betyr å merke alle punkter (x,y) i planet $\mathbb{R}^2$ som oppfyller ligningen

$$y = f(x)$$

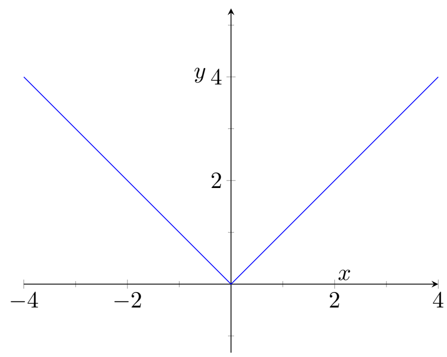
(jf. Eksempel 2.1.16). Vi får dermed en *kurve i planet* $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, som vist i følgende figur:



Grafen til en vilkårlig funksjon $f : X \to \mathbb{R}$, med $X \subset \mathbb{R}$, tegnet i planet

Det vi er vant til å kalle en graf i skolematematikken er altså kurven som gir den "geometriske visualiseringen" av grafen definert mengdeteoretisk ovenfor.

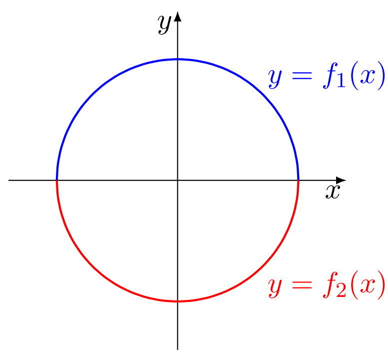
Eksempel 2.2.17. Grafen til absoluttverdifunksjonen i Eksempel 2.2.7, beskrevet i Eksempel 2.2.16(iii), ser slik ut:



GRAFEN TIL ABSOLUTTVERDIFUNKSJONEN

Som vist på de to siste figurene: starter vi med en x-verdi (i $D(f)$) og tegner den vertikale linjen som skjærer x-aksen i denne verdien, vil vi treffe grafen i nøyaktig ett punkt, nemlig $(x, f(x))$. Tegner vi den horisontale linjen gjennom dette punktet, vil den treffe y-aksen i punktet $f(x)$. Verdimengden til f kan vi "lese av" på y-aksen ved å projisere grafen horisontalt. Definisjonsmengden til f kan vi "lese av" på x-aksen ved å projisere grafen vertikalt.

EKSEMPEL 2.2.18. *Enhetssirkelen* gitt ved alle punkter (x, y) i planet som oppfyller ligningen $x^2 + y^2 = 1$ (jf. Eksempel 2.1.16) er *ikke* grafen til en funksjon, siden *to* y-verdier hører til samme x-verdi, nemlig $y = \pm\sqrt{1 - x^2}$. Imidlertid kan vi dele denne kurven i to deler, som hver er grafen til en funksjon, nemlig den øverste halvsirkelen, som er grafen til funksjonen $f_1 : [-1, 1] \to \mathbb{R}$ definert ved $f_1(x) = \sqrt{1 - x^2}$, og den nederste halvsirkelen, som er grafen til funksjonen $f_2 : [-1, 1] \to \mathbb{R}$ definert ved $f_2(x) = -\sqrt{1 - x^2}$.



ENHETSSIRKELEN BESKREVET SOM GRAFEN TIL TO FUNKSJONER

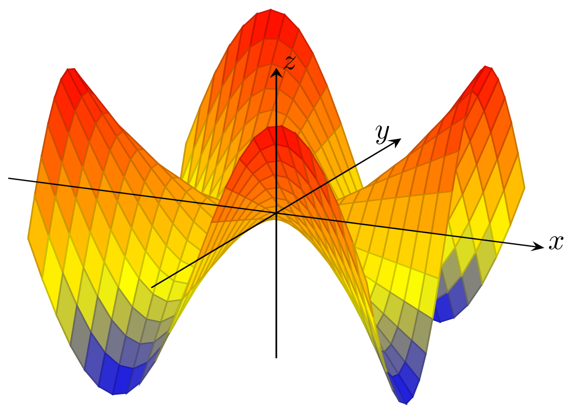
EKSEMPEL 2.2.19. Husk fra Eksempel 2.1.16 at $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ er mengden av alle ordnede par (x, y) med $x, y \in \mathbb{R}$. La $f : \mathbb{R}^2 \setminus \{0, 0\} \to \mathbb{R}$ være funksjonen gitt ved $f(x, y) = \frac{2xy(x^2-y^2)}{x^2+y^2}$. Grafen $\Gamma(f)$ til denne funksjonen er mengden av alle punkter (x, y, z) i $\left(\mathbb{R}^2 \setminus \{0,0\}\right) \times \mathbb{R} \subset \mathbb{R}^3$ ("rommet") som oppfyller

$$z = f(x, y),$$

med andre ord

$$\Gamma(f) = \left\{ (x, y, z) \in \left(\mathbb{R}^2 \setminus \{0,0\}\right) \times \mathbb{R} \;\middle|\; z = \frac{2xy(x^2 - y^2)}{x^2 + y^2} \right\} \subset \mathbb{R}^3.$$

Tegner vi $\Gamma(f)$ inn i rommet $\mathbb{R}^3$, får vi følgende flate:



Grafen til funksjonen i Eksempel 2.2.19

2.3. Injektive, surjektive og bijektive funksjoner

Vi har sett både i Eksempel 2.2.7 og i Eksempel 2.2.14(i) at det godt kan finnes *forskjellige elementer* fra definisjonsmengden $D(f)$ som "sendes på" samme element i verdimengden, det vil si

(34) det finnes $x_1, x_2 \in D(f)$ slik at $x_1 \neq x_2$, men $f(x_1) = f(x_2)$.

Som nevnt tidligere, er dette fullt tillatt og bryter ikke med definisjonen til en funksjon. Merk at det samme ikke skjer i Eksempel 2.2.14(iv), hvor forskjellige elementer i definisjonsmengden X alltid sendes på forskjellige elementer i verdimengden. Dette er en såpass pen egenskap at den har fått et eget navn:

Definisjon 2.3.1: Injektiv funksjon

En funksjon $f : X \to Y$ er *én-til-én*[a] eller *injektiv*[b] dersom vi har at $f(x_1) \neq f(x_2)$ for alle $x_1, x_2 \in D(f)$ slik at $x_1 \neq x_2$. (Med andre ord er f én-til-én dersom det *ikke* finnes punkter som oppfyller (34).)

En injektiv funksjon kalles også en *injeksjon*[c].

Noen ganger brukes pilen $\hookrightarrow$ for en injektiv funksjon.

[a] *One-to-one* på engelsk.
[b] *Injective* på engelsk.
[c] *Injecion* på engelsk.

MERKNAD 2.3.2. Vi kan også si at en funksjon $f : X \to Y$ er injektiv dersom vi for alle $y \in V(f)$ har at det inverse bildet $f^{-1}(\{y\})$ består av ett eneste element fra X.

Vi har dessuten sett både i Eksempel 2.2.14 og i Eksempel 2.2.7 at verdimengden til en funksjon $f : D(f) \to Y$ både kan være hele Y og ikke. Dette leder til neste definisjon:

Definisjon 2.3.3: Surjektiv funksjon

En funksjon $f : X \to Y$ er *på*[a] eller *surjektiv*[b] dersom $V(f) = Y$, med andre ord det finnes til enhver $y \in Y$ et element $x \in D(f)$ slik at $f(x) = y$.

En surjektiv funksjon kalles også en *surjeksjon*[c].

Noen ganger brukes pilen $\twoheadrightarrow$ for en surjektiv funksjon.

[a] *Onto* på engelsk.
[b] *Surjective* på engelsk.
[c] *Surjection* på engelsk.

En funksjon som både er injektiv og surjektiv har også fått et eget navn.

Definisjon 2.3.4: Bijektiv funksjon

En funksjon $f : X \to Y$ er en *én-til-én korrespondanse*[a] eller *bijektiv*[b] dersom den er både injektiv og surjektiv.

I dette tilfellet sies X og Y å *være i én-til-én korrespondanse* eller *ha samme kardinalitet*[c].

En bijektiv funksjon kalles også en *bijeksjon*[d].

[a] *One-to-one correspondence* på engelsk.
[b] *Bijective* på engelsk.
[c] *Cardinality* på engelsk.
[d] *Bijecion* på engelsk.

For en slik funksjon kan vi definere en "funksjon som gjør det motsatte":

Definisjon 2.3.5: Invers funksjon

Hvis f er bijektiv, defineres *den inverse funksjonen* eller *omvendte funksjonen*:

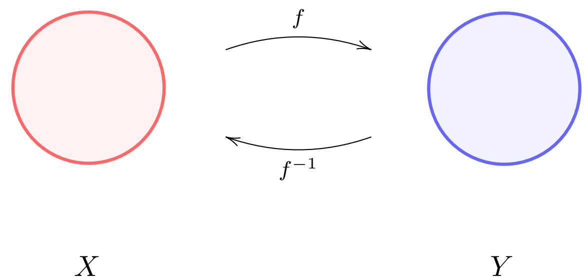
$$f^{-1} : Y \longrightarrow X,$$

ved regelen

$$f^{-1}(y) = \text{det entydige elementet } x \in X \text{ slik at } f(x) = y,$$

med andre ord

$$f^{-1}(y) = x \iff f(x) = y. \tag{35}$$



INVERSE FUNKSJONER

Det er klart fra definisjonen at

$$\begin{array}{rcl} f^{-1}\circ f : X & \longrightarrow & X \\ x & \mapsto & x \end{array}$$

og

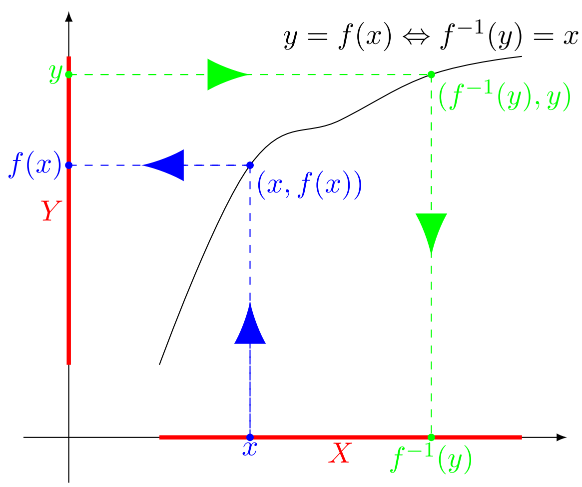
$$\begin{array}{rcl} f\circ f^{-1} : Y & \longrightarrow & Y \\ y & \mapsto & y. \end{array}$$

Slike funksjoner, som avbilder elementer på seg selv, kalles *identitets-funksjoner* eller *identitetsavbildninger*. Dessuten har vi at f^{-1} har f som sin inversfunksjon:

$$\left(f^{-1}\right)^{-1} = f.$$

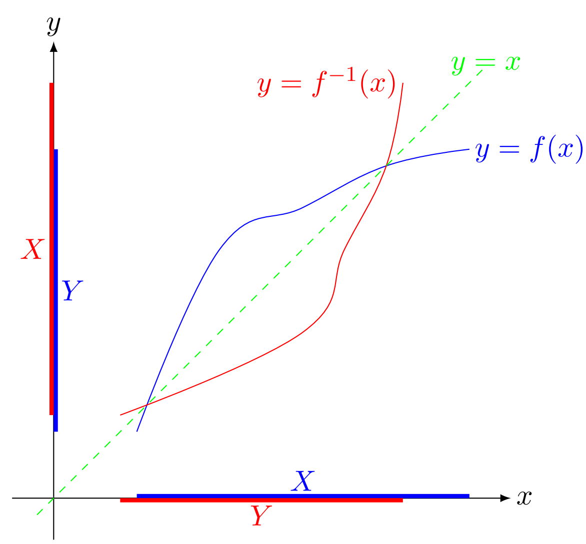
MERKNAD 2.3.6. For en vilkårlig funksjon $f : D(f) \to Y$ kan vi alltid erstatte Y med $V(f)$ og definere en surjektiv funksjon $f : D(f) \to V(f)$ med samme regel. (Strengt tatt burde vi kanskje gi et nytt navn til funksjonen.) Dersom f er injektiv, kan vi altså automatisk definere en bijektiv funksjon $f : D(f) \to V(f)$. Når vi da sier at $f : D(f) \to Y$ *har en invers funksjon* eller *er invertibel*, mener vi at funksjonen $f : D(f) \to Y$ er injektiv, slik at $f : D(f) \to V(f)$ er bijektiv og har derfor en invers $f^{-1} : V(f) \to D(f)$, selv om dette strengt tatt er en litt uforsiktig språkbruk.

For å få en bedre forståelse for inverse funksjoner kan det være lurt å tenke grafisk, ihvertfall i tilfeller der både definisjons- og verdimengde ligger i $\mathbb{R}$. La oss betrakte en bijektiv funksjon $f : X \to Y$ med $X, Y \subset \mathbb{R}$ og minne oss om hvordan vi tenker oss en slik funksjon ved hjelp av grafen. La oss markere definisjonsmengden X langs en horisontal akse og verdimengden Y langs en vertikal akse, som vist på figuren nedenfor. For hver "input"-verdi $x \in X$ finner vi "output"-verdien $f(x) \in Y$ ved å finne skjæringspunktet $(x, f(x))$ mellom den vertikale linjen gjennom x og grafen til f, og deretter finne skjæringspuniktet mellom den horisontale linjen gjennom $(x, f(x))$ og den vertikale aksen. Vi starter altså i X og finner den tilhørende funksjonsverdien i Y via grafen $y = f(x)$. Den inverse funksjonen $f^{-1} : Y \to X$ går motsatt vei: vi starter med en "input"-verdi $y \in Y$ og utfører den motsatte prosedyren for å finne den tilhørende "output"-verdien $f^{-1}(y) \in X$, som vist i følgende figur:



De inverse funksjonene $x \mapsto f(x)$ og $y \mapsto f^{-1}(y)$ visualisert grafisk

Grafen til f er som nevnt kurven i planet gitt ved ligningen $y = f(x)$, som er ekvivalent med $x = f^{-1}(y)$ ved (35). Vi kan likegodt betrakte dette som grafen til den inverse funksjonen $f^{-1} : Y \to X$, slik at forrige figur visualiserer like godt grafen til f som grafen til f^{-1}. Men for f^{-1} er dette ikke den måten vi er vant til å tegne en graf på, siden vi har tegnet inn definisjonsmengden Y langs den vertikale aksen og verdimengden X langs den horisontale aksen. For å tegne grafen til f^{-1} som vi er vant til, må vi derfor bytte om rollene på de to aksene, som vil tilsvare å speile hele figuren om linjen $y = x$. Gjør vi det, vil kurven $x = f^{-1}(y)$ få ligningen $y = f^{-1}(x)$ og vi får følgende figur som viser grafene til begge funksjonene f og f^{-1} tegnet "som vi er vant til":



Grafene til $f : X \to Y$ og $f^{-1} : Y \to X$ visualisert i samme koordinatsystem

Vi merker spesielt at vi på grunn av speilingen om $y = x$ får:

Observasjon 2.3.7: Symmetri av grafer til inverse funksjoner

Grafene $y = f(x)$ og $y = f^{-1}(x)$ til to inverse funksjoner f og f^{-1} er speilsymmetriske om linjen $y = x$.

Eksempler. Vi starter med et velkjent eksempel:

EKSEMPEL 2.3.8. La oss studere funksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved $f(x) = x^2$. For $a > 0$ finnes nøyaktig to verdier for x slik at $f(x) = a$, nemlig $x = \pm\sqrt{a}$; med andre ord har vi $f^{-1}(\{a\}) = \{-\sqrt{a}, \sqrt{a}\}$. For $a = 0$ finnes bare én slik x-verdi, nemlig $x = 0$. For $a < 0$ finnes ingen x-verdier slik at $f(x) = a$, fordi x^2 aldri kan være negativ når $x \in \mathbb{R}$. Som oppsummering viser dette at

- f ikke er injektiv, siden $f(-\sqrt{a}) = f(\sqrt{a}) = a$ for alle $a > 0$;
- f ikke er surjektiv, siden $V(f) = [0, \infty)$.

Vi kan imidlertid betrakte restriksjonen av f til $[0, \infty)$ definert ved

$$\begin{aligned} f|_{[0,\infty)} : [0, \infty) &\longrightarrow [0, \infty) \\ x &\mapsto x^2 \end{aligned}$$

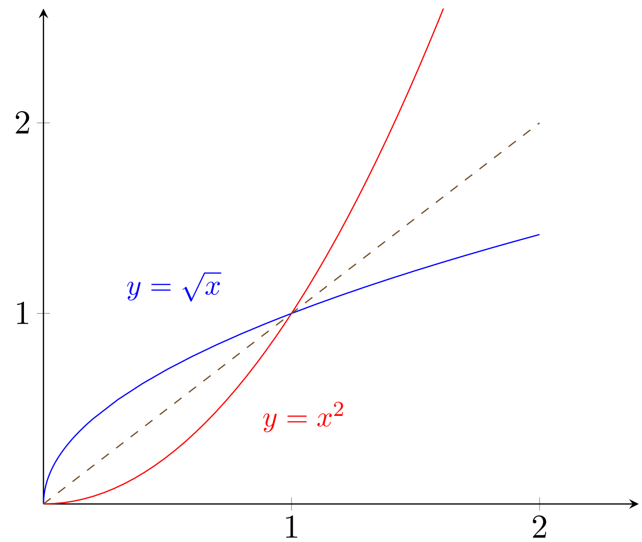
(jf. Definisjon 2.2.10). Vi betegner denne for enkelhets skyld med g. Denne har samme verdimengde som f og er dermed surjektiv. Dessuten er g injektiv, siden det nå kun er $\sqrt{a}$ i definisjonsmengden $[0, \infty)$ til g som sendes på $a \in [0, \infty)$, det vil si $g^{-1}(\{a\}) = \{\sqrt{a}\}$ for alle $a \in [0, \infty)$. Den inverse funksjonen $g^{-1} : [0, \infty) \to [0, \infty)$ er derfor gitt ved $g^{-1}(y) = \sqrt{y}$. Vi liker ofte å benytte samme bokstav for å betegne variabelen også for den inverse funksjonen og skriver derfor gjerne $g^{-1}(x) = \sqrt{x}$. Det er rett frem å sjekke at

$$g^{-1} \circ g(x) = g^{-1}(g(x)) = \sqrt{x^2} = x, \quad \text{for alle } \; x \in D(g) = [0, \infty)$$

og

$$g \circ g^{-1}(x) = g(g^{-1}(x)) = \left(\sqrt{x}\right)^2 = x, \quad \text{for alle } \; x \in D(g^{-1}) = [0, \infty).$$

Grafene til g og g^{-1} tegnet inn i samme koordinatsystem er, som bemerket i Observasjon 2.3.7, symmetriske om linjen $y = x$:



Grafene til g og g^{-1} fra Eksempel 2.3.8 tegnet inn i samme koordinatsystem

Eksempel 2.3.9. Betrakt funksjonen i Eksempel 2.2.14(iv). Ved å erstatte Y med $V(f) = \{p, q, r\}$, får vi en bijektiv funksjon

$$\begin{aligned} f : \{a, b, c\} \quad &\longrightarrow \quad \{p, q, r\} \\ a \quad &\mapsto \quad p \\ b \quad &\mapsto \quad q \\ c \quad &\mapsto \quad r \end{aligned}$$

(jf. Merknad 2.3.6). Den inverse funksjonen er gitt ved

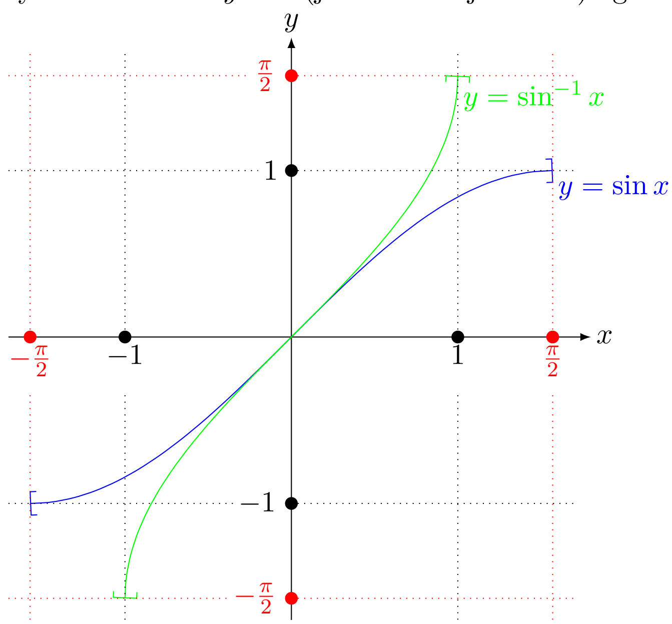
$$\begin{aligned} f^{-1} : \{p, q, r\} \quad &\longrightarrow \quad \{a, b, c\} \\ p \quad &\mapsto \quad a \\ q \quad &\mapsto \quad b \\ r \quad &\mapsto \quad c. \end{aligned}$$

Eksempel 2.3.10. Betrakt funksjonen $g : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1]$ definert ved $g(x) = \sin x$, som er restriksjonen (jf. Definisjon 2.2.10) av sinusfunksjonen (jf. Eksempel 2.2.8) til intervallet $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Tenker vi et øyeblikk på den geometriske definisjonen av sinusfunksjonen, vil vi innse at g er injektiv og surjektiv, altså bijektiv. Den inverse funksjonen $g^{-1} : [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$ assosierer til et tall $x \in [-1, 1]$ det entydige tallet $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ slik at $g(\theta) = \sin\theta = x$. (Her er θ den greske bokstaven "teta".) Vi betegner gjerne θ med $\sin^{-1} x$ eller $\arcsin x$, slik at den inverse funksjonen kan beskrives ved regelen

$$\begin{aligned} g^{-1} : [-1,1] &\longrightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \\ x &\mapsto \sin^{-1} x = \arcsin x. \end{aligned}$$

Denne funksjonen kalles *den inverse sinusfunksjonen* eller også *arcus-sinus-funksjonen.*

Grafene til sinusfunksjonen (restrisert til $[-\frac{\pi}{2}, \frac{\pi}{2}]$) og arcussinusfunksjonen er speilsymmetriske om $y = x$ (jf. Observasjon 2.3.7) og ser slik ut:



Grafene til sin og $\sin^{-1}$ tegnet inn i samme koordinatsystem

Samme resonnement som ovenfor kan brukes på den bijektive funksjonen

$$\begin{aligned} h : [0,\pi] &\longrightarrow [-1,1] \\ x &\mapsto \cos x, \end{aligned}$$

som har en inversfunksjon $h^{-1} : [-1,1] \to [0,\pi]$ som assosierer til et tall $x \in [-1,1]$ det entydige tallet $\theta \in [0,\pi]$ slik at $h(\theta) = \cos\theta = x$. Vi betegner gjerne θ med $\cos^{-1} x$ eller $\arccos x$, slik at den inverse funksjonen kan beskrives ved regelen

$$\begin{aligned} h^{-1} : [-1,1] &\longrightarrow [0,1] \\ x &\mapsto \cos^{-1} x = \arccos x. \end{aligned}$$

Denne funksjonen kalles *den inverse cosinusfunksjonen* eller også *arcus-cosinus-funksjonen.*

Vi henviser til Merknad 9.6.12 for den inverse tangensfunksjonen og til [**AE**, §3.5] for inverse trigonometriske funksjoner generelt.

EKSEMPEL 2.3.11. Inverse funksjoner dukker opp i mange sammenhenger i realfag og i dagliglivet.

Hvis vi for eksempel er i et land der temperaturen angis i Fahrenheit, trenger vi en funksjon som omdanner temperatur i Fahrenheit til Celsius for å få vite hvor varmt eller kaldt det er i henhold til det vi er vant til. Lar vi F angi temperaturen i Fahrenheit, og C temperaturen i Celsius, så har vi at $F \in [-459.67, \infty)$ og $C \in [-273.15, \infty)$ og at funksjonen

$$\begin{aligned} [-459.67, \infty) &\longrightarrow [-273.15, \infty) \\ F &\mapsto \frac{5}{9}(F-32). \end{aligned} \tag{36}$$

gir oss temperaturen i Celsius.

Er vi imidlertid interessert i å omdanne temperatur i Celsius til temperatur i Fahrenheit, da er vi interessert i den inverse funksjonen

$$\begin{aligned} [-273.15, \infty) &\longrightarrow [-459.67, \infty) \\ C &\mapsto F \text{ slik at } C = \frac{5}{9}(F-32). \end{aligned}$$

I dette tilfellet kan vi eksplisitt finne uttrykket for den inverse funksjonen, ved å løse ligningen $C = \frac{5}{9}(F-32)$ med hensyn på F. Enkel regning viser at

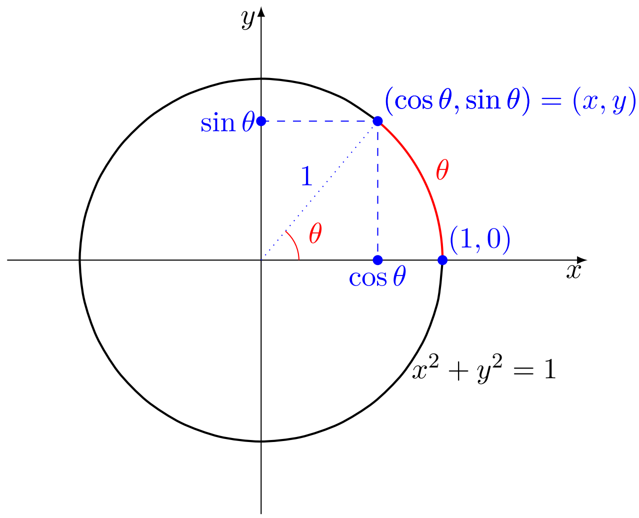
$$F = \frac{9}{5}C + 32.$$

Altså er den inverse funksjonen til (36) gitt ved

$$\begin{aligned} [-273.15, \infty) &\longrightarrow [-459.67, \infty) \\ C &\mapsto \frac{9}{5}C + 32. \end{aligned}$$

MERKNAD 2.3.12. Det er ytterst sjeldent at vi klarer å finne et eksplisitt uttrykk for den inverse funksjonen som i forrige eksempel. I alle tilfeller, gitt en bijektiv funksjon $x \mapsto y(x)$, så er prosedyren for å forsøke å finne uttrykket for den inverse funksjonen å løse ligningen $y = y(x)$ eksplisitt med hensyn på x for å ende opp med et uttrykk på formen $x = x(y)$. Da er $y \mapsto x(y)$ uttrykket for den inverse funksjonen.

EKSEMPEL 2.3.13. La $\mathbb{S}^1$ være *enhetssirkelen*, med andre ord delmengden av $\mathbb{R}^2$ bestående av alle punkter i avstand 1 fra origo, eller ekvivalent av alle punkter (x, y) som oppfyller ligningen $x^2 + y^2 = 1$ (jf. Eksempel 2.1.16). Ethvert punkt (x, y) på enhetssirkelen kan også beskrives ved $(\cos\theta, \sin\theta)$, hvor $\theta \in [0, 2\pi)$ er vinkelen mellom linjestykkene fra (x, y) til origo og fra origo til punktet $(1,0)$, som er per definisjon lengden av sirkelbuen fra $(1,0)$ til (x,y), som vist i følgende figur:

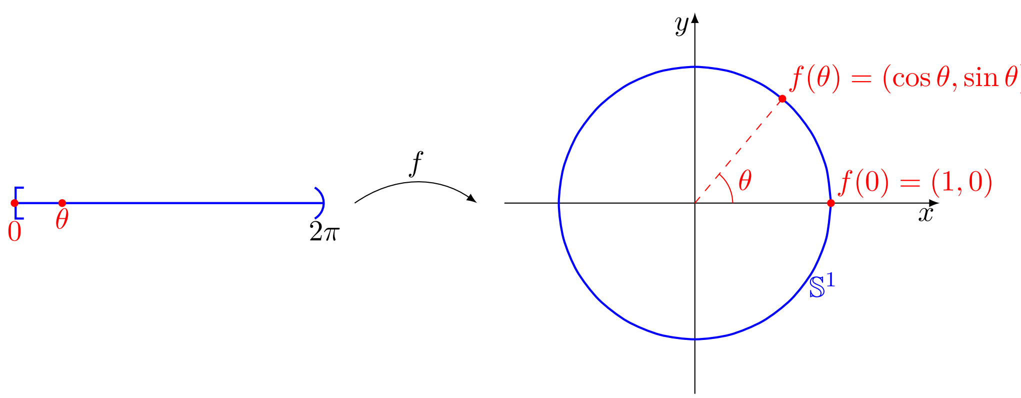


Punktet (x,y) på enhetssirkelen tilsvarer en entydig $\theta \in [0,2\pi)$

Dette betyr at det finnes en surjektiv funksjon

$$f : [0,2\pi) \to \mathbb{S}^1, \quad f(\theta) = (\cos\theta, \sin\theta),$$

visualisert i følgende figur:

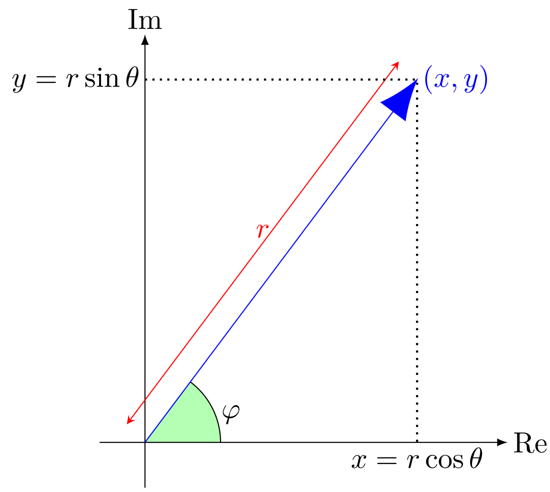


Funksjonen er dessuten injektiv, siden forskjellige θer definerer forskjellige punkter på enhetssirkelen. Den inverse funksjonen f^{-1} sender et punkt P på $\mathbb{S}^1$ til buelengden θ mellom $(1,0)$ og P.

EKSEMPEL 2.3.14. Et punkt (x,y) i $\mathbb{R}^2$ utenom origo kan beskrives ved hjelp av størrelsene

- $r = \sqrt{x^2+y^2}$ lik avstanden mellom punktet og origo, og
- φ (den greske bokstaven "fi") lik vinkelen bestemt av linjestykket mellom (x,y) og origo og den horisontale aksen.

Paret (r, φ) kalles *polarkoordinatene* til (x, y).



Polare koordinater r og φ til punktet med kartesiske koordinater (x, y).

Forskjellige punkter (x, y) har forskjellige polare koordinater (r, θ). Vi har $r \in (0, \infty)$ og $\theta \in [0, 2\pi)$, og ethvert par (r, θ) som oppfyller dette definerer et punkt i $\mathbb{R}^2$ med kartesiske koordinater $(x, y) = (r\cos\theta, r\sin\theta)$, som vist på figuren. Dette betyr at en vi har en bijektiv funksjon

$$\begin{array}{rcl} \mathbb{R}^2 \setminus \{(0,0)\} & \longrightarrow & (0,\infty) \times [0, 2\pi) \\ (x, y) & \mapsto & \left(\sqrt{x^2 + y^2}, \theta\right) \end{array}$$

med invers

$$\begin{array}{rcl} (0,\infty) \times [0, 2\pi) & \longrightarrow & \mathbb{R}^2 \setminus \{(0,0)\} \\ (r, \theta) & \mapsto & (r\cos\theta, r\sin\theta). \end{array}$$

Vi avslutter seksjonen med definisjonen av en endelig mengde, som er intuitivt ganske opplagt:

Definisjon 2.3.15: Endelig mengde

La $n \in \mathbb{N}$. En mengde A sies å være *endelig*[a], og å inneholde n elementer, dersom enten

- $n = 0$ og $A = \emptyset$, eller
- $n > 0$ og det finnes en bijeksjon $\{1, \ldots, n\} \to A$.

I disse tilfellene sier vi også at A har *kardinalitet*[b] n.

[a] *Finite* på engelsk.
[b] *Cardinality* på engelsk.

2.4. Relasjoner

Stoffet i denne seksjonen er nytt i forhold til skolen og kommer til å spille viktige roller senere i teksten.

Relasjoner. Delmengder av kartesiske produkter har fått egne navn:

Definisjon 2.4.1: Relasjon

La X og Y være mengder. En *relasjon*[a] R på X og Y er en delmengde R av $X \times Y$, det vil si $R \subset X \times Y$. En *relasjon* R på X er en delmengde av $X \times X$, det vil si $R \subset X \times X$.

At $(x, y) \in R$ betegnes ofte med xRy, og vi sier at x *er* R*-relatert til* y.

[a] *Relation* på engelsk.

Eksempel 2.4.2. La S være mengden av alle mennesker på jorden, i nå- og fortid. Da er

$$R = \{(x, y) \in S \times S \mid x \text{ er en etterkommer av } y\}$$

en relasjon på S. Her er altså paret (x, y) av mennesker R-relatert hvis og bare hvis x er etterkommer av y.

Eksempel 2.4.3. La $X = \{1, 2, 3, 4\}$ og

$$R = \{(1, 1), (1, 2), (1, 3), (1, 4), (2, 2), (2, 4), (3, 3), (4, 4)\}.$$

Da er R en relasjon på X. Ved nærmere ettersyn ser vi at xRy (eller x er R-relatert til y) hvis og bare hvis x deler y, for $x, y \in X$.

Eksempel 2.4.4. Grafen til en funksjon $f : X \to Y$ er en relasjon på X og Y, siden den er en delmengde av $X \times Y$ (jf. Definisjon 2.2.15).

Vi kan nå definere funksjonsbegrepet noe mer formelt som følger (jf. Oppgave 2.30).

Definisjon 2.4.5: Funksjon

En *funksjon* fra en mengde X til en mengde Y er en relasjon R på X og Y som oppfyller

(i) til enhver $x \in X$ finnes en $y \in Y$ slik at xRy;
(ii) hvis xRy og xRz, da er $y = z$.

Ekvivalensrelasjoner. Følgende begrep vil spille en viktig rolle i konstruksjonen av heltallene, de rasjonale tallene og de reelle tallene.

Definisjon 2.4.6: Ekvivalensrelasjon

En *ekvivalensrelasjon*[a] på en mengde X er en relasjon R som oppfyller følgende betingelser:

(i) *Refleksivitet*[b]: xRx for alle $x \in X$.
(ii) *Symmetri*[c]: hvis xRy, da er yRx, for alle $x, y \in X$.
(iii) *Transitivitet*[d]: hvis xRy og yRz, da er xRz, for alle $x, y, z \in X$.

[a] *Equivalence relation* på engelsk.
[b] *Reflexivity* på engelsk.
[c] *Symmetry* på engelsk.
[d] *Transitivity* på engelsk.

Ofte betegner vi en ekvivalensrelasjon R med $\sim$, slik at betingelsene ovenfor kan uttrykkes som:

(i) Refleksivitet: $x \sim x$ for alle $x \in X$.
(ii) Symmetri: hvis $x \sim y$, da er $y \sim x$, for alle $x, y \in X$.
(iii) Transitivitet: hvis $x \sim y$ og $y \sim z$, da er $x \sim z$, for alle $x, y, z \in X$.

Vi leser "$\sim$" som "er ekvivalent med".

Eksempel 2.4.7. Som et veldig opplagt eksempel på en ekvivalensrelasjon på en hvilken som helst mengde X har vi *likhetsrelasjonen* $x = y$.

Eksempel 2.4.8. Relasjonen i Eksempel 2.4.2 er ikke en ekvivalensrelasjon. Transitivitetsbetingelsen holder, men verken refleksivitet eller symmetri holder. Til nød kan man kverulere om hvorvidt man kan si at en person er etterkommer av seg selv, men det er ihvertfall slik at hvis x er en etterkommer av y, og $x \neq y$ så er *ikke* y en etterkommer av x. (Ålreit, med mindre man er hovedpersonen i science fiction novellen *'–All You Zombies–'* av den amerikanske forfatteren Robert A. Heinlein (1907–1988) fra 1958, filmatisert som *Predestination* i 2014.)

Eksempel 2.4.9. Relasjonen R i Eksempel 2.4.3 er ikke en ekvivalensrelasjon. Transitivitet og refleksivitet holder, men symmetri holder ikke, da for eksempel $(1, 2) \in R$, mens $(2, 1) \notin R$.

Eksempel 2.4.10. La $\mathcal{M}$ være en mengde som består av mengder. Relasjonen på $\mathcal{M}$ definert ved at

$$A \sim B \iff \text{det finnes en bijektiv funksjon } f : A \to B$$

(som betyr at A og B er i én-til-én korrespondanse eller har samme kardinalitet, i språkbruken til Definisjon 2.3.4) er en ekvivalensrelasjon: Refleksivitet holder siden *identitetsavbildningen* $id : A \to A$ som avbilder x til x er bijektiv. Symmetri holder siden en bijektiv $f : A \to B$ har en invers funksjon $f^{-1} : B \to A$ som også er bijektiv. Transitivitet holder fordi dersom $g : A \to B$ og $f : B \to C$ er bijektive, da er den sammensatte funksjonen $f \circ g : A \to C$ (jf. Definisjon 2.2.12) bijektiv ved Oppgave 2.22.

Et viktig poeng med ekvivalensrelasjoner er at vi kan bruke dem til å "dele opp" eller "partere" mengden vår i delmengder, bestående av alle elementene som er relatert til hverandre, som vi nå forklarer.

Definisjon 2.4.11: Ekvivalensklasse

Gitt en ekvivalensrelasjon $\sim$ på en mengde X og $x \in X$, så definerer vi *ekvivalensklassen*[a] $[x]$ *til* x som mengden:

$$[x] = \{y \in X \mid y \sim x\}.$$

Et element i ekvivalensklassen kalles en *representant*[b] for ekvivalensklassen.

Vi betegner ofte *mengden av ekvivalensklasser* med $X/\sim$.

[a] *Equivalence class* på engelsk.
[b] *Representative* på engelsk.

Følgende er en konsekvens av definisjonen og egenskapene til ekvivalensrelasjoner:

Observasjon 2.4.12

Vi har at $x \sim y \iff [x] = [y]$.

Bevis
Vi viser først "bare hvis"-delen ("$\Rightarrow$"). Anta at $x \sim y$. Hvis $z \in [x]$ da er per definisjon $z \sim x$. Transitivitet gir at $z \sim y$, det vil si $z \in [y]$ per definisjon. Dette viset at $[x] \subset [y]$. Motsatt inklusjon fås ved å bytte roller på x og y.

Vi viser dernest "hvis"-delen ("$\Leftarrow$"). Anta $[x] = [y]$. Siden $x \sim x$ ved refleksivitet, er $x \in [x]$. Dermed er $x \in [y]$, som per definisjon betyr at $x \sim y$. $\square$

Følgende er en viktig egenskap til ekvivalensrelasjoner:

Hjelpesetning 2.4.13

Enhver $x \in X$ er med i nøyaktig én ekvivalensklasse. Sagt på en annen måte: to ekvivalensklasser er enten like eller disjunkte.

Bevis

La $x \in X$. Ved refleksivitet har vi $x \sim x$, slik at $x \in [x]$, hvilket betyr at x ligger i minst én ekvivalensklasse. Hvis $x \in [y_1]$ og $x \in [y_2]$, da vil $x \sim y_1$ og $x \sim y_2$. Ved symmetri har vi $y_1 \sim x$. Dette, sammen med at $x \sim y_2$ og transitivitetsegenskapen, gir at $y_1 \sim y_2$, det vil si $[y_1] = [y_2]$ ved Observasjon 2.4.12. Dette betyr at x ligger i kun én ekvivalensklasse. $\square$

Dette sier oss at $X/\sim$ består av *disjunkte delmengder* av X, og unionen av alle disse er hele X. Vi kaller en slik union en *partisjon*[35]. Merk også at vi har en naturlig surjektiv avbildning

$$\begin{array}{ccc} X & \longrightarrow & X/\sim \\ x & \mapsto & [x]. \end{array}$$

Eksempel 2.4.14. La $m \in \mathbb{Z}^+ = \{1, 2, 3, \ldots\}$ og $X = \mathbb{Z}$.

La oss vise at relasjonen på $\mathbb{Z}$ definert ved

$$x \sim y \Longleftrightarrow m \text{ deler } |x - y|$$

er en ekvivalensrelasjon.

Vi har at m deler $0 = x - x$, som betyr at $x \sim x$. Altså er relasjonen refleksiv.

Vi har at $|x - y|{=}|y - x|$, som betyr at m deler $|x - y|$ hvis og bare hvis m deler $|y - x|$. Altså har vi at $x \sim y \Leftrightarrow y \sim x$, hvilket betyr at relasjonen er symmetrisk.

Hvis $x \sim y$ og $y \sim z$, da har vi per definisjon at m deler $|x - y|$ og $|y - z|$, hvilket betyr at $x - y = km$ og $y - z = lm$ for $k, l \in \mathbb{Z}$. Dette gir at $x - z = (x - y) + (y - z) = km + lm = (k + l)m$, hvilket betyr at m deler $|x - z|$, altså at $x \sim z$. Dette viser at relasjonen er transitiv.

Vi har dermed vist at $\sim$ er en ekvivalensrelasjon.

Ekvivalensklassen til en vilkårlig $n \in \mathbb{Z}$ er

$$[n] = \{k \in \mathbb{Z} \mid m \text{ deler } |n - k|\} = \{n, n \pm m, n \pm 2m, n \pm 3m, \ldots\}.$$

Det er klart at $[n]$ inneholder et entydig element n_0 som oppfyller betingelsen $0 \leq n_0 \leq m - 1$. Dermed er mengden av ekvivalensklasser $\mathbb{Z}/\sim$, som vi ofte betegner med $\mathbb{Z}/m\mathbb{Z}$ eller $\mathbb{Z}_m$, i én-til-én korrespondanse med mengden

$$\{0, 1, 2, \ldots, m - 1\}.$$

[35] *Partition* på engelsk.

MERKNAD 2.4.15. Ekvivalensrelasjonen $\sim$ i forrige eksempel betegnes vanligvis som $x \equiv_m y$ eller som $x \equiv y \pmod{m}$ og leses som "x er kongruent med y modulo m" eller (på grunn av symmetri) "x og y er kongruente modulo m". Hvis x og y ikke er kongruente modulo m, skriver vi $x \not\equiv y \pmod{m}$. Ekvivalensklassen til $n \in \mathbb{Z}$ betegnes vanligvis med $[n]_m$. Altså kan vi skrive

$$\mathbb{Z}_m = \{[0]_m, [1]_m, [2]_m, \ldots, [m-1]_m\}.$$

EKSEMPEL 2.4.16. Relasjonen på planet $\mathbb{R}^2$ gitt ved

$$z_1 \sim z_2 \iff z_1 \text{ og } z_2 \text{ er i samme avstand fra origo}$$

er en ekvivalensrelasjon. Vi overlater til Oppgave 2.31 å vise dette.

La $z \in \mathbb{R}^2$. Da er ekvivalensklassen til z lik

$$[z] = \{c \in \mathbb{R}^2 \mid c \text{ er i samme avstand fra origo som } z\},$$

som består av alle punkter i planet som ligger på sirkelen om origo som innholder punktet z. Dermed er mengden av ekvivalensklasser i én-til-én korrespondanse med mengden av sirkler om origo i planet (inkludert sirkelen av radius 0, som er kun origo selv). Denne mengden er igjen i én-til-én korrespondanse med mengden $\mathbb{R}^{\geq 0}$ av alle ikke-negative reelle tall, siden enhver sirkel om origo er bestemt av sin radius.

EKSEMPEL 2.4.17. La $\mathbf{0} = (0, \ldots, 0)$ være origo i $\mathbb{R}^{n+1}$, for et positivt heltall n. Relasjonen på $\mathbb{R}^{n+1} \setminus \mathbf{0}$ gitt ved

$$\mathbf{x} \sim \mathbf{y} \iff x \text{ og } y \text{ ligger på samme linje gjennom origo i } \mathbb{R}^{n+1}$$

er en ekvivalensrelasjon. Vi overlater til Oppgave 2.36 å vise dette, samt å vise at $\mathbf{x} = (x_1, \ldots, x_{n+1})$ og $\mathbf{y} = (y_1, \ldots, y_{n+1})$ ligger på samme linje gjennom origo i $\mathbb{R}^{n+1}$ hvis og bare hvis det finnes en $k \in \mathbb{R} \setminus \{0\}$ slik at $(x_1, \ldots, x_{n+1}) = k(y_1, \ldots, y_{n+1})$. Dette betyr at det kun er *forholdet* mellom koordinatene som bestemmer ekvivalensklassen, og vi betegner derfor ekvivalensklassen til $\mathbf{x} = (x_1, \ldots, x_{n+1})$ med $(x_1 : \cdots : x_{n+1})$.

En ekvivalensklasse består altså av alle punkter på den samme linjen gjennom origo (bortsett fra origo selv). Mengden $(\mathbb{R}^{n+1} \setminus \mathbf{0}) \,/ \sim$ av alle ekvivalensklasser består dermed av alle linjer gjennom origo i $\mathbb{R}^{n+1}$. Vi betegner $(\mathbb{R}^{n+1} \setminus \mathbf{0}) \,/ \sim$ med $\mathbb{P}^n_{\mathbb{R}}$ og kaller det *det projektive n-rommet (over $\mathbb{R}$)*[36].

[36] *Projective n-space* på engelsk.

Oppgaver

Oppgaver til §2.1

OPPGAVE 2.1. List opp alle elementene i mengden $\{x \in \mathbb{R} \mid x^2 = 1\}$.

OPPGAVE 2.2. Betrakt mengdene $A = \{2, 4, 6\}$, $B = \{2, 6\}$, $C = \{4, 6\}$, $D = \{4, 6, 8\}$. Avgjør hvilke som er delmengder av hverandre.

OPPGAVE 2.3. Gi en kortere beskrivelse av mengdene $[0, 2] \cup [1, 3]$ og $[0, 2] \cap [1, 3]$.

OPPGAVE 2.4. Vis (så mange som du orker av) mengdeidentitetene i tabellen på side 41, bortsett fra den som er vist i Setning 2.1.14.

OPPGAVE 2.5. Bevis eller motbevis (ved å gi et moteksempel) følgende utsagn om mengder:

(i) $A \setminus (B \setminus C) = (A \setminus B) \setminus C$.
(ii) $A \setminus (B \setminus C) = (A \setminus B) \cup C$.
(iii) Hvis $A \cap B = A \cup B$, da er $A = B$.

OPPGAVE 2.6. La A og B være delmengder av en mengde X. Vis at $A \subset B$ hvis og bare hvis $A \cap (X \setminus B) = \emptyset$.

OPPGAVE 2.7. La S være en mengde. *Potensmengden*[37] *av* S, noen ganger betegnet som $P(S)$ eller 2^S, er *mengden av alle delmengder av* S. Finn potensmengden til $\{0, 1, 2\}$.

OPPGAVE 2.8. La $A = \{a, b, c, d\}$ og $B = \{x, y\}$. Finn $A \times B$ og $B \times A$.

OPPGAVE 2.9. Anta A og B er mengder. Kan det skje at $A \times B = B \times A$?

OPPGAVE 2.10. Anta A og B er mengder slik at $A \times B = \emptyset$. Hva kan du da konkludere?

OPPGAVE 2.11. Denne oppgaven presenterer *Russels paradoks*, oppdaget i 1901 av den britiske filosofen, logikeren, matematikeren og forfatteren Bertrand Russell (1872–1970).

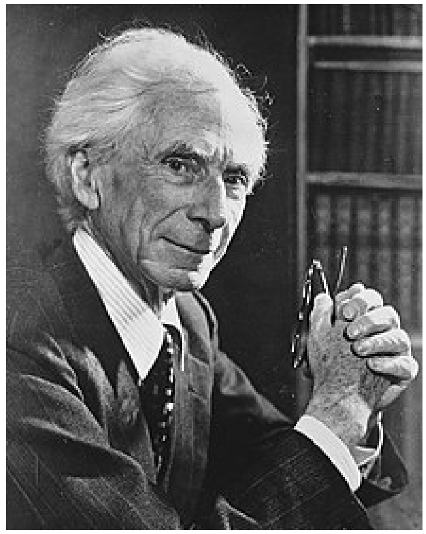
La S være mengden av alle mengder som ikke er element i seg selv, det vil si $S = \{x \mid x \notin x\}$.

Vis at begge antagelsene $S \in S$ og $S \notin S$ leder til en motsigelse.

Av dette konkluderer vi at når vi definerer en mengde ved hjelp av karakteristiske egenskaper ved objektene den inneholder, så kan dette lede til selvmotsigelser (jf. Merknad 2.1.3).

(Synes du denne oppgaven er vanskelig å forstå, kan du først se på neste oppgave, som gir en velkjent populærvitenskapelig variant av paradokset.)

[37] *Power set* på engelsk.



Bertrand Arthur William Russell (Kilde: Wikipedia, public domain)

Oppgave 2.12. En forenklet variant av *Russels paradoks* fra forrige oppgave, som Russell selv brukte for å illustrere sitt paradoks på en populærvitenskapelig måte, er *paradokset om barbereren*[38]:

I en landsby langt borte er barbereren definert som *den som barberer alle menn i landsbyen som ikke barberer seg selv og kun menn som ikke barberer seg selv.* Barberer han da også seg selv?

Oppgaver til §2.2–2.3

Oppgave 2.13. La $f : \mathbb{R} \to \mathbb{R}$ være gitt ved $f(x) = \cos x$. Finn $f^{-1}(\{0\})$ og $f^{-1}\left([0, \frac{\pi}{3})\right)$.

Oppgave 2.14. Betrakt funksjonen $f : X \to Y$ i Eksempel 2.2.14(i). Hva er $f(\{a, b, c\})$ og $f^{-1}(\{p, q\})$?

Oppgave 2.15. La $f : X \to Y$ være en funksjon og $B \subset Y$. Vis at $f^{-1}(B) = f^{-1}(B \cap V(f))$, der $V(f)$ som vanlig er verdimengden til f.

Oppgave 2.16. La $f : X \to Y$ og la $A_1, A_2 \subset X$ og $B_1, B_2 \subset Y$. Vis at

(a) $A_1 \subset A_2 \Longrightarrow f(A_1) \subset f(A_2)$;
(b) $B_1 \subset B_2 \Longrightarrow f^{-1}(B_1) \subset f^{-1}(B_2)$;
(c) $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$;
(d) $f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2)$;
(e) $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$;
(f) $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$.

Oppgave 2.17. Anta at $f : X \to Y$ med X og Y delmengder av $\mathbb{R}$. Betrakt grafen $\Gamma(f) \subset X \times Y \subset \mathbb{R}^2$ og tenk den som "tegnet inn" i planet

[38] *The barber paradox* på engelsk.

som vi er vant til fra skolen. Hvordan kan du ut fra grafen avgjøre om f er injektiv?

Oppgave 2.18. Betrakt funksjonene $f, g : \mathbb{Z} \to \mathbb{Z}$ definert ved $f(n) = 2n + 1$ og $g(n) = 2\lfloor \frac{n}{2} \rfloor$. Her betegner $\lfloor x \rfloor$ "nedrundingen" av tallet x, det vil si det største hele tallet m slik at $m \leq x$.

(a) Hva er verdimengden til hver av funksjonene f, g, $f \circ g$ og $g \circ f$, henholdsvis? Hvilke av funksjonene f, g, $f \circ g$ og $g \circ f$ er surjektive?
(b) Hvilke av funksjonene f, g, $f \circ g$ og $g \circ f$ er injektive?

Oppgave 2.19. Betrakt funksjonen $f : \mathbb{R} \to \mathbb{R}$ definert ved $f(x) = x^2$. Finn følgende:

(a) $f([-1, 1))$,
(b) $f^{-1}(f([-1, 1)))$,
(c) $f^{-1}([-1, 1])$,
(d) $f(f^{-1}([-1, 1]))$,
(e) $f^{-1}((0, 1))$.

Oppgave 2.20. La $f : X \to Y$ være en funksjon.

(a) Vis at $f\left(f^{-1}(B)\right) \subset B$ for alle $B \subset Y$.
(b) Vis ved et eksempel at likhet ikke trenger å holde i (a).
(c) Vis at likhet holder i (a) hvis f er surjektiv.
(d) Vis at $f^{-1}(f(A)) \subset A$ for alle $A \subset X$.
(e) Vis ved et eksempel at likhet ikke trenger å holde i (d).
(f) Vis at likhet holder i (e) hvis f er injektiv.

Oppgave 2.21. La $f : A \to B$ og $g : C \to D$ være funksjoner mellom ikke-tomme mengder. Vi definerer funksjonen

$$\begin{aligned} f \times g : A \times C &\to B \times D \\ (a, c) &\mapsto (f(a), g(c)) \end{aligned}$$

Vis at $f \times g$ er injektiv (henholdsvis surjektiv) hvis og bare hvis f og g begge er injektive (henholdsvis surjektive).

Oppgave 2.22. La $g : X \to Y$ og $f : Y \to Z$ være bijektive funksjoner. Vis at sammensetningen $f \circ g : X \to Z$ er bijektiv og at den inverse funksjonen er gitt ved $(f \circ g) = g^{-1} \circ f^{-1}$.

Oppgave 2.23. Vis at hvis A har kardinalitet m og kardinalitet n (jf. Definisjon 2.3.15), da er $m = n$.

Oppgave 2.24. Ola og Kari diskuterer begrepene *bilde* og *invers bilde* til en funksjon $f : X \to Y$. Ola er forvirret: "Dersom f har en invers funksjon, og vi betegner denne som vanlig med f^{-1}, og A er en delmengde av Y, da vil jo $f^{-1}(A)$ både kunne betegne bildet av A ved f^{-1} *og* det inverse bildet av A ved f. Notasjonen er forvirrende og tvetydig", sier han.

Kari på sin side tar det hele med ro og sier "Slapp av, det er egentlig ikke så farlig". Hvorfor synes Kari det?

Oppgave 2.25. La A og B være endelige mengder (jf. Definisjon 2.3.15) og $f : A \to B$ en funksjon. Vis at følgende er ekvivalent:

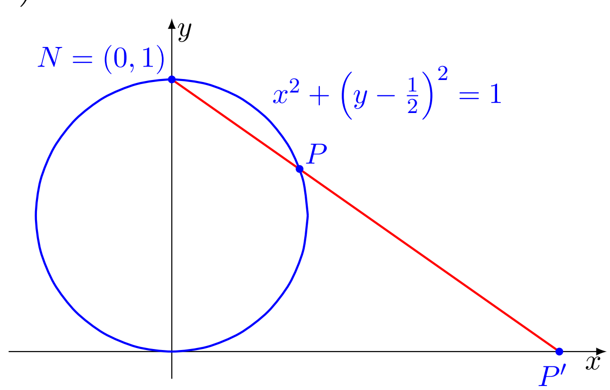
- f er injektiv,
- f er surjektiv,
- f er bijektiv.

Oppgave 2.26. La $[a, b]$ og $[b, c]$ være to lukkede, begrensede intervaller i $\mathbb{R}$. Vis at det finnes en bijeksjon mellom $[a, b]$ og $[b, c]$.

Oppgave 2.27. (a) Vis at det finnes en bijeksjon mellom intervallene $(0, 1)$ og $[0, 1)$. (Hint: Lag en tegning av $(0, 1)$ og tegn inn punktene $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots$; lag en tegning av $[0, 1)$ rett nedenfor og marker at $\frac{1}{2}$ avbildes på 0 og $(0, \frac{1}{2})$ avbildes på $(0, \frac{1}{2})$. La deretter $\frac{2}{3}$ avbildes på $\frac{1}{2}$ og $(\frac{1}{2}, \frac{2}{3})$ på $(\frac{1}{2}, \frac{2}{3})$. Fortsett prosessen.)

(b) La (a, b) og $[c, d)$ være begrensede intervaller i $\mathbb{R}$. Vis at det finnes en bijeksjon mellom dem.

Oppgave 2.28. La $\mathbb{S}^1$ være enhetssirkelen. La oss betrakte den som en delmengde av $\mathbb{R}^2$ ved å legge den inn som en sirkel av radius 1 med sentrum i $\left(0, \frac{1}{2}\right)$, slik at topp-punktet N ligger i $(0, 1)$ og bunnpunktet ligger i origo, som vist i figuren nedenunder. Ligningen som beskriver denne sirkelen er da $x^2 + \left(y - \frac{1}{2}\right)^2 = 1$.



Stereografisk projeksjon (illustrasjon til Oppgave 2.28(a)-(c))

(a) Begrunn at det finnes en bijeksjon $\mathbb{S}^1 \setminus \{N\} \to \mathbb{R}$ som assosierer til et punkt $P \in \mathbb{S}^1 \setminus \{N\}$ skjæringspunktet P' mellom x-aksen og linjen gjennom N og P, som vist i figuren. Denne avbildningen kalles *stereografisk projeksjon.*

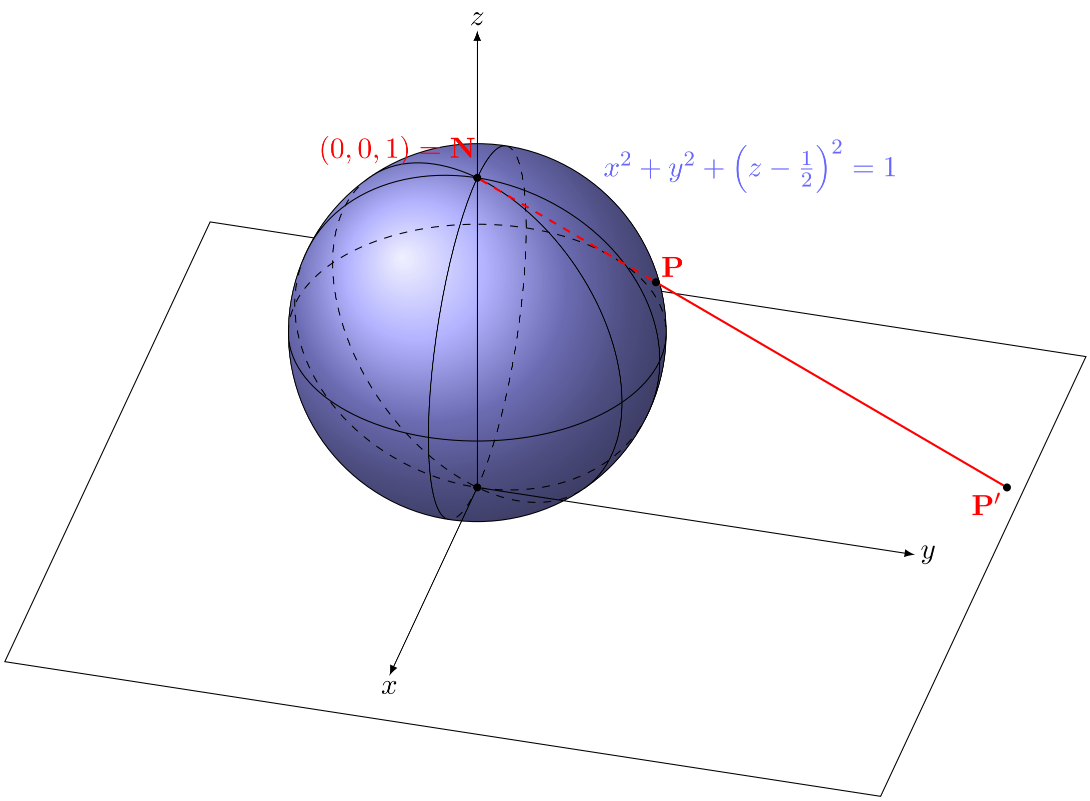
(b) Finn uttrykket for denne bijeksjonen og dens inverse funksjon.

(c) La ∞ være et symbol og la som vanlig $\{\infty\}$ være mengden som består av dette ene elementet kalt ∞. Begrunn at det finnes en bijeksjon gitt ved

$$\begin{aligned} \mathbb{S}^1 &\longrightarrow \mathbb{R} \cup \{\infty\} \\ P &\mapsto \begin{cases} P', & \text{hvis } P \neq N \\ \infty, & \text{hvis } P = N. \end{cases} \end{aligned}$$

(Når P nærmer seg N, vil linjen NP nærme seg den horisontale linjen gjennom N. Denne linjen skjærer ikke x-aksen, siden den er parallell med den, slik at avbildningen fra (a) ikke kan utvides til en avbildning $\mathbb{S}^1 \to \mathbb{R}$. Vi kan likevel tenke oss at denne linjen skjærer x-aksen i et punkt "uendelig langt borte", og dette er grunnen til at vi betegner det "ekstra punktet" vi trenger for å avbilde N i (c) for ∞. Mengden $\mathbb{R} \cup \{\infty\}$, som er bijektiv til $\mathbb{S}^1$, kalles *ettpunktskompaktifiseringen*[39] *av* $\mathbb{R}$. Intuitivt kan vi tenke på denne mengden som "tallinjen pluss et punkt i det uendelige".)

(d) Avbildningene ovenfor har en to-dimensjonal analog. La $\mathbb{S}^2$ være enhetskulen lagt inn i $\mathbb{R}^3$ slik at den har sentrum i $\left(0, 0, \frac{1}{2}\right)$. Da vil "nordpolen" N ligge i $(0, 0, 1)$, mens "sørpolen" ligger i origo, som vist i figuren:



Stereografisk projeksjon (illustrasjon til Oppgave 2.28(d))

Konstruér en bijektiv avbildning $\mathbb{S}^2 \to \mathbb{R}^2 \cup \{\infty\}$ ved å ta utgangspunkt i det du har gjort tidligere i oppgaven og figuren ovenfor.

(Mengden $\mathbb{R}^2 \cup \{\infty\}$ kalles *ettpunktskompaktifiseringen av* $\mathbb{R}^2$ og har samme "intuitive tolkning" som ettpunktkompaktifiseringen av $\mathbb{R}$. Betrakter vi $\mathbb{R}^2$ som det komplekse planet $\mathbb{C}$ (se §8.2), kalles denne mengden for *(den kompakte) Riemannsfæren*, etter den tyske matematikeren Bernhard Riemann (1826–1866).)

[39] *One-point compactification* på engelsk.

OPPGAVE 2.29. Bruk resultatet fra Oppgave 2.28(a) og Eksempel 2.3.13 til å begrunne at det finnes en bijeksjon mellom $\mathbb{R}$ og intervallet $(0, 1)$.

Oppgaver til §2.4

OPPGAVE 2.30. (a) Vis at Definisjonene 2.2.1 og 2.4.5 av en funksjon sier nøyaktig det samme.

(b) Hvordan vil du definere *verdimengden* til en funksjon f ved hjelp av språket i Definisjon 2.4.5?

(c) Hvordan vil du uttrykke egenskapene *injektiv*, *surjektiv* og *bijektiv* til en funksjon f ved hjelp av språket i Definisjon 2.4.5?

OPPGAVE 2.31. Vis at relasjonen i Eksempel 2.4.16 er en ekvivalensrelasjon.

OPPGAVE 2.32. Betrakt relasjonen R på heltallene $\mathbb{Z}$ definert ved at nRm hvis og bare hvis $nm \geq 0$. Avgjør om dette er en ekvivalensrelasjon.

OPPGAVE 2.33. En *bitstreng* er et endelig "ord" formet av "bokstavene" 0 og 1, altså en endelig rekke av 0-ere og 1-ere.

(a) Vis at relasjonen på mengden av alle bitstrenger

$$R = \{(x, y) \mid x \text{ og } y \text{ er bitstrenger som inneholder samme antall 0er}\}$$

er en ekvivalensrelasjon.

(b) Hva er ekvivalensklassene til bitstrengene 1, 00 og 101, henholdsvis, under relasjonen R?

OPPGAVE 2.34. La X være en mengde og n et positivt heltall. Definér

$$(x_1, \ldots, x_n) \sim (y_1, \ldots, y_n)$$

dersom det finnes en bijeksion $j : \{1, \ldots, n\} \to \{1, \ldots, n\}$ slik at $y_i = x_{j(i)}$ for alle $i \in \{1, \ldots, n\}$. (Med andre ord er koordinatene i de to n-tuplene permutasjoner av hverandre.) Vis at dette definerer en ekvivalensrelasjon på X^n.

Mengden av ekvivalensklasser X^n kalles det *$n$te symmetriske produktet av* X og betegnes gjerne med $\mathrm{Sym}^n(X)$. Dens elementer er *uordnede* n-tupler av elementer fra X.

OPPGAVE 2.35. Et flyselskap binder sammen fire byer, A, B, C og D, med tre fly som gjør følgende ruter:

- et fly går fra A til B og deretter fra B tilbake til A;
- et fly går fra A til C og deretter fra C tilbake til A;
- et fly går fra B til C, deretter fra C til D, og så fra D tilbake til B.

La R være relasjonen på mengden $\{A, B, C, D\}$ definert ved

$$(x, y) \in R \text{ hvis og bare hvis det finnes et direkte fly fra } x \text{ til } y.$$

(a) List opp alle elementene i R.
(b) Avgjør om R oppfyller noen av egenskapene refleksivitet, symmetri og transitivitet i Definisjon 2.4.6.

Oppgave 2.36. Vi viser til Eksempel 2.4.17.

(a) La $\mathbf{x} = (x_1, \ldots, x_{n+1})$ og $\mathbf{y} = (y_1, \ldots, y_{n+1})$ være to punkter i $\mathbb{R}^{n+1}$ forskjellig fra origo $\mathbf{0} = (0, \ldots, 0)$. Vis at $\mathbf{x}$ og $\mathbf{y}$ ligger på samme linje gjennom origo hvis og bare hvis det finnes en $k \in \mathbb{R} \setminus \{0\}$ slik at $(x_1, \ldots, x_{n+1}) = k(y_1, \ldots, y_{n+1})$.

(b) Vis at

$$\mathbf{x} \sim \mathbf{y} \Longleftrightarrow x \text{ og } y \text{ ligger på samme linje gjennom origo i } \mathbb{R}^{n+1}$$

definerer en ekvivalensrelasjon på $\mathbb{R}^{n+1} \setminus \{\mathbf{0}\}$. (Se Eksempel 2.4.17.)

Oppgave 2.37. Vi viser til Eksempel 2.4.17 og Oppgave 2.36 for definisjonen av *det projektive n-rommet* $\mathbb{P}^n_{\mathbb{R}}$. Vi betrakter i denne oppgaven tilfellene $n = 1$ og $n = 2$; i disse tilfellene kalles $\mathbb{P}^1_{\mathbb{R}}$ for *den projektive linjen* og $\mathbb{P}^2_{\mathbb{R}}$ for *det projektive plan.* Vi minner om at $\mathbb{P}^1_{\mathbb{R}}$ kan betraktes som mengden av linjer gjennom origo i $\mathbb{R}^2$, mens $\mathbb{P}^2_{\mathbb{R}}$ kan betraktes som mengden av linjer gjennom origo i $\mathbb{R}^3$.

(a) Legg inn enhetssirkelen $\mathbb{S}^1$ i planet $\mathbb{R}^2$ som en sirkel av radius 1 med sentrum i $\left(0, \frac{1}{2}\right)$, som i Oppgave 2.28(a) (se den tilhørende figuren). Begrunn at enhver linje gjennom origo, bortsett fra x-aksen, skjærer $\mathbb{S}^1$ i nøyaktig ett punkt utenfor origo.

(b) Bruk observasjonen fra (a) til å begrunne at det finnes en naturlig bijeksjon mellom $\mathbb{P}^1_{\mathbb{R}}$ og $\mathbb{S}^1$.

(c) Konklueér at det finnes en naturlig bijeksjon mellom $\mathbb{P}^1_{\mathbb{R}}$ og *ettpunktskompaktifiseringen* $\mathbb{R} \cup \{\infty\}$ av $\mathbb{R}$ konstruert i Oppgave 2.28(c). (Av denne grunnen tenker vi ofte på den projektive linjen $\mathbb{P}^1_{\mathbb{R}}$ som ettpunktskompaktifiseringen av $\mathbb{R}$.)

(d) Legg inn enhetskulen $\mathbb{S}^2$ i rommet $\mathbb{R}^3$ som en sirkel av radius 1 med sentrum i $\left(0, 0, \frac{1}{2}\right)$, slik at "nordpolen" ligger i $(0, 0, 1)$, mens "sørpolen" ligger i origo, som i Oppgave 2.28(d) (se den tilhørende figuren). Begrunn at enhver linje gjennom origo som ikke er parallell med xy-planet skjærer $\mathbb{S}^2$ i nøyaktig ett punkt utenfor origo.

(e) Bruk observasjonen fra (d) til å konkludere at det finnes en én-til-én korrespondanse mellom elementene i $\mathbb{P}^2_{\mathbb{R}}$ som tilsvarer linjer gjennom origo som ikke ligger i xy-planet og punktene i $\mathbb{S}^2 \setminus \{0, 0, 0\}$.

(f) Konkludér at det finnes en naturlig bijeksjon mellom $\mathbb{P}^2_{\mathbb{R}}$ og mengden $\mathbb{R}^2 \cup \mathbb{P}^1_{\mathbb{R}}$. (Av denne grunnen ser vi på det projektive planet som en "utvidelse" av planet $\mathbb{R}^2$, av en litt annen type enn ettpunktskompaktifiseringen vi så i Oppgave 2.28(d).)

(g) Bruk samme fremgangsmåte til å vise at det finnes naturlige bijeksjoner mellom $\mathbb{P}^n_{\mathbb{R}}$ og $\mathbb{R}^n \cup \mathbb{P}^{n-1}_{\mathbb{R}}$ for hver $n \in \mathbb{Z}^+$. (Tilfellet $n = 1$ tilsvarer det vi fant i (c) siden $\mathbb{P}^0_{\mathbb{R}}$ er lik mengden av alle linjer i tallinjen gjennom origo, som er tallinjen selv, slik at $\mathbb{P}^0_{\mathbb{R}}$ består av ett enkelt punkt.)

KAPITTEL 3

Naturlige tall, heltall og rasjonale tall

Vi har i de to første kapitlene tatt skolematematikken for gitt. Dette gjelder spesielt eksistensen av og egenskapene til heltallene, de rasjonale tallene og de reelle tallene. Fra nå av skal vi imidlertid glemme det vi har lært på skolen og aksiomatisk bygge opp disse tallsystemene og deres egenskaper helt fra begynnelsen. Vi starter med de naturlige tallene, heltallene og de rasjonale tallene i dette kapitlet.

Opprinnelsen av de naturlige tallene er umulig å datere. Fra skriflige opptegnelser vet vi sikkert at både egypterne og babylonerne hadde avanserte systemer for å telle allerede 2000 f.Kr, og det samme gjelder mayaene i Sør-Amerika og sivilisasjoner i India og Kina. Tallenes opprinnelse er imidlertid trolig eldre: et interessant og diskutert funn er Ishango-benet fra Edward-sjøen på grensen mellom Kongo og Uganda. Dette er datert til å være ca. 20000 år gammelt og har risset i seg vertikale streker delt i forskjellige grupper. En teori går derfor ut på at benet har blitt brukt til telling.

Ishango-benet, utstilt ved Royal Belgian Institute of Natural Sciences
(Kilde: Wikipedia, public domain)

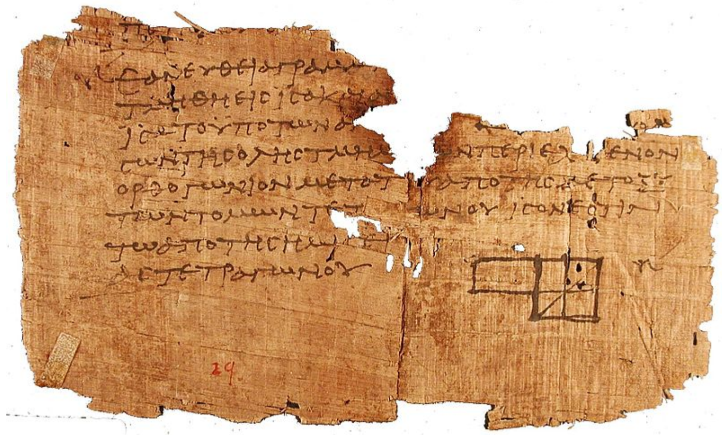
Både egypterne, babylonerne, inderne og kineserne kunne regne med brøker flere tusen år før Kristi fødsel. Interessant er det at egypterne bare godtok brøker på formen $1/n$, såkalte *stammbrøker*, og uttrykte alle andre brøker som sum av stammbrøker. De hadde riktignok egne symboler for 2/3 og 3/4, til hjelp for å forenkle uttrykk. Fremdeles i dag kalles en sum av stammbrøker for en *Egyptisk brøk*. Rhind-papyrusen fra ca. 1500 f.Kr. inneholder tabeller for alle brøkene på formen $2/n$ som Egyptisk brøk.

Et stykke av Rhind-papyrusen (Kilde: Wikipedia, public domain)

De gamle grekerne hadde så å si dagens intuitive forståelse for hva et rasjonalt tall $\frac{a}{b}$ er: tenk deg en fiksert lengdeenhet, som vi kaller 1. Del denne lengdeenheten i b like store deler (da trenger vi $b \neq 0$), og ta så a ganger denne mindre lengdeenheten. Lengden vi da får er $\frac{a}{b}$. Dette resonnementet bygde på de gamle grekernes teori om *kommensurable linjestykker*: to linjestykker er kommensurable hvis det alltid finnes et lite, tredje linjestykke som går opp et helt antall ganger i begge de to opprinnelige; eller sagt på en annen måte: forholdet mellom lengdene på to kommensurable linjestykker er et rasjonalt tall.

Den første systematiske studien av tall tilegnes vanligvis de gamle grekerne, også fordi dette er den første matematikktradisjonen hvor vi kjenner enkeltmenneskers bidrag og kan følge utviklingen med tiden. Spesielt bør nevnes bidragene til Pythagoras (ca. 570–495 f.Kr.) og Arkimedes (ca. 287–212 f.Kr.). Mye av denne matematikken er oppsummert i *Elementer* fra ca. 300 f.Kr., skrevet av Euklid. Verket omhandler først of fremst geometri (herfra kommer aksiomene for det som i dag kalles *euklidsk geometri*), men innholder også tallteori. Euklid viser blant annet uendeligheten av primtall (se Teorem 3.7.10) og en variant av *aritmetikkens fundamentalteorem* (se Teorem 3.7.7). Det store tallteoretiske verket innenfor gresk matematikk er imidlertid *Arithmetika* skrevet av Diofantus rundt år 250. Den greske storhetstiden varte til den kvinnelige matematikeren, astronomen og filosofen Hypatia[1] (ca. 350/370–415) ble myrdet av kristne fanatikere.

[1] Hypatia regnes ofte som den første kvinnelige matematikeren. Imidlertid mener mange at matematikeren Pandrosion av Alexandria, beskrevet i samlingene av Pappus av Alexandria (ca. 290–ca. 350) og kreditert med en metode for å beregne kubikkrøtter, var en kvinne. I så fall er Pandrosion den tidligste kvinnelige matematikeren vi kjenner til.



ET AV DE ELDSTE BEVARTE FRAGMENTENE AV EUKLIDS *Elementer*, FRA CA. ÅR 100. (KILDE: WIKIPEDIA, PUBLIC DOMAIN)

I India og Kina blomstret tallteorien gjennom oldtiden og langt inn i middelalderen. I Europa ble det araberne som overtok arven etter grekerne, etter at ekspansjonen av det arabiske riket begynte på 600-tallet. Blant annet ble Egypt erobret, og araberne fikk tilgang til mange gamle greske manuskripter. Både Euklids *Elementer* og Diofantus' *Arithmetika* (og bearbeidelsen av denne av Hypatia) ble oversatt. På 1200-tallet begynte den arabiske vitenskapen å forfalle, samtidig som universiteter begynte å dukke opp i mange europeiske byer. Det eldste nålevende er Universitetet i Bologna, som ble stiftet i 1088. Her ble for øvrig historiens første kvinnelige matematikkprofessor utnevnt i 1750, nemlig Maria Gaetana Agnesi (1718–1799)[2].

En algoritme for å uttrykke brøker som egyptiske brøker er gjengitt i boken *Liber Abaci*[3] (1202) av den italienske matematikeren Fibonacci (ca. 1170–1245), som de fleste i dag forbinder med *Fibonaccitallene* (se Eksempel 5.3.5). Interessant er det å merke seg at selv om de positive hele tallene som sagt var kjent i forhistorisk tid, og både inderne og kineserne tidlig tok i bruk negative tall, fikk ikke sistnevnte gjennomslag i Europa før på 15/1600-tallet.

Viktige matematikere ble etter hvert franskmennene Pierre de Fermat (1607–1665), Blaise Pascal (1623–1662), Adrien-Marie Legendre (1752–1833), sveitseren Leonhard Euler (1707–1783) og fransk-italieneren Joseph-Louis Lagrange (1736–1813). Det store gjennombruddet kom imidlertid i 1801, da den tyske matematikeren og fysikeren Johann Carl Friedrich Gauss (1777–1855) ga ut *Disquisitiones Arithmeticae*[4], hvor tallteorien slik vi kjenner den i dag ble grunnlagt. Gauss regnes som en av de fremste matematikerne gjennom tidene.

---

[2]Maria Agnesi er for øvrig historiens *andre* kvinnelige professor, og den første tilhører også Universitetet i Bologna, nemlig Maria Caterina Bassi (1711–1778), som ble utnevt til professor i vitenskapsfilosofi i 1732.

[3]Fra latin *liber=bok* og *abaci=beregninger.*

[4]Latin for *aritmetiske undersøkelser/diskusjoner.*

JOHANN CARL FRIEDRICH GAUSS (KILDE: WIKIPEDIA, PUBLIC DOMAIN)

Selv om forståelsen av de hele tallene og beregninger med dem rundt år 1800 var slik vi kjenner dem i dag, og bygde på en nesten 2000 år lang tradisjon, vokste det i løpet av 1800–tallet et behov for å introdusere mer logisk rigorøsitet i grunnleggende begreper innnefor matematikk, deriblant heltallene. Dette til tross for at det pågikk en intens debatt i matematiske og filosofiske kretser om naturen til heltallene. Den tyske matematikeren Leopold Kronecker (1823–1891) hevdet for eksempel at "Gud skapte heltallene, alt annet er menneskeverk"[5]. Den aksiomatiske oppbyggingen av de naturlige tallene, det vil si de *ikke-negative heltallene*, som etter hvert vant frem, kan tilegnes den amerikanske matematikeren Charles Sanders Peirce (1839–1914), den tyske matematikeren Richard Dedekind (1831–1916) og den italienske matematikeren Giuseppe Peano (1858–1932). Som vi vil se, går den formelle konstruksjonen av heltallene og de rasjonale tallene rimelig greit gitt eksistensen av de naturlige tallene.

Vi skal i dette kapitlet se på den formelle definisjonen av naturlige tall, hele tall og rasjonale tall, samt deres aritmetiske egenskaper (§3.1–3.5). Deretter skal vi se hvordan deres egenskaper passer inn i et større bilde innenfor matematikken, nemlig i mer generelle algebraiske strukturer som *ringer* og *kropper* (§3.6 og Kapittel 4). Vi tar også med en seksjon om *primtall* (§3.7), på grunn av den spesielle rollen disse har i matematikk, og om *binomialteoremet* (§3.8), siden det har mange bruksområder.

3.1. Aksiomatisk definisjon av de naturlige tallene

Det finnes forskjellige tradisjoner om hvorvidt tallet 0 skal inkluderes i de naturlige tallene. Den vanligste konvensjonen i matematikk og informatikk,

[5]"Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk".

som vi skal følge, er at 0 er med. Mengden av de naturlige tallene $\mathbb{N}$ kan defineres ved følgende aksiomer:

Peano-aksiomene for mendgen av heltall $\mathbb{N}$

(N1) Det finnes et element $0 \in \mathbb{N}$.
(N2) For enhver $n \in \mathbb{N}$ finnes et element $S(n) \in \mathbb{N}$ kalt *etterfølgeren*[a] til n.
(N3) For alle $m, n \in \mathbb{N}$ gjelder $S(m) = S(n)$ hvis og bare hvis $m = n$.
(N4) 0 er ikke etterfølgeren til et element i $\mathbb{N}$ (det vil si: det finnes ikke noen $n \in \mathbb{N}$ slik at $S(n) = 0$).
(N5) **Induksjonsaksiomet**[b]. Hvis $K \subset \mathbb{N}$ er en delmengde slik at
- $0 \in K$ og
- for alle $n \in K$ er også $S(n) \in K$,

da er $K = \mathbb{N}$.

[a] *Successor på engelsk*
[b] *Axiom of induction* på engelsk.

Merk at vi kan betrakte S som en funksjon

$$S : \mathbb{N} \longrightarrow \mathbb{N}$$

og da sier (N3) at S er injektiv, og (N4) sier at 0 ikke er med i verdimengden til S.

Merk at (N4) sier spesielt at $S(0) \neq 0$, det vi si at *etterfølgeren til* 0 *ikke er* 0 *selv.* Ut fra dette kan vi si at $\mathbb{N}$ inneholder ihvertfall to forskjellige elementer, nemlig 0 og $S(0)$. Betrakt $S(0)$. Ved (N4) igjen må vi ha at $S(S(0)) \neq 0$. Videre, siden S er injektiv ved (N3) og $S(0) \neq 0$, har vi at også $S(S(0)) \neq S(0)$. Dette betyr at vi har tre forskjellige elementer $0, S(0), S(S(0)) \in \mathbb{N}$. Fortsetter vi som dette, ser vi at alle elementene $0, S(0), S(S(0)), S(S(S(0)), \ldots$ er forskjellige elementer i $\mathbb{N}$. Dette betyr at aksiomene (N1)-(N4) gjør oss i stand til å generere uendelig mange forskjellige elementer i $\mathbb{N}$, det vil si

$$\{0, S(0), S(S(0)), S(S(S(0))), \ldots\} \subset \mathbb{N}. \tag{37}$$

Aksiomene (N1)-(N4)garanterer imidlertid *ikke* likhet i (37); til det trenger vi å vite at *alle* elementer i $\mathbb{N}$ utenom 0 er en etterfølger av et annet element i $\mathbb{N}$. Til dette trengs aksiom (N5): siden mengden til venstre i (37) oppfyller egenskapen til mengden K i (N5), gir (N5) at

$$\mathbb{N} = \{0, S(0), S(S(0)), S(S(S(0))), \ldots\}.$$

Med andre ord betyr dette at

(38) alle elementer i $\mathbb{N} \setminus \{0\}$ er på formen $S(n)$ for en $n \in \mathbb{N}$,

eller ekvivalent at

(39) verdimengden til S er $\mathbb{N} \setminus \{0\}$.

For å lette notasjonene gir vi hvert av elementene i mengden et eget tegn, slik vi kjenner:

$$\begin{aligned} S(0) &= 1 \\ S(1) &= 2 \\ S(2) &= 3 \\ S(3) &= 4 \\ S(4) &= 5 \end{aligned}$$

osv.

3.2. Aritmetikk på $\mathbb{N}$

La oss nå se på hvordan man definerer de velkjente regneoperasjonene *addisjon* og *multiplikasjon* på $\mathbb{N}$.

Addisjon på $\mathbb{N}$. Vi definerer en funksjon som vi kaller addisjon:

$$\begin{aligned} + : \mathbb{N} \times \mathbb{N} &\longrightarrow \mathbb{N} \\ (m, n) &\mapsto m + n \end{aligned}$$

ved de to egenskapene:

$$n + 0 = n \text{ for alle } n \in \mathbb{N}, \tag{40}$$

$$m + S(n) = S(m + n) \text{ for alle } m, n \in \mathbb{N}. \tag{41}$$

Det er kanskje ikke helt klart at de to reglene (40)-(41) definerer $m + n$ for alle $m, n \in \mathbb{N}$. Det viser vi i følgende setning:

Setning 3.2.1

Egenskapene (40)-(41) gir en veldefinert funksjon $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ ved $(m, n) \mapsto m + n$.

Bevis

Dette beviset er et godt eksempel på hvordan man bruker *induksjonsaksiomet* (N5) i bevis, en teknikk vi vil bruke flere ganger nedenunder.

Betrakt delmengden av $\mathbb{N}$ gitt ved

$$M = \{n \in \mathbb{N} \mid m + n \text{ er definert for alle } m \in \mathbb{N}\}.$$

Å vise setningen er ekvivalent med å vise at $M = \mathbb{N}$, og til dette bruker vi *induksjonsaksiomet* (N5). Vi vil vise at

- $0 \in M$,
- $n \in M \Longrightarrow S(n) \in M$.

Dersom vi gjør det, vil nemlig *induksjonsaksiomet* (N5) implisere at $M = \mathbb{N}$.

For alle $m \in \mathbb{N}$ er $m + 0 = 0$ ved (40), slik at $m + 0$ er definert. Dette viser at $0 \in M$.

Anta nå at $n \in M$, det vil si at $m + n$ er definert for alle $m \in \mathbb{N}$. Vi vil vise at da er også $S(n) \in M$, det vil si at også $m + S(n)$ er definert for alle $m \in \mathbb{N}$. Ved (41) har vi at $m + S(n) = S(m+n)$, og dette er definert, siden $m + n$ er definert. Altså er $S(n) \in M$.

Vi har dermed vist de to punktene vi ønsket å vise. *Induksjonsaksiomet* (N5) impliserer da at $M = \mathbb{N}$. □

For eksempel gir dette at:

$$\begin{aligned} n+1 &= n+S(0) && \text{(per definisjon av 1)} \\ &= S(n+0) && \text{(ved (41))} \\ &= S(n) && \text{(ved (40))}, \end{aligned} \tag{42}$$

altså er $n + 1$ etterfølgeren til n. Videre har vi:

$$\begin{aligned} n+2 &= n+S(1) && \text{(per definisjon av 2)} \\ &= S(n+1) && \text{(ved (41))} \\ &= S(S(n)) && \text{(ved (42))}, \end{aligned} \tag{43}$$

$$\begin{aligned} n+3 &= n+S(2) && \text{(per definisjon av 3)} \\ &= S(n+2) && \text{(ved (41))} \\ &= S(S(S((n))) && \text{(ved (43))}, \end{aligned} \tag{44}$$

osv. Dette gir oss den velkjente addisjonen på ℕ som vi er vant til.

EKSEMPEL 3.2.2. Vi regner ut $S(S(0)) + S(S(S(0)))$, eller $2 + 3$ i moderne notasjon, som følger:

$$\begin{aligned} S(S(0)) + S(S(S(0)) &= S(S(S(0)) + S(S(0))) && \text{(ved (41))} \\ &= S(S(S(S(0) + S(0)))) && \text{(ved (41))} \\ &= S(S(S(S(S(0) + 0))) && \text{(ved (41))} \\ &= S(S(S(S(S(0)))) && \text{(ved (40))}, \end{aligned}$$

som er det vi kaller 5.

Alle andre egenskaper ved addisjon som vi kjenner fra skolen utledes direkte fra aksiomene (40) og (41). Merk at det *ikke* er en del av aksiomene at operasjonen $+$ er *kommutativ*, det vil si at $m+n = n+m$ for alle $m, n \in \mathbb{N}$, og heller ikke at operasjonen er *assosiativ*, det vil si uavhengig av hvilke objekter som adderes først, eller sagt matematisk: $m+(n+r) = (m+n)+r$. Det er heller ikke en del av aksiomene at $0 + n = n$ for alle $n \in \mathbb{N}$. At alle disse egenskapene gjelder må vises ved hjelp av *induksjonsaksiomet* (N5). La oss gjøre det.

Setning 3.2.3

$0 + n = n$ for alle $n \in \mathbb{N}$.

BEVIS

Vi bruker *induksjonsaksiomet* (N5).

La M være delmengden av $\mathbb{N}$ bestående av alle $n \in \mathbb{N}$ som oppfyller $0 + n = n$. Med andre ord:

$$M = \{n \in \mathbb{N} \mid 0 + n = n\}.$$

Vi ønsker å vise at $M = \mathbb{N}$. Ved *Induksjonsaksiomet* (N5) holder det å vise at $0 \in M$ og at $S(n) \in M$ hvis $n \in M$. La oss gjøre det.

Siden $0 + 0 = 0$ ved (40), er $0 \in M$.

Hvis $n \in M$, da har vi

$$\begin{array}{rcll} 0 + S(n) & = & S(0 + n) & \text{(ved (41))} \\ & = & S(n) & \text{(siden } n \in M\text{)}, \end{array}$$

slik at også $S(n) \in M$. *Induksjonsaksiomet* (N5) gir derfor at $M = \mathbb{N}$, som beviser setningen. □

Ved setningen og (40) har vi altså at

$$n + 0 = 0 + n = n \text{ for alle } n \in \mathbb{N}, \tag{45}$$

og vi uttrykker dette gjerne ved å si at elementet 0 er et *nøytralt element for addisjon*.

For å vise kommutativitet av addisjon, trenger vi et resultat på veien:

Hjelpesetning 3.2.4

$m + S(n) = S(m) + n$ for alle $m, n \in \mathbb{N}$.

Bevis

Vi bruker *induksjonsaksiomet* (N5) igjen. La

$$M = \{n \in \mathbb{N} \mid m + S(n) = S(m) + n \text{ for alle } m \in \mathbb{N}\}.$$

Vi ønsker å vise at $M = \mathbb{N}$. Ved *Induksjonsaksiomet* (N5) holder det å vise at $0 \in M$ og at $S(n) \in M$ hvis $n \in M$.

Siden

$$\begin{array}{rcll} m + S(0) & = & S(m + 0) & \text{(ved (41))} \\ & = & S(m) & \text{(ved (40))} \\ & = & S(m) + 0 & \text{(ved (40))}, \end{array}$$

er $0 \in M$.

Hvis $n \in M$, da har vi

$$\begin{array}{rcll} m + S(S(n)) & = & S(m + S(n)) & \text{(ved (41))} \\ & = & S(S(m) + n) & \text{(siden } n \in M\text{)} \\ & = & S(m) + S(n) & \text{(ved (41))}, \end{array}$$

slik at også $S(n) \in M$. *Induksjonsaksiomet* (N5) gir derfor at $M = \mathbb{N}$, som beviser setningen. □

Setning 3.2.5: Kommutativitet av addisjon

Addisjon i $\mathbb{N}$ er *kommutativ*[a]: For alle $m, n \in \mathbb{N}$ er $m + n = n + m$.

[a] *Commutative* på engelsk.

Bevis
Vi bruker *induksjonsaksiomet* (N5) nok en gang, sammen med hjelpesetningen. La

$$M = \{n \in \mathbb{N} \mid m + n = n + m \text{ for alle } m \in \mathbb{N}\}.$$

Ved (45) har vi at $0 \in M$.
Anta $n \in M$. Da er

$$\begin{aligned} m + S(n) &= S(m) + n && \text{(ved Hjelpesetning 3.2.4)} \\ &= n + S(m) && \text{(siden } n \in M) \\ &= S(n) + m && \text{(ved Hjelpesetning 3.2.4),} \end{aligned}$$

slik at også $S(n) \in M$. *Induksjonsaksiomet* (N5) gir derfor at $M = \mathbb{N}$, som beviser setningen. □

Setning 3.2.6: Assosiativitet av addisjon

Addisjon i $\mathbb{N}$ er *assosiativ*[a]: For alle $m, n, r \in \mathbb{N}$ har vi at $m + (n + r) = (m + n) + r$.

[a] *Associative* på engelsk.

Bevis
Dette ligner på bevisene ovenfor og overlates til Oppgave 3.2. □

På grunn av assosiativiteten, skriver vi ofte kun $m + n + r$ istedenfor $m + (n + r) = (m + n) + r$.

Multiplikasjon på ℕ. Vi kan videre definere en funksjon som vi kaller multiplikasjon:

$$\begin{aligned} \cdot : \mathbb{N} \times \mathbb{N} &\longrightarrow \mathbb{N} \\ (m, n) &\mapsto m \cdot n \end{aligned}$$

ved de to egenskapene:

$$n \cdot 0 = 0 \text{ for alle } n \in \mathbb{N} \tag{46}$$

$$n \cdot S(m) = n + (n \cdot m) \text{ for alle } m, n \in \mathbb{N} \tag{47}$$

Igjen må vi vise at de to reglene (46)-(47) definerer $m \cdot n$ for alle $m, n \in \mathbb{N}$:

Setning 3.2.7

Egenskapene (46)-(47) gir en veldefinert funksjon $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ ved $(m, n) \mapsto m \cdot n$.

Bevis
Dette ligner på beviset for Setning 3.2.1 og overlates til Oppgave 3.3. □

Som en konsekvens kan vi utlede at $1 = S(0)$ er den *multiplikative identiteten* til operasjonen $\cdot$:

Setning 3.2.8: Multiplikativ identitet

For alle $n \in \mathbb{N}$ er $n \cdot S(0) = S(0) \cdot n = n$.

Bevis
Vi har
$$n \cdot S(0) \overset{(47)}{=} n + (n \cdot 0) \overset{(46)}{=} n + 0 \overset{(40)}{=} n.$$
For å vise at $S(0) \cdot n = n$ bruker vi atter *induksjonsaksiomet* (N5). La
$$M = \{n \in \mathbb{N} \mid S(0) \cdot n = n\}.$$
Siden $S(0) \cdot 0 = 0$ ved (46), er $0 \in M$.
Anta nå at $n \in M$, det vil si at $S(0) \cdot n = n$. Da er
$$\begin{array}{rcll} S(0) \cdot S(n) & = & S(0) + (S(0) \cdot n) & \text{(ved (47))} \\ & = & S(0) + n & \text{(ved antagelsen } n \in M) \\ & = & n + S(0) & \text{(ved komm. av addisjon, Setn. 3.2.5)} \\ & = & S(n+0) & \text{(ved (41))} \\ & = & S(n) & \text{(ved (40)).} \end{array}$$
Altså er $S(n) \in M$. Ved *induksjonsaksiomet* (N5) er $M = \mathbb{N}$. □

Man kan utlede følgende egenskaper:

Setning 3.2.9: Egenskaper til multiplikasjon

(i) Multiplikasjon i $\mathbb{N}$ er *kommutativ*: $m \cdot n = n \cdot m$ for alle $m, n \in \mathbb{N}$.
(ii) Multiplikasjon i $\mathbb{N}$ er *assosiativ*: $m \cdot (n \cdot r) = (m \cdot n) \cdot r$ for alle $m, n, r \in \mathbb{N}$.
(iii) Den *distributive lov*[a] gjelder: $r \cdot (m + n) = r \cdot m + r \cdot n$ for alle $m, n, r \in \mathbb{N}$.

[a] *Distributive law* på engelsk.

Bevis
Dette ligner på bevisene ovenfor og overlates til Oppgave 3.4. □

Merk at vi også har:

Setning 3.2.10

For $a, n \in \mathbb{N}$ gjelder:
$$\underbrace{a + \cdots + a}_{n \text{ ganger}} = n \cdot a.$$

Bevis
Ethvert naturlig tall n oppnås ved å anvende funksjonen S på 0 et visst antall ganger (n ganger):
$$S(S(S(\ldots S(0)\ldots))) = n.$$

Samtidig vet vi at $S(0) = 1$ (per definisjon), $S(1) = 1 + 1$ (ved (42)), $S(1+1) = 1+1+1$ (ved (43)), osv., slik at

$$\underbrace{1 + \cdots + 1}_{n \text{ ganger}} = n.$$

Ganger vi begge sider av likheten med a og bruker den distributive loven (n ganger), får vi likheten i setningen. □

Merk at vi ofte dropper multiplikasjonstegnet når det ikke fører til misforståelser og skriver for eksempel kun na for $n \cdot a$. På grunn av assosiativiteten, skriver vi ofte kun mnr istedenfor $m(nr) = (mn)r$. På grunn av kommutativitet og assosiativitet av multiplikasjon ikke trenger å bry oss om rekkefølgen av multiplikasjon. Vi pleier å bruke notasjonen

$$n^m = \underbrace{n \cdot n \cdots n}_{m \text{ ganger}},$$

slik at vi for eksempel skriver

$$3 \cdot 2 \cdot 5 \cdot 2 \cdot 3 \cdot 2 = 2^3 \cdot 3^2 \cdot 5.$$

Ytterligere egenskaper. Følgende lover, velkjent fra skolen, vil være nyttige i fortsettelsen:

Setning 3.2.11: Kanselleringslovene

For alle $m, n, r \in \mathbb{N}$ gjelder:

(i) Hvis $m + n = r + n$, da er $m = r$;
(ii) Hvis $m + n = n$, da er $m = 0$;
(iii) Hvis $mn = rn$ og $n \neq 0$, da er $m = r$;
(iv) Hvis $mn = 0$, da er $m = 0$ eller $n = 0$.

Bevis
(i) Vi vil vise utsagnet

$$m + n = r + n \Longrightarrow m = r \tag{48}$$

for alle $m, n, r \in \mathbb{N}$. La $M = \{n \in \mathbb{N} \mid (48) \text{ er oppfylt for alle } m, r \in \mathbb{N}\}$. Vi vil vise at $M = \mathbb{N}$ og vil bruke *induksjonsaksiomet* (N5) nok en gang.

Vi har at $0 \in M$, siden $m + 0 = m$ og $r + 0 = r$ ved (40).

Anta nå at $n \in M$. Hvis $m + S(n) = r + S(n)$, da er $S(m+n) = S(r+n)$ ved (41), og dermed $m + n = r + n$ ved Peanoaksiom (N3). Siden $n \in M$, medfører dette at $m = r$. Dermed har vi vist at $S(n) \in M$. Det følger at $M = \mathbb{N}$ ved *induksjonsaksiomet* (N5).

(ii) Dette er et spesialtilfelle av (i) ved å sette $r = 0$.

(iv) Vi viser dette punktet før (iii), siden vi vil trenge det i beviset av (iii). Vi utfører et kontrapositivt bevis og viser at dersom $m, n \neq 0$, da er også $mn \neq 0$. Hvis $m, n \neq 0$, da kan vi på grunn av (38) skrive

$m = S(m_1)$ og $n = S(n_1)$ for $m_1, n_1 \in \mathbb{N}$. Dermed har vi:

$$\begin{array}{rcll} mn & = & S(m_1)S(n_1) & \\ & = & S(m_1) + S(m_1)n_1 & \text{(ved (47))} \\ & = & S(m_1)n_1 + S(m_1) & \text{(ved komm. av addisjon, Setn. 3.2.5)} \\ & = & S(S(m_1)n_1 + m_1) & \text{(ved (41))} \\ & \neq & 0 & \text{(ved (N4)).} \end{array}$$

(iii) Vi vil vise utsagnet

$$mn = rn \Longrightarrow m = r \tag{49}$$

for alle $m, r \in \mathbb{N}$, $n \in \mathbb{N} \setminus \{0\}$. La

$$M = \{m \in \mathbb{N} \mid \text{(49) er oppfylt for alle } r \in \mathbb{N} \text{ og } n \in \mathbb{N} \setminus \{0\}\}.$$

Vi vil vise at $M = \mathbb{N}$ og vil bruke *induksjonsaksiomet* (N5) nok en gang.

Hvis $0n = rn$, da er $0 = rn$ ved (46) og kommutativitet av multiplikasjon (Setning 3.2.9(i)). Hvis $n \neq 0$, da må $r = 0$ ved (iv). Dette betyr at $0 \in M$.

Anta nå at $m \in M$. Vil vil vise at $S(m) \in M$, og vil derfor vise at hvis $S(m)n = rn$ for $r \in \mathbb{N}$, $n \in \mathbb{N} \setminus \{0\}$, da er $S(m) = r$. Siden $S(m) \neq 0$ ved Peanoaksiom (N4) og $n \neq 0$ per antagelse, er $S(m)n \neq 0$ ved (iv), og dermed også $rn \neq 0$. Da må $r \neq 0$ ved (46) og kommutativitet av multiplikasjon (Setning 3.2.9(i)). Vi kan derfor skrive $r = S(r_1)$ for en $r_1 \in \mathbb{N}$ ved (38). Dette betyr at vi har $S(m)n = S(r_1)n$, som ved (47) og kommutativitet av multiplikasjon (Setning 3.2.9(i)) kan skrives som $mn + n = r_1 n + n$. Ved (i) har vi da $mn = r_1 n$. Siden $m \in M$, medfører dette at $m = r_1$. Da er $S(m) = S(r_1) = r$, som er det vi ønsket å vise.

Det følger at $M = \mathbb{N}$ ved *induksjonsaksiomet* (N5). □

Vi tar også med følgende:

Setning 3.2.12

For alle $m, n \in \mathbb{N}$ gjelder:

(i) Hvis $m + n = 0$, da er $m = n = 0$;
(ii) Hvis $mn = 1$, da er $m = n = 1$. (Husk at $1 = S(0)$ per definisjon.)

Bevis

(i) Anta $m + n = 0$. Vi vil først vise at $m = 0$ og argumenterer ved selvmotsigelse. Dersom $m \neq 0$, kan vi skrive $m = S(m_1)$ for en $m_1 \in \mathbb{N}$, ved (38). Da er $0 = m + n = S(m_1) + n = S(m_1 + n)$, ved kommutativitet av addisjon (Setning 3.2.5) og (40). Dette motsier Peanoaksiom (N4). Dette viser at $m = 0$. Ved symmetri, siden addisjon er kommutativ, har vi også $n = 0$.

(ii) Anta $mn = 1$. Ved kommutativitet av multiplikasjon (Setning 3.2.9(i)) og (46), har vi at både $m \neq 0$ og $n \neq 0$. Vi kan derfor ved (38) skrive $m = S(m_1)$ og $n = S(n_1)$ for $m_1, n_1 \in \mathbb{N}$. Da har vi

$$S(0) = 1 = mn = mS(n_1) \overset{(47)}{=} m + mn_1 = S(m_1) + mn_1 \overset{(47)}{=} S(m_1 + mn_1)$$

(hvor vi også har brukt kommutativitet av addisjon). Ved Peanoaksiom (N3) er $0 = m_1 + mn_1$. Det følger fra (i) at $m_1 = 0$ og $mn_1 = 0$, og siden $m \neq 0$, medfører dette siste at $n_1 = 0$ ved Setning 3.2.11(iv). Dermed er $m = S(m_1) = S(0) = 1$ og likeledes $n = S(n_1) = S(0) = 1$. □

Siste setning har et par interessante konsekvenser. Del (i) sier at en $n \in \mathbb{N} \setminus \{0\}$ ikke kan en *additiv invers*, det vi si det ikke finnes noen $m \in \mathbb{N}$ slik at $m+n = 0$. Del (ii) sier at en $n \in \mathbb{N} \setminus \{0, 1\}$ ikke kan ha en *multiplikativ invers*, det vi si at det ikke finnes noen $m \in \mathbb{N}$ slik at $mn = 1$. De hele tall og de rasjonale tall vi skal konstruere om litt vil bøte på disse manglene.

Representasjon av de naturlige tall i 10-tallsystemet. Som nevnt tidligere, gir vi egne symboler til de naturlige tallene, istedenfor å uttrykke dem ved $S(S(S(\ldots S(0)\ldots)))$:

$$\begin{aligned} S(0) &= 1 \\ S(1) &= 2 \\ S(2) &= 3 \\ S(3) &= 4 \\ S(4) &= 5 \\ S(5) &= 6 \\ S(6) &= 7 \\ S(7) &= 8 \\ S(8) &= 9 \end{aligned}$$

Man har så funnet at det er praktisk å ikke innføre flere tallsymboler, men setter $S(9) = 10$ og uttrykker videre alle resterende naturlige tall kun ved hjelp av tegnene $0, 1, \ldots, 9$: med det naturlige tallet $s_n s_{n-1} s_{n-2} \ldots s_3 s_2 s_1 s_0$ mener vi tallet

$$s_n s_{n-1} s_{n-2} \ldots s_3 s_2 s_1 s_0 = 10^n s_n + 10^{n-1} s_{n-1} + \cdots + 10^2 s_2 + 10 s_1 + s_0.$$

Dette kalles *representasjonen av tallet i 10-tallssystemet* og tallene $s_0, \ldots, s_n$ kalles *siffre*[6].

De fleste gamle sivilisasjoner brukte 10-tallssystemer, muligens fordi mennesker har 10 fingre. Et interessant unntak er Maya-sivilisasjonen fra Sør-Amerika, som brukte et 20-tallssystem (kanskje fordi de også telte med tærne). Symbolene $0, \ldots, 9$ som vi bruker i dag kalles det *hindu–arabiske tallsystemet*, og ble utviklet i tiden 100-400 e.Kr. av indiske matematikere, for så å bli brukt av arabiske matematikere, som brakte dem til Europa fra

[6] *Digits* på engelsk.

år 1000 og utover. Heldigvis, kan man si, for romertallene man hadde brukt frem til da er ikke så praktiske å regne med.

3.3. De hele tall

Vi har sett at de naturlige tallene $\mathbb{N}$ inneholder et bestemt element 0 som er spesielt på to måter:

- det er det eneste tallet i $\mathbb{N}$ som ikke er etterfølgeren av et annet tall;
- det fungerer som et *nøytralt element* for addisjon, jf. (45).

Det som imidlertid mangler i $\mathbb{N}$ er en måte å "nøytralisere et ikkenull element ved addisjon", det vil si gitt en $n \in \mathbb{N} \setminus \{0\}$, da finnes ingen $m \in \mathbb{N}$ slik at $m + n = 0$, som vi observerte rett etter Setning 3.2.12. Men vi kan utvide mengden $\mathbb{N}$ ved å *aksiomatisk* legge til disse elementene, slik at vi får *mengden av de hele tallene*, som har fått symbolet $\mathbb{Z}$ etter det tyske ordet "Zahlen" for tall (i flertall). Dette kan vi gjøre på følgende måte:

Intuitivt vil det hele gå ut på at tallparet $(a, b) \in \mathbb{N}^2$ vil beskrive tallet som resulterer ved å subtrahere b fra a. Da vil altså for eksempel $(1, 2)$ og $(4, 5)$ representere samme tall, nemlig det vi til vanlig kaller -1. Vi introduserer derfor en ekvivalensrelasjon på $\mathbb{N}^2$ som gjør disse parene ekvivalente, som vi nå skal beskrive.

Vi definerer ekvivalensrelasjonen $\sim$ på $\mathbb{N}^2 = \mathbb{N} \times \mathbb{N}$ ved at

$$(a, b) \sim (c, d) \iff a + d = b + c. \tag{50}$$

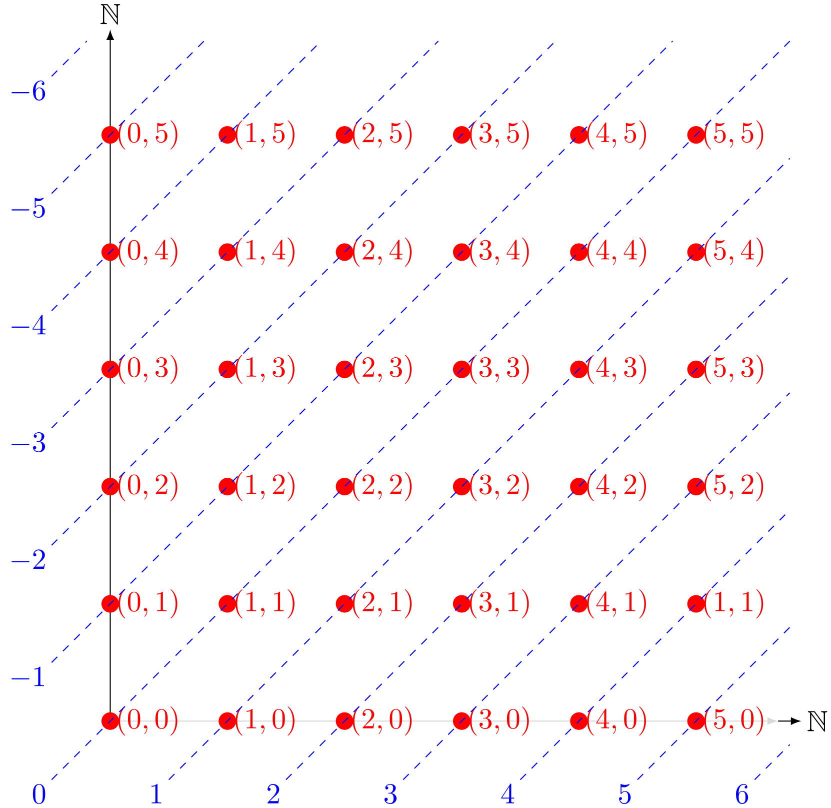
Å sjekke at dette virkelig er en ekvivalensrelasjon (jf. Definisjon 2.4.6), overlates til Oppgave 3.5. Så definerer vi $\mathbb{Z}$, *mengden av heltallene*, til å være mengden av ekvivalensklasser til denne relasjonen.

Eksempel 3.3.1. Vi har $(1, 2) \sim (4, 5)$, siden $1 + 5 = 2 + 4$. Mer generelt er $(1, 2) \sim (n, n + 1)$ for alle $n \in \mathbb{N}$, siden $1 + (n + 1) = 2 + n$. For eksempel $(0, 1)$ en representant for ekvivalensklassen, det vil si $[(n, n + 1)] = [(0, 1)]$, der klammeparentesene som vanlig betegner ekvivalensklasser.

Noen flere eksempler er, når vi som vanlig betegner ekvivalensklassen til (a, b) med $[(a, b)]$:

$$\begin{array}{ccccccc}
[(0,0)] & = & [(1,1)] & = & \cdots & = & [(k,k)] \\
[(1,0)] & = & [(2,1)] & = & \cdots & = & [(k+1,k)] \\
[(0,1)] & = & [(1,2)] & = & \cdots & = & [(k,k+1)] \\
[(2,0)] & = & [(3,1)] & = & \cdots & = & [(k+2,k)] \\
[(0,2)] & = & [(1,3)] & = & \cdots & = & [(k,k+2)] \\
[(3,0)] & = & [(4,1)] & = & \cdots & = & [(k+3,k)] \\
[(0,3)] & = & [(1,4)] & = & \cdots & = & [(k,k+3)]
\end{array}$$

Følgende figur oppsummerer situasjonen: Alle par $(a, b) \in \mathbb{N} \times \mathbb{N}$ i samme ekvivalensklassene ligger på samme diagonal, markert i blått.



VISUALISERING AV $\mathbb{Z}$

Det er lett å se, som ovenfor, at alle ekvivalensklassene har et entydig element på formen $(n, 0)$ eller $(0, n)$ med $n \in \mathbb{N}$ som representant (eller begge, i tilfellet $n = 0$). På figuren er dette representantene lengst nede langs den horisontale aksen og lengst til venstre langs den vertikale aksen. Vi har altså:

$$\mathbb{Z} = \{[(n, 0)] \mid n \in \mathbb{N}\} \cup \{[(0, n)] \mid n \in \mathbb{N} \setminus \{0\}\}. \tag{51}$$

Funksjonen

$$\begin{array}{ccc} \mathbb{N} & \longrightarrow & \mathbb{Z} \\ n & \mapsto & [(n, 0)] \end{array}$$

er injektiv (jf. Definisjon 2.3.1) og gjør at vi kan identifisere $\mathbb{N}$ med en delmengde av $\mathbb{Z}$, nemlig verdimengden av denne funksjonen. Gjennom denne identifiseringen sier vi at $\mathbb{N}$ er en delmengde av $\mathbb{Z}$, og skriver $\mathbb{N} \subset \mathbb{Z}$. Ser vi på (51), så er $\mathbb{N} = \{[(n, 0)] \mid n \in \mathbb{N}\}$ gjennom denne identifiseringen. På grunn av dette, og for å forenkle notasjonen, betegner vi $[(n, 0)]$ med n.

Den resterende mengden $\mathbb{Z} \setminus \mathbb{N} = \{[(0, n)] \mid n \in \mathbb{N} \setminus \{0\}\}$ er det som vi kaller *de negative heltallene*, og denne mengden betegnes gjerne med $\mathbb{Z}^-$. For å forenkle notasjonen betegner vi $[(0, n)]$ med $-n$.

Med denne notasjonen er det blå tallet i enden av diagonalen på figuren ovenfor uttrykket for ekvivalensklassen.

Mengden $\mathbb{N} \setminus \{0\}$ kaller vi *de positive heltallene*, og vi betegner denne gjerne med $\mathbb{Z}^+$. Altså har vi

$$\mathbb{Z} = \mathbb{Z}^- \cup \{0\} \cup \mathbb{Z}^+.$$

Aritmetikk på $\mathbb{Z}$. Operasjonene addisjon og multiplikasjon som vi definerte på $\mathbb{N}$ kan lett utvides til mengden $\mathbb{Z}$:

Definisjon 3.3.2: Addisjon og multiplikasjon på $\mathbb{Z}$

Vi definerer to operasjoner $+, \cdot : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$, kalt henholdsvis *addisjon* og *multiplikasjon*, ved:

- $[(a,b)] + [(c,d)] = [(a+c, b+d)]$.
- $[(a,b)] \cdot [(c,d)] = [(a \cdot c + b \cdot d, a \cdot d + b \cdot c)]$.

Merk at vi må vise at disse operasjonene er veldefinert. Vi har nemlig definert uttrykkene ved å velge representanter for ekvivalensklassene og vi må dermed vise at svarene blir de samme hvis vi velger andre representanter:

Setning 3.3.3

Dersom $[(a,b)] = [(a',b')]$ og $[(c,d)] = [(c',d')]$, da er

(i) $[(a+c, b+d)] = [(a'+c', b'+d')]$, og

(ii) $[(a \cdot c + b \cdot d, a \cdot d + b \cdot c)] = [(a' \cdot c' + b' \cdot d', a' \cdot d' + b' \cdot c')]$.

Bevis

At $[(a,b)] = [(a',b')]$ og $[(c,d)] = [(c',d')]$ betyr at $(a,b) \sim (a',b')$ og $(c,d) \sim (c',d')$, som igjen per definisjon betyr at

$$a + b' = b + a' \quad \text{og} \quad c + d' = d + c'. \tag{52}$$

Dette er vår hypotese, som vi kan bruke for å vise (i) og (ii).

(i) Adderer vi venstre og høyre side av likhetene i (52) og bruker kommutativitet av addisjon på $\mathbb{N}$ (Setning 3.2.5), får vi

$$\begin{aligned} a + b' + c + d' &= b + a' + d + c' \\ (a+c) + (b'+d') &= (b+d) + (a'+c'), \end{aligned}$$

og siste linje er per definisjon det samme som $(a+c, b+d) \sim (a'+c', b'+d')$, som igjen betyr at $[(a+c, b+d)] = [(a'+c', b'+d')]$.

(ii) Å vise at $[(a \cdot c + b \cdot d, a \cdot d + b \cdot c)] = [(a' \cdot c' + b' \cdot d', a' \cdot d' + b' \cdot c')]$ er ekvivalent med å vise at

$$ac + bd + a'd' + b'c' = ad + bc + a'c' + b'd'. \tag{53}$$

For å dedusere denne likheten starter vi med venstre side og adderer noe som gjør at vi kan benytte oss av (52). Vi bruker også kommutativitet av addisjon og multiplikasjon på $\mathbb{N}$ (Setningene 3.2.5 og 3.2.9(i)).

$$\begin{aligned}
(ac+bd+a'd'+b'c') + (a'c+b'c+a'd+b'd) &= \\
\underbrace{(a+b')}_{\overset{\|}{(b+a')}}c + \underbrace{(b+a')}_{\overset{\|}{(a+b')}}d + a'\underbrace{(c+d')}_{\overset{\|}{(d+c')}} + b'\underbrace{(d+c')}_{\overset{\|}{(c+d')}} &= \\
(b+a')c + (a+b')d + a'(d+c') + b'(c+d') &= \\
(ad+bc+a'c'+b'd') + (a'c+b'c+a'd+b'd).&
\end{aligned}$$

sammenligner vi første og siste uttrykk i kjeden og benytter oss av kanselleringsloven i Setning 3.2.11(i), får vi nå (53), som ønsket. □

Operasjonene i Definisjon 3.3.2 er kommutative, assosiative og følger den distributive lov, siden operasjonene oppfylte disse betingelsene på $\mathbb{N}$.

Med denne definisjonen har vi et nøytralt element for addisjon i $\mathbb{Z}$ og ethvert element i $\mathbb{Z}$ en *additiv invers*, i motsetning til $\mathbb{N}$:

Setning 3.3.4

(i) Elementet $[(0,0)]$ er et nøytralt element for addisjon på $\mathbb{Z}$, det vil si at $[(0,0)] + [(a,b)] = [(a,b)]$ for alle elementer $[(a,b)] \in \mathbb{Z}$.

(ii) Ethvert element $[(a,b)] \in \mathbb{Z}$ har $[(b,a)]$ som en additiv invers; med andre ord: $[(b,a)] + [(a,b)] = [(0,0)]$.

Bevis

Dette følger omtrent umiddelbart av definisjonene og overlates til Oppgave 3.6. □

Husker vi at ethvert element utenom $[(0,0)]$ kan representeres ved $[(n,0)]$ eller $[(0,n)]$ for en $n \in \mathbb{N} \setminus \{0\}$, og at vi betegner $[(n,0)]$ med n og $[(0,n)]$ med $-n$, ser vi at egenskapen (ii) i siste setning kan uttrykkes som

$$n + (-n) = 0 \ \text{ for alle } \ n \in \mathbb{N}.$$

Videre sier Definisjon 3.3.2 at:

$$\begin{aligned}
m+n &= [(m,0)] + [(n,0)] = [(m+n,0)] \\
m+(-n) &= [(m,0)] + [(0,n)] = [(m,n)] \\
(-m)+(-n) &= [(0,m))] + [(0,n)] = [(0,m+n)] = -(m+n) \\
m\cdot n &= [(m,0)]\cdot[(n,0)] = [(m\cdot n,0)] \\
m\cdot(-n) &= [(m,0)]\cdot[(0,n)] = [(0,m\cdot n)] = -(m\cdot n) \\
(-m)\cdot(-n) &= [(0,m))]\cdot[(0,n)] = [(m\cdot n,0)] = m\cdot n.
\end{aligned}$$

Merk spesielt at:

$$m,n \in \mathbb{Z}^+ \implies m+n \in \mathbb{Z}^+, \tag{54}$$

$$m,n \in \mathbb{Z}^- \implies m+n \in \mathbb{Z}^-, \tag{55}$$

$$mn \in \mathbb{Z}^+ \iff m,n \in \mathbb{Z}^+ \text{ eller } m,n \in \mathbb{Z}^-, \tag{56}$$

$$mn = 1 \iff m = n = 1 \text{ eller } m = n = -1 \tag{57}$$

(hvor det siste følger fra Setning 3.2.12(ii)), og dette er det vi er vant til. Slik får vi altså representasjonen av heltallene med de to operasjonene $+$ og $\cdot$ som vi kjenner i dag. Vi merker også at (57) sier at det kun er 1 og -1 i $\mathbb{Z}$ som har multiplikative inverser.

Vi avslutter med å oppsummere de aritmetiske egenskapene til $\mathbb{Z}$:

Setning 3.3.5: Egenskaper til addisjon og multiplikasjon på $\mathbb{Z}$

Operasjonene $+ : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ og $\cdot : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ på $\mathbb{Z}$ oppfyller egenskapene:

- (A1) Addisjon er kommutativ: $a + b = b + a$ for alle $a, b \in \mathbb{Z}$;
- (A2) Addisjon er assosiativ: $(a+b)+c = a+(b+c)$ for alle $a, b, c \in \mathbb{Z}$;
- (A3) $\mathbb{Z}$ inneholder et element 0 (kalt *nøytralt element*) som oppfyller $0 + a = a + 0 = a$ for alle $a \in \mathbb{Z}$;
- (A4) Til ethvert element $a \in \mathbb{Z}$ finnes et element $-a \in \mathbb{Z}$ (kalt *det additive inverselementet til* a) slik at $a + (-a) = (-a) + a = 0$;
- (M1) Multiplikasjon er kommutativ: $a \cdot b = b \cdot a$ for alle $a, b \in \mathbb{Z}$;
- (M2) Multiplikasjon er assosiativ: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for alle $a, b, c \in \mathbb{Z}$;
- (M3) $\mathbb{Z}$ inneholder et element $1 \neq 0$ (kalt *identitetselement*) som oppfyller $1 \cdot a = a \cdot 1 = a$ for alle $a \in \mathbb{Z}$;
- (D) Den *distributive loven* $a \cdot (b + c) = a \cdot b + a \cdot c$ holder for alle $a, b, c \in \mathbb{Z}$.

På samme måte som før dropper vi gjerne multiplikasjonstegnet og skriver mn istedenfor $m \cdot n$ når det ikke er fare for misforståelser. Dessuten pleier vi også å forenkle og skrive

$$m + (-n) = m - n,$$

som vi er vant til fra skolen.

3.4. De rasjonale tallene

Vi vil nå definere de rasjonale tallene aksiomatisk. Vi kan starte med å observere at det finnes en asymmetri mellom egenskapene oppfylt av addisjon og multiplikasjon i $\mathbb{Z}$ i Setning 3.3.5. Egenskapen tilsvarende (A4) mangler for multiplikasjon: det finnes nemlig ikke noe multiplikativt invers til noe tall i $\mathbb{Z}$ utenom 1 og -1; med andre ord, dersom $a \neq \pm 1$, finnes ingen $b \in \mathbb{Z}$ slik at $ab = 1$, som vi bemerket mot slutten av forrige seksjon rett etter (57). På samme måte som vi utvidet de naturlige tallene til de hele tallene ved å legge til additive inverser, kan vi utvide heltallene til de rasjonale tall ved å legge til multiplikative inverser. For å gjøre dette presist og utvide operasjonene addisjon og multiplikasjon, kan vi tenke litt på hva et rasjonalt tall $\frac{a}{b}$ er for oss i dag: det representeres ved et tallpar $(a, b) \in \mathbb{Z} \times \mathbb{Z} \setminus \{0\}$, men noen tallpar gir oss samme rasjonale tall: vi vet nemlig at

$$\frac{a}{b} = \frac{c}{d} \iff ad - bc = 0. \tag{58}$$

Dette tar vi som utgangspunkt.

Vi definerer ekvivalensrelasjonen $\sim$ på $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$ ved at

$$(a, b) \sim (c, d) \iff ad - bc = 0. \tag{59}$$

Å vise at dette virkelig er en ekvivalensrelasjon (jf. Definisjon 2.4.6) overlates til Oppgave 3.8. Så definerer vi $\mathbb{Q}$, *mengden av de rasjonale tallene*, til å være mengden av ekvivalensklasser til denne relasjonen. Bokstaven $\mathbb{Q}$ ble først brukt av Giuseppe Peano etter det italienske ordet "quoziente", som betyr "kvotient". Ekvivalensklassen til (a, b) betegnes med $\frac{a}{b}$. Da er (58) ekvivalent med (59). Tallet a kalles *telleren*[7], mens b kalles *nevneren*[8].

For alle $a, b, n \in \mathbb{Z}$ er $anb - bna = 0$, slik at (58) gir at $\frac{a}{b} = \frac{an}{bn}$. Dette er den velkjente regelen om at vi kan forkorte felles faktorer i teller og nevner. Spesielt innebærer dette at ethvert rasjonalt tall kan uttrykkes på en entydig måte som en *irredusibel brøk*, det vil si som $\frac{a}{b}$ der a og b ikke har felles primtallsfaktorer. Dette følger av *aritmetikkens fundamentalteorem* 3.7.7, som vi skal komme til senere i kapitlet. Dette uttrykket kalles gjerne *den kanoniske formen* til det rasjonale tallet. Funksjonen

$$\begin{array}{ccc} \mathbb{Z} & \longrightarrow & \mathbb{Q} \\ n & \mapsto & \dfrac{n}{1} \end{array}$$

er injektiv (jf. Definisjon 2.3.1) og gjør at vi kan identifisere $\mathbb{Z}$ med en delmengde av $\mathbb{Q}$, nemlig verdimengden til denne funksjonen. Gjennom denne identifiseringen sier vi at $\mathbb{Z}$ er en delmengde av $\mathbb{Q}$. Dessuten pleier vi å skrive n istedenfor $\frac{n}{1}$ i $\mathbb{Q}$.

Aritmetikk på $\mathbb{Q}$. Operasjonene addisjon og multiplikasjon defineres på følgende måte:

Definisjon 3.4.1: Addisjon og multiplikasjon på $\mathbb{Q}$

Vi definerer to operasjoner $+, \cdot : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$, kalt henholdsvis *addisjon* og *multiplikasjon*, ved:

- $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$.
- $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$.

Igjen må vi vise at disse operasjonene er veldefinert, det vil si uavhengige av valg av representanter for ekvivalensklassene:

Setning 3.4.2

Hvis $\frac{a'}{b'} = \frac{a}{b}$ og $\frac{c'}{d'} = \frac{c}{d}$, er $\frac{a'd'+b'c'}{b'd'} = \frac{ad+bc}{bd}$ og $\frac{a'c'}{b'd'} = \frac{ac}{bd}$.

[7] *Enumerator* på engelsk.

[8] *Denominator* på engelsk.

Bevis
Overlates til Oppgave 3.9. □

Vi kaller $\frac{a}{b} + \frac{c}{d}$ og $\frac{a}{b} \cdot \frac{c}{d}$ for henholdsvis *summen* og *produktet* av $\frac{a}{b}$ og $\frac{c}{d}$.

Merk også at dersom $\frac{a}{b} = \frac{a}{1}$ og $\frac{c}{d} = \frac{c}{1}$ er heltall, da er operasjonene de samme som de vi har definert på $\mathbb{Z}$ tidligere, siden $\frac{a}{1} + \frac{c}{1} = \frac{a+c}{1}$ og $\frac{a}{1} \cdot \frac{c}{1} = \frac{ac}{1}$.

Man kan vise at operasjonene er kommutative, assosiative og følger den distributive lov. Operasjonen addisjon har et nøytralt element, nemlig $0 = \frac{0}{1} = \frac{0}{n}$ for vilkårlig $n \in \mathbb{Z} \setminus \{0\}$, som altså oppfyller

$$0 + \frac{a}{b} = \frac{0}{1} + \frac{a}{b} = \frac{a}{b} \quad \text{for alle} \quad \frac{a}{b} \in \mathbb{Q},$$

og ethvert element $\frac{a}{b} \in \mathbb{Q}$ har $\frac{-a}{b}$ som additiv invers. Videre har operasjonen $\cdot$ et identitetselement, nemlig $1 = \frac{1}{1} = \frac{n}{n}$ for vilkårlig $n \in \mathbb{Z} \setminus \{0\}$, som altså oppfyller

$$1 \cdot \frac{a}{b} = \frac{1}{1} \cdot \frac{a}{b} = \frac{a}{b} \quad \text{for alle} \quad \frac{a}{b} \in \mathbb{Q}.$$

Og i motsetning til $\mathbb{Z}$, så har alle elementer i $\mathbb{Q} \setminus \{0\}$ nå en multiplikativ invers: inversen til $\frac{a}{b}$ er $\frac{b}{a}$. Denne betegnes gjerne også med $\left(\frac{a}{b}\right)^{-1}$. Vi oppsummerer de aritmetiske egenskapene til $\mathbb{Q}$ og vi overlater deltaljene med å sjekke disse til Oppgave 3.10. Merk at egenskapene er de samme som for $\mathbb{Z}$, bortsett fra at $(M4)$ har kommet til.

Setning 3.4.3: Egenskaper til addisjon og multiplikasjon på $\mathbb{Q}$

Operasjonene $+ : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ og $\cdot : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ på $\mathbb{Q}$ oppfyller egenskapene:

- (A1) Addisjon er kommutativ: $a + b = b + a$ for alle $a, b \in \mathbb{Q}$;
- (A2) Addisjon er assosiativ: $(a+b)+c = a+(b+c)$ for alle $a, b, c \in \mathbb{Q}$;
- (A3) $\mathbb{Q}$ inneholder et element 0 (kalt *nøytralt element*) som oppfyller $0 + a = a + 0 = a$ for alle $a \in \mathbb{Q}$;
- (A4) Til ethvert element $a \in \mathbb{Q}$ finnes et element $-a \in \mathbb{Q}$ (kalt *det additive inverselementet til* a) slik at $a + (-a) = (-a) + a = 0$;
- (M1) Multiplikasjon er kommutativ: $a \cdot b = b \cdot a$ for alle $a, b \in \mathbb{Q}$;
- (M2) Multiplikasjon er assosiativ: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for alle $a, b, c \in \mathbb{Q}$;
- (M3) $\mathbb{Q}$ inneholder et element $1 \neq 0$ (kalt *identitetselement*) som oppfyller $1 \cdot a = a \cdot 1 = a$ for alle $a \in \mathbb{Q}$;
- (M4) Til ethvert element $a \in \mathbb{Q}$ slik at $a \neq 0$ finnes et element $a^{-1} \in \mathbb{Q}$ (kalt *det multiplikative inverselementet til* a) slik at $a \cdot a^{-1} = a^{-1} \cdot a = 1$;
- (D) Den *distributive loven* $a \cdot (b + c) = a \cdot b + a \cdot c$ holder for alle $a, b, c \in \mathbb{Q}$.

MERKNAD 3.4.4. Vi har definert de to operasjonene $+$ (addisjon) og $\cdot$ (multiplikasjon) på $\mathbb{Q}$. Hva med de to operasjonene $-$ (subtraksjon) og $\div$ (divisjon), som vi kjenner fra skolen? Disse kan defineres ved hjelp av de første to, ved

$$\begin{aligned} a - b &= a + (-b) \qquad \text{og} \\ a \div b &= a \cdot b^{-1}, \quad \text{for} \quad b \neq 0. \end{aligned}$$

Vi kaller $a - b$ en *differanse* og $a \div b$ en *kvotient*[9].

3.5. Ordensrelasjonen på $\mathbb{Z}$ og $\mathbb{Q}$

Vi har hittil snakket om de aritmetiske egenskapene til $\mathbb{Z}$ og $\mathbb{Q}$, nemlig egenskapene til operasjonene $+$ og $\cdot$. Det finnes imidlertid en veldig viktig tilleggsstruktur på $\mathbb{Z}$ og $\mathbb{Q}$, som går ut på at vi kan "sammenligne elementer" ved å si hvilket av to gitte forskjellige elementer i mengdene som er størst.

La oss ta en titt på figuren på side 97 som "visualiserer" $\mathbb{Z}$. Denne gir en ordning av heltallene ved å gå "nedover og til høyre", det vil si at vi kan definere et heltall m til å være større enn n dersom m ligger til høyre for eller under n.

Noe mer formelt kan vi gå frem på følgende ekvivalente måte: Vi så i §3.3 på side 98 at vi kan skrive

$$\mathbb{Z} = \mathbb{Z}^- \cup \{0\} \cup \mathbb{Z}^+,$$

hvor

$$\mathbb{Z}^+ = \mathbb{N} \setminus \{0\} = \{1, 2, 3, 4, 5, \ldots\},$$

som vi har kalt *mengden av positive heltall*, og

$$\mathbb{Z}^- = \{-1, -2, -3, -4, -5, \ldots\},$$

som vi har kalt *mengden av negative heltall*.

Vi bruker notasjonen

$$n > 0 \;\text{ eller }\; 0 < n \;\text{ hvis } n \text{ er positiv, det vil si } n \in \mathbb{Z}^+$$

og

$$n < 0 \;\text{ eller }\; 0 > n \;\text{ hvis } n \text{ er negativ, det vil si } n \in \mathbb{Z}^-.$$

For $m, n \in \mathbb{Z}$ definerer vi:

$$m > n \Longleftrightarrow n < m \Longleftrightarrow m - n > 0 \tag{60}$$

og sier da at m *er større enn* n eller, ekvivalent, at n *er mindre enn* m. Det er lett å se at dette stemmer med at m ligger til høyre for eller under n i figuren på side 97.

For $p, q \in \mathbb{Q}$ går vi frem på følgende måte: Skriv først $p = \frac{a}{b}$ og $q = \frac{c}{d}$ med $b, d \in \mathbb{Z}^+$ og $a, b \in \mathbb{Z}$; dette kan vi alltid gjøre ved å eventuelt gange

[9] *Quotient* på engelsk.

teller og nevner med -1 dersom sistnevnte er negativ. (Ved (59) har vi nemlig $(a,b) \sim (-a,-b)$, med andre ord $\frac{a}{b} = \frac{-a}{-b}$.) Siden

$$p = \frac{a}{b} = \frac{ad}{bd} \quad \text{og} \quad q = \frac{c}{d} = \frac{cb}{db},$$

kan vi nå sammenligne p og q ved å se på tellerne og definere:

$$p > q \iff q < p \iff ad - bc > 0 \tag{61}$$

og sier da at p *er større enn* q, eller, ekvivalent, at q *er mindre enn* p.

Vi må som vanlig vise at dette er veldefinert, det vil si uavhengig av valg av representanter for ekvivalensklassene:

Setning 3.5.1

Hvis $\frac{a'}{b'} = \frac{a}{b}$ og $\frac{c'}{d'} = \frac{c}{d}$, med $b, d, b', d' \in \mathbb{Z}^+$, er

$$ad - bc > 0 \iff a'd' - b'c' > 0.$$

Bevis

At $\frac{a'}{b'} = \frac{a}{b}$ og $\frac{c'}{d'} = \frac{c}{d}$ betyr at

$$a'b = ab' \quad \text{og} \quad c'd = cd'.$$

Ved hjelp av dette kan vi regne ut

$$(b'd')(ad+bc) = \underbrace{(ab')}_{\substack{\| \\ (ba')}}dd' + \underbrace{(cd')}_{\substack{\| \\ (c'd)}}bb' = (ba')dd' + (c'd)bb' = (bd)(a'd' + b'c').$$

Antagelsen om at $b, d, b', d' \in \mathbb{Z}^+$ gir at $(bd), (b'd') \in \mathbb{Z}^+$ ved (56), og da er $ad - bc \in \mathbb{Z}^+$ hvis og bare hvis $a'd' - b'c' \in \mathbb{Z}^+$ igjen ved (56). □

I både $\mathbb{Z}$ og $\mathbb{Q}$ sier vi dessuten at p *er større eller lik* q, ekvivalent at q *er mindre eller lik* p, dersom $p > q$ eller $p = q$, og betegner dette med $p \geq q$ eller $p \leq q$. Med andre ord er "$p \geq q$" negasjonen til "$p < q$".

Setning 3.5.2

Følgende gjelder:

(i) For alle $a, b \in \mathbb{Q}$ gjelder nøyaktig én av følgende:
$$a < b, \quad a = b, \quad a > b.$$

(ii) For alle $a, b, c \in \mathbb{Q}$, hvis $a < b$ og $b < c$, da er $a < c$.

Bevis

Overlates til Oppgave 3.11. □

På samme måte som for $\mathbb{Z}$ sier vi at $x \in \mathbb{Q}$ er *positiv* om $x > 0$ og *negativ* om $x < 0$. Det følger av egenskap (i) i siste setning at et element $x \in \mathbb{Q} \setminus \{0\}$ enten er positivt eller negativt, men ikke begge deler. Vi betegner mengden av alle positive rasjonale tall med $\mathbb{Q}^+$ og mengden av alle negative rasjonale tall med $\mathbb{Q}^-$, slik at

$$\mathbb{Q} = \mathbb{Q}^+ \cup \{0\} \cup \mathbb{Q}^-.$$

Det er også vanlig å bruke notasjonene

$$\mathbb{Q}^{\geq 0} = \mathbb{Q}^+ \cup \{0\}$$

for mengden av de ikke-negative rasjonale tallene og

$$\mathbb{Q}^{\leq 0} = \mathbb{Q}^- \cup \{0\}$$

for mengden av de ikke-positive rasjonale tallene.

Det er ikke vanskelig å se at delmengden

$$R = \{(p, q) \in \mathbb{Q} \times \mathbb{Q} \mid p < q\}$$

er en relasjon på $\mathbb{Q}$ (og likeledes på $\mathbb{Z}$), jf. Definisjon 2.4.1. En relasjon på en hvilken som helst mengde som oppfyller egenskapene i Setning 3.5.2 kalles en *ordensrelasjon*. Vi skal komme tilbake til disse i neste seksjon (jf. Definisjon 3.6.9).

Relasjonen $<$ på $\mathbb{Q}$ (og $\mathbb{Z}$) og operasjonene $+$ og $\cdot$ oppfyller følgende viktige egenskaper, som vi er vant til:

Setning 3.5.3

Følgende gjelder:

(i) hvis $x, y, z \in \mathbb{Q}$ og $y < z$, da er $x + y < x + z$;
(ii) hvis $x, y \in \mathbb{Q}$ og $x > 0$, $y > 0$, da er $xy > 0$.

Bevis
Overlates til Oppgave 3.13. □

Vi tar også med følgende resultat, som virker helt opplagt fordi vi er så vant til at det er sant, men som likevel har et ikke helt banalt bevis:

Setning 3.5.4

La $n \in \mathbb{Z}$. Det finnes ingen $k \in \mathbb{Z}$ slik at $n < k < n + 1$.

Bevis
Vi argumenterer ved selvmotsigelse. Anta at det finnes et heltall k slik at $n < k < n+1$. Ved Setning 3.5.3(i) kan vi addere $-n$ på begge sider av en ulikhet og bevare den. Dette medfører at $0 = n-n < k-n < n+1-n = 1$, altså at det finnes et heltall $l = k - n$ mellom 0 og 1. Siden $l > 0$, er $l \in \mathbb{Z}^+ \subset \mathbb{N}$, slik at l må være et naturlig tall.

Betrakt delmengden $M = \{n \in \mathbb{N} \mid n \geq 1\} \cup \{0\}$. Ved å argumentere som flere steder ovenfor ved *induksjonsaksiomet* (N5) kan vi konkludere at $M = \mathbb{N}$. (Detaljene overlates til Oppgave 3.14.) Dette viser at det ikke finnes noen naturlige tall mellom 0 og 1, som motsier eksistensen av l. □

3.6. Ringer og kropper

Vi har sett at operasjonene $+$ og $\cdot$ på $\mathbb{Z}$ og $\mathbb{Q}$ oppfyller de aritmetiske egenskapene i Setningene 3.3.5 og 3.4.3. Når vi regner med $\mathbb{Z}$ og $\mathbb{Q}$ er vi imidlertid vant til å bruke langt flere regneregler enn de som er listet opp i disse setningene. For eksempel er vi vant til å bruke kansellering og forkortning når vi løser ligninger. Har vi uttrykket

$$x + 3 = y + 3,$$

da er vi vant til å kunne konkludere (ved "kansellering" av 3 på begge sider av ligningen) at $x = y$. Har vi uttrykket

$$4x = 4y,$$

da er vi også vant til å kunne konkludere (ved "kansellering" eller "forkortning" av 4 på begge sider av ligningen) at $x = y$.

Det samme gjelder ordensrelasjonen $<$ på $\mathbb{Z}$ og $\mathbb{Q}$, som oppfyller egenskapene i Setningene 3.5.2 og 3.5.3. Vi er for eksempel vant til å kunne dedusere ut fra

$$3x < 3y$$

at $x < y$ (ved "forkortning" av 3 på begge sider av ulikheten). Og vi er vant til å kunne multiplisere begge sider av en ulikhet med et positivt tall og "bevare ulikheten", eller multiplisere begge sider av en ulikhet med et negativt tall og "snu ulikheten".

Hvor kommer disse reglene (og andre vi er vant til) fra?

Svaret er heldigvis at alle reglene er direkte konsekvenser av egenskapene listet i Setningene 3.3.5 og 3.4.3 (når det gjelder $+$ og $\cdot$), og i Setningene 3.5.2 og 3.5.3 (når i tillegg $<$ er involvert). De kan utledes relativt enkelt. Siden reglene kun bygger på de grunnleggende egenskapene i disse setningene, og det er mange flere mengder enn $\mathbb{Z}$ og $\mathbb{Q}$ som oppfyller disse egenskapene i matematikk (som for eksempel de reelle tallene $\mathbb{R}$, som vi skal komme til i Kapittel 7, og de komplekse tallene $\mathbb{C}$, som vi skal behandle i Kapittel 8) er det imidlertid en fordel på dette tidspunktet å sile ut nettopp de mengdene som oppfyller egenskapene i Setningene 3.3.5 og 3.4.3, og deretter i Setningene 3.5.2 og 3.5.3, og så utlede alle andre regneregler som konsekvenser av disse. Det er det vi skal gjøre i denne seksjonen.

Kroppsaksiomene. Mengder med to operasjoner som oppfyller egenskapene i Setning 3.4.3 er såpass viktige at de har fått et eget navn, som unektelig høres litt morsomt ut på norsk: Det norske navnet *kropp*[10] kommer fra tysk "Körper", som var det opprinnelige begrepet introdusert av den tyske matematikeren Richard Dedekind (1831-1916) i 1871[11].

[10] *Field* på engelsk.

[11] Det engelske navnet "field" ble gitt i 1893 av den amerikanske matematikeren E. H. Moore (1862-1932).

Definisjon 3.6.1: Kropp

En *kropp* er en mengde K sammen med to funksjoner

$$+ : K \times K \longrightarrow K$$

(kalt *addisjon*) og

$$\cdot : K \times K \longrightarrow K$$

(kalt *multiplikasjon*) som oppfyller følgende aksiomer:

(A1) Addisjon er kommutativ: $a + b = b + a$ for alle $a, b \in K$;

(A2) Addisjon er assosiativ: $(a + b) + c = a + (b + c)$ for alle $a, b, c \in K$;

(A3) K inneholder et element 0 (kalt *nøytralt element*) som oppfyller $0 + a = a + 0 = a$ for alle $a \in K$;

(A4) Til ethvert element $a \in K$ finnes et element $-a \in K$ (kalt *det additive inverselementet til* a) slik at $a+(-a) = (-a)+a = 0$;

(M1) Multiplikasjon er kommutativ: $a \cdot b = b \cdot a$ for alle $a, b \in K$;

(M2) Multiplikasjon er assosiativ: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for alle $a, b, c \in K$;

(M3) K inneholder et element $1 \neq 0$ (kalt *identitetselement*) som oppfyller $1 \cdot a = a \cdot 1 = a$ for alle $a \in K$;

(M4) Til ethvert element $a \in K$ slik at $a \neq 0$ finnes et element $a^{-1} \in K$ (kalt *det multiplikative inverselementet til* a) slik at $a \cdot a^{-1} = a^{-1} \cdot a = 1$;

(D) Den *distributive loven* $a \cdot (b + c) = a \cdot b + a \cdot c$ holder for alle $a, b, c \in K$.

Vi kaller $a + b$ og $a \cdot b$ for henholdsvis *summen* og *produktet* av a og b.

Som vi har sett, oppfyller mengden av rasjonale tall $\mathbb{Q}$ med operasjonene $+$ og $\cdot$ alle disse aksiomene, slik at $\mathbb{Q}$ er en kropp. (Vi vil også se senere at også mengden av de reelle tallene $\mathbb{R}$ er en kropp, slik som mengden av alle komplekse tall $\mathbb{C}$.) Mengden av heltall $\mathbb{Z}$ oppfyller alle disse bortsett fra eksistensen av det multiplikative inverselement (M4), som vi diskuterte ovenfor. Slike mengder kalles *kommutative ringer*:

Definisjon 3.6.2: Kommutativ ring

En *kommutativ ring*[a] er en mengde R sammen med to funksjoner

$$+ : R \times R \longrightarrow R,$$

$$\cdot : R \times R \longrightarrow R$$

som oppfyller alle aksiomene ovenfor bortsett fra muligens (M4).

[a] *Commutative ring* på engelsk.

I en kommutativ ring vil det altså ikke nødvendigvis for alle $a \neq 0$ i ringen finnes et element a^{-1} slik at $a \cdot a^{-1} = a^{-1} \cdot a = 1$. En slik a^{-1}, når den finnes, kalles for et *multiplikativt inverselement* til a. For eksempel har vi sett tidligere, i (57), at det kun er elementene 1 og -1 i $\mathbb{Z}$ som har multiplikative inverser, som er 1 og -1 selv, henholdsvis.

Det finnes også eksempler på mengder som oppfyller enda færre aksiomer. For eksempel kalles en mengde som oppfyller alle aksiomene bortsett fra muligens (M4) og (M1) kun for en *ring*, og dersom (M1) *ikke* er oppfylt, spesifiserer vi gjerne at *ringen er ikke-kommutativ*[12]. De kanskje mest kjente eksemplene på ikke-kommutative ringer er *matriseringer*, som dere vil møte i kurset *MAT121-Lineær algebra.* Vi gir en liten forsmak i §4.4.

Strukturer som ringer og kropper er eksempler på *algebraiske strukturer*. Mer om slike strukturer lærer dere om i kurset *MAT220–Algebra.* Selv om vi videre skal konsentrere oss om de rasjonale tallene $\mathbb{Q}$ og hvordan man bygger de reelle tallene $\mathbb{R}$ (og deretter de komplekse tallene $\mathbb{C}$) ut i fra dem, skal vi i Kapittel 4 se på noen andre eksempler på ringer og kropper.

Før det skal vi utlede noen egenskaper for kropper og ringer, som da spesielt gjelder for $\mathbb{Q}$ og $\mathbb{Z}$.

I enhver kropp skriver vi gjerne også $\frac{1}{a}$ for a^{-1} og vi forenkler uttrykk på følgende måte (noe vi er vant til å gjøre for både $\mathbb{Q}$ og $\mathbb{R}$ fra skolen):

$$\begin{array}{rcl}
a + (-b) & = & a - b \\
a \cdot b^{-1} = a \cdot \dfrac{1}{b} & = & \dfrac{a}{b} \\
(a + b) + c = a + (b + c) & = & a + b + c \\
a \cdot b & = & ab \\
(a \cdot b) \cdot c = a \cdot (b \cdot c) & = & abc \\
a + a & = & 2a \\
a + a + a & = & 3a \\
\cdots & \cdots & \cdots \\
\underbrace{a + \cdots + a}_{n \text{ ganger}} & = & na \\
a \cdot a & = & a^2 \\
a \cdot a \cdot a & = & a^3 \\
\cdots & \cdots & \cdots \\
\underbrace{a \cdots a}_{n \text{ ganger}} & = & a^n \\
\underbrace{a^{-1} \cdots a^{-1}}_{n \text{ ganger}} & = & a^{-n}
\end{array}$$

Vi kaller $a - b$ en *differanse* og $\frac{a}{b}$ en *kvotient*[13].

[12] *Non-commutative* på engelsk.

[13] *Quotient* på engelsk.

Vi sier at a^n er en *potens*[14], nærmere bestemt *n-te potens*, av a, at a er *grunntallet*[15] og at n er *eksponenten*. Per definisjon setter vi dessuten $a^1 = a$ og

$$a^0 = 1, \quad \text{for} \;\; a \neq 0. \tag{62}$$

Som konsekvens får vi regelen

$$a^n \cdot a^m = a^{n+m} \;\; \text{for alle} \;\; n, m \in \mathbb{Z}.$$

Merk ellers at uttrykkene $2a, 3a, \ldots, na$ ovenfor også gir mening som produkt av to elementer i kroppen ved å betrakte 2 som $1+1$, 3 som $1+1+1$ og mer generelt n som $\underbrace{1+\cdots+1}_{n \text{ ganger}}$ og bruke den distributive loven.

Grunnleggende egenskaper til ringer og kropper. La R være en kommutativ ring. (Dette inkluderer muligheten at R er en kropp.)

Vi starter med en elementær, men viktig, observasjon. Siden operasjonene $+$ og $\cdot$ er definert som funksjoner $R \times R \to R$, har vi at dersom $x = y$ i R og $z \in R$, da er $(x, z) = (y, z) \in R \times R$, slik at funksjonsverdiene i (x, z) og (y, z) er de samme. Funksjonsverdiene til $+$ er $x+z$ og $x+y$, og funksjonsverdiene til $\cdot$ er xz og xy. Vi kan altså konkludere at

$$x = y \;\text{ i }\; R \;\text{ og }\; z \in R \Longrightarrow x + z = y + z \;\text{ og }\; xz = yz.$$

Dette er den velkjente regelen om at vi kan addere eller multiplisere med samme element på begge sider av en likhet og bevare likheten. Vi vil bruke dette videre uten ytterligere bemerkninger.

En hel rekke egenskaper som vi er vant til gjelder for $\mathbb{Z}$, $\mathbb{Q}$ (og $\mathbb{R}$) fra skolen er konsekvenser av kroppsaksiomene ovenfor:

Setning 3.6.3: Aritmetiske egenskaper i kropper

La K være en kropp. Følgende egenskaper gjelder for $a, b, c \in K$

(i) Hvis $a + b = a + c$, er $b = c$.
(ii) Hvis $a + b = a$, er $b = 0$.
(iii) Hvis $a + b = 0$, er $b = -a$.
(iv) $-(-a) = a$.
(v) Hvis $ab = ac$ og $a \neq 0$, er $b = c$.
(vi) Hvis $ab = a$ og $a \neq 0$, er $b = 1$.
(vii) Hvis $ab = 1$ og $a \neq 0$, er $b = a^{-1}$.
(viii) Hvis $a \neq 0$, er $(a^{-1})^{-1} = a$.
(ix) $0a = 0$.
(x) Hvis $a \neq 0$ og $b \neq 0$, er $ab \neq 0$.
(xi) $(-a)b = -(ab) = a(-b)$.
(xii) $(-a)(-b) = ab$.

I en kommutativ ring gjelder egenskapene (i)-(iv), (ix), (xi) og (xii).

[14] *Power* på engelsk.
[15] *Base* på engelsk.

Merknad 3.6.4. Punktene (ii), (iii), (vi) og (vii) hevder entydigheten av elementene som oppfyller (A3), (A4), (M3) og (M4). Entydigheten av det multiplikative inverselementet 1 gjelder også i en kommutativ ring, hvor det kan vises direkte (jf. Oppgave 3.15).

Merknad 3.6.5. Reglene (i) og (v) går også under navnet "kanselleringslovene". Regel (v) omtales også som "forkortningsloven".

Bevis for Setning 3.6.3
Vi viser først (i). Vi starter med hypotesen $a + b = a + c$ og adderer $-a$ på begge sider:

$$
\begin{aligned}
(-a) + (a + b) = (-a) + (a + c) \quad &\overset{(A2)}{\Longrightarrow} \quad (-a + a) + b = (-a + a) + c \\
&\overset{(A4)}{\Longrightarrow} \quad 0 + b = 0 + c \overset{(A3)}{\Longrightarrow} b = c,
\end{aligned}
$$

hvor vi har markert hvilke aksiomer vi har brukt. Konklusjonen gjelder altså også i en kommutativ ring.

Egenskapene (ii)-(iv) er konsekvenser av (i); vi overlater detaljene til Oppgave 3.16.

Egenskapen (v) vises på tilsvarende måte (i en kropp) ved å starte med hypotesen $ab = ac$ og multiplisere med a^{-1} på begge sider. Her trenger vi at $a \neq 0$ og at K er en kropp for at a^{-1} skal eksistere. Detaljene overlates igjen til Oppgave 3.16.

Egenskapene (vi)-(viii) er konsekvenser av (v); vi overlater detaljene til Oppgave 3.16.

For å vise (ix), skriver vi

$$(0a + 0a) \overset{(D)}{=} (0 + 0)a \overset{(A3)}{=} 0a.$$

Fra dette og egenskap (ii) konkluderer vi at $0a = 0$.

Vi overlater (x)-(xii) til Oppgave 3.16. □

Merknad 3.6.6. Siden $\mathbb{Z} \subset \mathbb{Q}$ med de samme operasjonene, så gjelder også alle reglene i Setning 3.6.3 som ikke involverer inverser (det vil si alle utenom (vii) og (viii)) også for $\mathbb{Z}$.

Merknad 3.6.7. Kommutative ringer oppfyller generelt ikke alle egenskapene i Setning 3.6.3. Vi vil i §4.2 og §4.3 se eksempler på kommutative ringer der egenskap (x) ikke er oppfylt, og heller ikke forkortningsloven (v).

MERKNAD 3.6.8. Prosedyren med å "utvide" $\mathbb{Z}$ til $\mathbb{Q}$ ved å "legge til" multiplikative inverselementer kan generaliseres til visse typer ringer, og den resulterende kroppen kalles *kvotientkroppen*[16] til ringen (jf. Oppgave 4.15).

Ordnede kropper og deres egenskaper. En relasjon på en mengde som oppfyller egenskapene i Setning 3.5.2 har fått et eget navn:

Definisjon 3.6.9: Ordensrelasjon

En *ordensrelasjon*[a] på en mengde X er en relasjon R som oppfyller følgende betingelser:

(O1) Trikotomi[b]/totalitet: for alle $x, y \in X$ gjelder nøyaktig én av følgende:
$$(x, y) \in R, \quad x = y, \quad (y, x) \in R.$$

(O2) Transitivitet[c]: for alle $x, y, z \in X$, has at hvis $(x, y) \in R$ og $(y, z) \in R$, da er $(x, z) \in R$.

Vi sier at en mengde X sammen med en ordensrelasjon på X er en *ordnet mengde*.

[a] *Order relation* på engelsk.
[b] Dette ordet betyr "oppdeling i tre deler"; på engelsk heter det *trichotomy*.
[c] *Transitivity* på engelsk.

Vi bruker notasjonen
$$x < y \text{ eller } y > x \text{ for } (x, y) \in R$$
og leser "$<$" som "mindre enn" og "$>$" som "større enn". Da kan betingelsene (O1) og (O2) ovenfor uttrykkes som:

(O1) *Trikotomi/totalitet*: for alle $x, y \in X$ gjelder nøyaktig én av følgende:
$$x < y, \quad x = y, \quad x > y.$$

(O2) *Transitivitet*: for alle $x, y, z \in X$, hvis $x < y$ og $y < z$, da er $x < z$.

Ved Setning 3.5.2 er relasjonene definert ovenfor i (60) på $\mathbb{Z}$ og (61) på $\mathbb{Q}$ ordensrelasjoner.

På en ordnet mengde X innfører vi også notasjonene:
$$x \leq y \iff y \geq x \iff x < y \text{ eller } x = y$$
og leser "$\leq$" som "mindre enn eller lik" og "$\geq$" som "større enn eller lik". Med andre ord er "$x \leq y$" negasjonen av påstanden "$x > y$".

La oss se på et par ordensrelasjoner på mengder forskjellig fra $\mathbb{Z}$ og $\mathbb{Q}$.

[16] *Quotient field* på engelsk.

EKSEMPEL 3.6.10. Betrakt en mengde $X = \{a, b, c\}$ med tre elementer og følgende relasjoner på X:

$$\begin{aligned} R_1 &= \{(a,b),(b,c)\}, \\ R_2 &= \{(a,b),(b,c),(c,a)\}, \\ R_3 &= \{(a,b),(b,c),(a,c)\}. \end{aligned}$$

Relasjonen R_1 oppfyller ikke trikotomibetingelsen i Definisjon (3.6.9), siden $a \neq c$, $(a,c) \notin R_1$ og $(c,a) \notin R_1$.

Relasjonen R_2 oppfyller trikotomibetingelsen i Definisjon (3.6.9), men ikke transitivitetsbetingelsen, for siden $(a,b) \in R_2$ og $(b,c) \in R_2$, ville vi måttet ha at $(a,c) \in R_2$ og *ikke* $(c,a) \in R_2$.

Relasjonen R_3 oppfyller både trikotomibetingelsen og transitivitetsbetingelsen i Definisjon (3.6.9) og er derfor en ordensrelasjon.

EKSEMPEL 3.6.11. Vi kan lage en ordensrelasjon på $\mathbb{Q}^2 = \mathbb{Q} \times \mathbb{Q}$ ved å sette $(a,b) < (c,d)$ hvis og bare hvis enten $a < c$, eller $a = c$ og $b < d$. Vi overlater til Oppgave 3.17 å verifisere at dette er en ordensrelasjon. Den kalles *leksikografisk orden*[17].

Denne ordensrelasjonen kan generaliseres til et kartesisk produkt av vilkårlig mange ordnede mengder: hvis $X_1, \ldots, X_n$ er ordnede mengder, definerer vi $(x_1, \ldots, x_n) < (y_1, \ldots, y_n)$ hvis og bare hvis enten $x_1 < y_1$ eller det finnes en $m \in \{2, \ldots, n\}$ slik at $x_i = y_i$ for alle $1 \leq i < m$ og $x_m < y_m$.

Som vi så i Setning 3.5.3, oppfyller ordensrelasjonen $<$ på $\mathbb{Q}$ (og $\mathbb{Z}$) visse egenskaper knyttet til operasjonene $+$ og $\cdot$. Dette danner grunnlaget for følgende definisjon som gjelder generelt for kropper:

Definisjon 3.6.12: Ordnet kropp

En *ordnet kropp* er en kropp K som også er en ordnet mengde, og som oppfyller:

(OK1) hvis $x, y, z \in K$ og $y < z$, da er $x + y < x + z$;
(OK2) hvis $x, y \in K$ og $x > 0$, $y > 0$, da er $xy > 0$.

Hvis $x \in K \setminus \{0\}$, da sier vi at x er *positiv* hvis $x > 0$ og *negativ* hvis $x < 0$. (Trikotonomi-/totalitetsegenskapen gjør at enhver $x \in K \setminus \{0\}$ enten er positiv eller negativ.)

Ved Setning 3.5.3 er $\mathbb{Q}$ en ordnet kropp. De resterende egenskapene til $\mathbb{Q}$ som vi er vant til gjelder generelt for ordnede kropper:

[17] *Lexicographical order* på engelsk.

Setning 3.6.13: Ordensegenskaper i ordnede kropper

La K være en ordnet kropp, og $x, y, z \in K$. Da gjelder følgende egenskaper:

(i) $x > 0$ hvis og bare hvis $-x < 0$.
(ii) Hvis $x > 0$ og $y < z$, da er $xy < xz$.
(iii) Hvis $x < 0$ og $y < z$, da er $xy > xz$.
(iv) Hvis $x \neq 0$, da er $x^2 > 0$. Spesielt er $1 > 0$.
(v) Hvis $0 < x < y$, da er $0 < \frac{1}{y} < \frac{1}{x}$.

Bevis

Vi viser kun (i) og (ii) og overlater resten til Oppgave 3.18.

(i) Hvis $x > 0$, er $-x + x > -x + 0$ ved aksiom (OK1) for ordnet kropp (Definisjon 3.6.12). Kroppsaksiom (A4) gir at $-x + x = 0$ og kroppsaksiom (A3) gir at $-x + 0 = -x$. Det følger at $0 > -x$, som er det vi skulle bevise. På helt tilsvarende måte viser man at $x > 0$, hvis $-x < 0$.

(ii) Fra antagelsen om at $y < z$, får vi $-y + y < -y + z$ ved aksiom (OK1) for ordnet kropp (Definisjon 3.6.12). Siden $-y+y = 0$ ved kroppsaksiom (A4), får vi at $0 < -y + z$. Siden også $x > 0$ per antagelse, gir aksiom (OK2) for ordnet kropp (Definisjon 3.6.12) at $x(-y+z) > 0$. Ved aksiom (OK1) i Definisjon 3.6.12 nok en gang får vi at

$$xy + x(-y + z) > xy + 0. \tag{63}$$

Siden $xy + 0 = xy$ ved kroppsaksiom (A3), og

$$\begin{aligned} xy + x(-y + z) \overset{(D)}{=} x\,(y + (-y + z)) &\overset{(A2)}{=} x\,((y + (-y)) + z)) \\ &\overset{(A4)}{=} x\,(0 + z) \overset{(A3)}{=} xz \end{aligned}$$

(hvor vi har markert hvilke kroppsaksiomer vi har brukt), er (63) ekvivalent med $xz > xy$, som er det vi skulle vise. $\square$

Merknad 3.6.14. Siden $\mathbb{Z} \subset \mathbb{Q}$ med den samme ordensrelasjonen, så gjelder reglene (i)-(iv) også for $\mathbb{Z}$.

Definisjon 3.6.15: Absoluttverdi

La K være en ordnet kropp og $a \in K$. *Absoluttverdien*[a] til a, betegnet med $|a|$, er definert som

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

[a] *Absolute value* på engelsk.

Setning 3.6.16: Egenskaper til absoluttverdi

Absoluttverdien i en ordnet kropp oppfyller følgende:

(i) $|ab| = |a||b|$.
(ii) $|a + b| \leq |a| + |b|$ (Trekantulikheten).

Bevis
Overlates til Oppgave 3.19. □

Trekantulikheten kommer vi til å bruke ved mange anledninger senere.

Merknad 3.6.17. Siden $|a - b| = |a + (-b)|$ og $|-b| = |b|$, sier trekantulikheten også at $|a - b| \leq |a| + |b|$.

Hvis vi tenker oss alle elementene i en ordnet kropp K som "ordnet i en tallinje" (som er det vi gjør for $\mathbb{Q}$), så kan vi tolke $|a|$ som avstanden mellom a og det nøytrale elementet 0. Mer generelt kan vi definere *avstanden*[18] mellom to elementer a og b som $|a - b|$.

3.7. Primtall og divisjon

Siden primtall har en helt særegen rolle i matematikken, tar vi med en seksjon om dette temaet. Vi starter med å definere hva vi mener med *deling* eller *divisjon*:

Definisjon 3.7.1: Delelighet

La $m \in \mathbb{N}$ og $n \in \mathbb{Z}^+$. Vi sier at n *deler*[a] m, ekvivalent at m *er delelig på*[b] n, dersom det finnes en $k \in \mathbb{N}$ slik at $m = n \cdot k$. Vi betegner dette med $n|m$.

I dette tilfellet sier vi også at n er en *faktor* eller *divisor* i m og at m er et *multippel* av n.

[a] *Divides* på engelsk.
[b] *Divisible by* på engelsk.

Vi legger merke til at alle $n \in \mathbb{Z}^+$ deler 0, per definisjon.
Vi lister opp noen elementære egenskaper av delelighet:

Setning 3.7.2: Delelighetsegenskaper

La $a, b, c \in \mathbb{N}$, med $a \neq 0$. Da gjelder:

(i) Hvis $a|b$ og $a|c$, da vil $a|(b + c)$ og $a|\ |b - c|$.
(ii) Hvis $a|b$, da vil $a|bc$.
(iii) Hvis $a|b$ og $b|c$ (her må også $b \neq 0$), da vil $a|c$.

[18] *Distance* på engelsk.

Bevis

Overlates til Oppgave 3.20(b). □

Vi tar også med den velkjente *divisjonsalgoritmen*:

Setning 3.7.3: Divisjonsalgoritmen

Hvis $a \in \mathbb{Z}$ og $d \in \mathbb{Z}^+$, da finnes entydige $q, r \in \mathbb{Z}$ slik at $0 \leq r < d$ og $a = dq + r$.

Definisjon 3.7.4: Kvotient og rest

Vi kaller q for *kvotienten*[a] i divisjonen og r for *resten*[b] i divisjonen.

[a] *Quotient* på engelsk.
[b] *Remainder* på engelsk.

Eksempel 3.7.5. La $a = 29$ og $d = 3$. Da er $29 = 3 \cdot 9 + 2$, slik at $q = 9$ og $r = 2$. Vi skriver dette også som $\frac{29}{3} = 9 + \frac{2}{3}$.

Beviset for divisjonsalgoritmen bygger på *induksjonsaksiomet* (N5) for de naturlige tallene, eller rettere sagt, en konsekvens av dette, som heter *velordningsprinsippet*[19] for de naturlige tallene. Vi skal komme nærmere inn på dette prinsippet i §5.1 og vil vente med beviset for divisjonsalgoritmen til da.

Noen heltall skiller seg spesielt ut:

Definisjon 3.7.6: Primtall

Et *primtall*[a] er et heltall større enn 1 som kun er delelig på seg selv og 1.

[a] *Prime number* på engelsk.

Altså er per definisjon 1 *ikke* et primtall, mens 2 er det. De laveste primtallene er $2, 3, 5, 7, 11, 13, 17, 19, \ldots$ Tallet 2 er det eneste primtallet som er et partall.

Primtallenes særegne, universelle egenskap gjør at (de mulige) romvesenene i science fiction romanen *Contact* fra 1985 av den amerikanske astronomen og kosmologen Carl Sagan (1934–1996), filmatisert av Robert Zemeckis med Jodie Foster og Matthew McConaughey i hovedrollene i 1997, bruker primtallsekvensen til å kommunisere med SETI[20]-forskerne; se filmscenen på

https://www.youtube.com/watch?v=GZDT-FsO9Uc.

[19] *Principle of well-ordering* på engelsk.
[20] Search for extraterrestrial intelligence.

Aritmetikkens fundamentalteorem ble først vist av Euklid i Bok VII av *Elementer* ca. 300 f.Kr. Det viser at primtallene er de fundamentale byggestenene i alle heltall:

Teorem 3.7.7: Aritmetikkens fundamentalteorem

Ethvert heltall $n \geq 2$ kan skrives som et produkt av primtall, på en *entydig* måte (bortsett fra rekkefølgen av faktorene).

Definisjon 3.7.8: Primfaktorer

Primtallene som dukker opp i et talls entydige primtallsfaktorisering kalles *primfaktorene* til tallet.

EKSEMPEL 3.7.9. Vi har $2835 = 3^4 \cdot 5 \cdot 7$.

Vi har igjen allerede tilgjengelig det verktøyet som skal til for å bevise *aritmetikkens fundamentalteorem*, nemlig *induksjonsaksiomet* (N5) for de naturlige tallene. Likevel skal vi vente med beviset til §5.3, etter at vi har finpusset bruken av dette aksiomet noe.

Vi tar med følgende velkjente resultat, også bevist av Euklid.

Teorem 3.7.10: Uendelighet av primtall

Det finnes uendelig mange primtall.

BEVIS

Vi utfører et bevis ved selvmotsigelse (jf. §1.3) og antar derfor at det finnes endelig mange primtall $p_1, \ldots, p_n$. La

$$P = p_1 \cdots p_n + 1.$$

Ved *aritmetikkens fundamentalteorem* 3.7.7 har P en entydig primtallsfaktorisering (som består av kun P selv, om P er primtall). Siden vi har antatt at $p_1, \ldots, p_n$ er de eneste primtallene, er det kun disse som kan fremkomme i primtallsfaktoriseringen til P. Det finnes derfor minst ett av disse primtallene, si p_j, som fremkommer i primtallsfaktoriseringen, og dermed deler P. Samtidig deler p_j tallet $p_1 \cdots p_n$. Dermed vil p_j også dele $P - p_1 \cdots p_n = 1$ ved Setning 3.7.2(i), som er umulig, siden $p_j \geq 2$. Vi har dermed oppnådd en selvmotsigelse. □

Per dags dato er det største kjente primtallet

$$2^{82\,589\,933} - 1,$$

som ble "oppdaget" 21. desember 2018. Følg med på siden

https://www.mersenne.org/

Et spørsmål som har fascinert matematikere i flere hundre år er hvordan primtallene er fordelt, eller *distribuert*, på tallinjen. Man ble raskt klar over at fordelingen er ganske ujevn, men mot slutten av 1700-tallet oppdaget den franske matematikeren Adrien-Marie Legendre (1752–1833) og den tyske matematikeren og fysikeren Johann Carl Friedrich Gauss (1777–1855) uavhengig av hverandre en overraskende lovmessighet: definerer vi funksjonen $\pi : \mathbb{R}^+ \to \mathbb{N}$ ved $\pi(x) = \text{antall primtall } < x$, ser det ut til at $\pi(x)$ er tilnærmet lik $x/\ln x$ når x er stor. (Her er $\ln x$ den naturlige logaritmen til x, jf. Definisjon 9.5.13.) Med andre ord ser det ut til at

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\ln x} = 1.$$

(Se Definisjon 9.10.17 for definisjonen av $\lim\limits_{x\to\infty}$.) Verken Legendre eller Gauss kunne bevise dette resultatet. Det ble bevist først i 1896 av den franske matematikeren Jacques Salomon Hadamard (1865–1963) og den belgiske matematikeren Charles Jean de la Vallée Poussin (1866–1962) uavhengig av hverandre. Resultatet ble kjent som *Primtallsteoremet* og beviset bruker i stor grad hele teorien for *kompleks analyse* (grovt sagt funksjonsteori med komplekse tall, jf. Kapittel 8) som var utviklet på 1800–tallet, blant annet av den tyske matematikeren Bernhard Riemann (1826–1866), som selv hadde skrevet et arbeid om primtallenes fordeling i 1859, som vi nevnte i delseksjonen om *Riemannhypotesen* i §1.4.

Atle Selberg (Kilde: Wikipedia, public domain)

Mange trodde at det ikke var mulig å bevise primtallsteoremet uten bruk av kompleks analyse, og det var derfor en stor sensasjon da den norske matematikeren Atle Selberg (1917–2007) og den ungarske matematikeren Paul Erdős (1913–1996) uavhengig av hverandre publiserte et slikt bevis i 1947. Det hersket uenighet blant de to om hvor uavhengige av hverandre bevisene virkelig var, siden de hadde snakket sammen om problemet før, og de ble aldri helt enige om saken. Uansett brukte begge bevisene et resultat Selberg hadde publisert tidligere. For dette arbeidet, og det "elementære beviset" for primtallsteoremet, ble Atle Selberg tildelt den prestisjefylte *Fields-medaljen* i 1950, den mest eksklusive prisen som deles ut i matematikk. Selberg er så langt den eneste norske vinner av denne prisen.

Åpne formodninger om primtall. Det finnes mange interessante åpne formodninger om primtall. Vi nevner her to:

FORMODNING 3.7.11 (Goldbachs formodning). *Ethvert partall større enn to kan skrives som summen av to primtall.*

For eksempel ser vi at $4 = 2+2$, $6 = 3+3$, $8 = 5+3$, $10 = 7+3$, $12 = 7+5$, ... Før dataalderen ble Goldbachs formodning sjekket for hånd for alle primtall opptil millionstørrelse. Etter at man har datamaskiner tilgjengelig, har man klart å sjekke formodningen opp til primtall av størrelsesorden $2 \cdot 10^7$.

Goldbachs formodning har navnet etter den tyske matematikeren Christian Goldbach (1690–1764), som fremsatte formodningen i et brev til Leonhard Euler i 1742. Det er blitt laget en kortfilm om formodingen i 2011 kalt *Calculus of Love*; en liten snutt er tilgjengelig på

http://people.math.harvard.edu/˜ knill/mathmovies/m4v/calculuslove.m4v.

To primtall sies å være *tvillingsprimtall*[21] hvis deres differanse er 2. De laveste parene av tvillingsprimtall er derfor

$$(3, 5), (5, 7), (11, 13), (17, 19), \dots$$

I 1849 la den franske matematikeren Alphonse de Polignac (1826–1863) frem følgende formodning:

FORMODNING 3.7.12 (Tvillingsprimtallformodningen). *Det finnes uendelig mange tvillingsprimtall.*

For å vise hvor vanskelig dette problemet er, er det nok å nevne at et arbeid av den kinesisk-amerikanske matematikeren Yitang Zhang (1955–) publisert i det prestisjefylte tidsskriftet *Annals of Mathematics* i 2013 viser eksistensen av uendelig mange primtallspar i avstand mindre enn 70.000.000 fra hverandre. Dette var det første arbeidet som viste eksistensen av uendelig mange primtallspar i endelig avstand fra hverandre. (Senere har dette resultatet blitt forbedret til å gjelde avstand mindre eller lik 246.)

Den norske matematikeren Viggo Brun (1885–1978) viste i 1919 et viktig resultat som omhandler tvillingsprimtall, nemlig at *summen av inversene til alle tvillingsprimtall konvergerer.* Dette betyr at summen

$$\left(\frac{1}{3} + \frac{1}{5}\right) + \left(\frac{1}{5} + \frac{1}{7}\right) + \left(\frac{1}{11} + \frac{1}{13}\right) + \left(\frac{1}{17} + \frac{1}{19}\right) + \cdots$$

[21] *Twin primes* på engelsk.

enten har endelig mange ledd (som er tilfellet hvis *tvillingsprimtallsformodningen* er sann) eller har uendelig mange ledd, men konvergerer, det vil si nærmer seg en bestemt sum. (Vi skal snakke om konvergens senere.) Teoremet er kjent som *Bruns teorem* og summen som *Bruns konstant*.

VIGGO BRUN (KILDE: WIKIPEDIA, PUBLIC DOMAIN)

3.8. Binomialteoremet

Vi har lært at

$$(x+y)^2 = x^2 + 2xy + y^2.$$

Dette var allerede kjent hos de gamle greske matematikerne. Videre er det lett å regne ut (og huske utenat) at

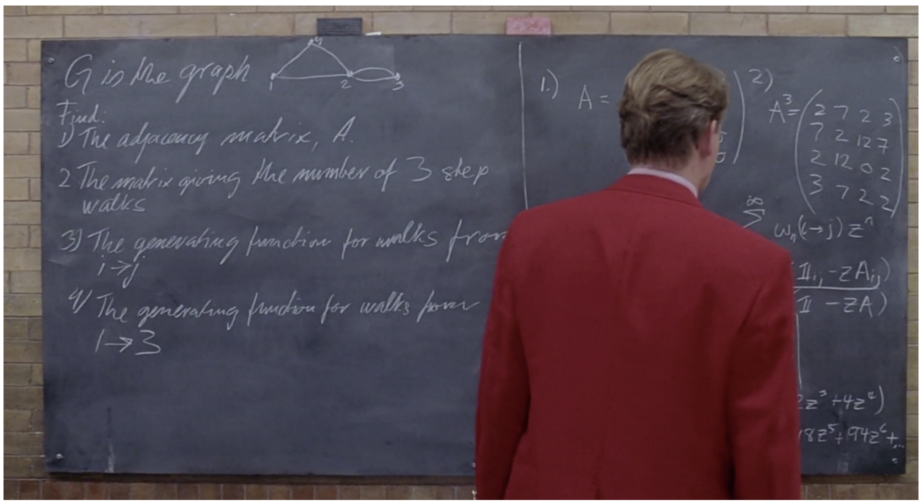
$$(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3,$$

og denne formelen virker å ha vært kjent hos indiske matematikere allerede 600 år f.Kr. Videre kan vi regne ut

$$(x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4.$$

Vi begynner å ane hva mønsteret er: dersom vi multipliserer ut $(x+y)^n$ får vi ledd av typen x^iy^{n-i} for $0 \leq i \leq n$; spørsmålet er hvilke tall, også kalt *koeffisienter*, som skal stå foran leddene. Det er dette det berømte (og svært nyttige!) *binomialteoremet* gir svar på. Teoremet gjelder for alle x og y i en hvilken som helst kommutativ ring. Den første formuleringen av binomialteoremet ble gitt i boken *al-fakhri fi al-jabr wa al-muqabala* fra ca. år 1000 av den persiske matematikeren Al-Karaji (ca. 953–ca. 1029), som beviste teoremet med en implisitt form for induksjon, en teknikk vi skal komme inn på i neste kapittel. Man kan imidlertid også gi et *kombinatorisk* bevis for binomialteoremet, som er det vi skal gjøre i denne seksjonen. Uansett trenger vi en rask gjennomgang av litt kombinatorikk for å forklare teoremet.

Litt kombinatorikk. *Kombinatorikk* er den grenen av matematikk som beskjeftiger seg med studiet av hvor mange kombinasjoner som finnes av ulike typer. Det har derfor mange anvendelsesområder, innenfor for eksempel statistisk fysikk, sannsynlighetsteori og evolusjonsbiologi. Merk at problemene som Will Hunting, spilt av Matt Damon, løser på tavlen på Massachusetts Institute of Technology (MIT) i filmen *Good Will Hunting* fra 1997 regissert av Gus van Sant faller inn under dette fagfeltet (nærmere bestemt i grenen *grafteori*).



PROFESSOR LAMBEAU (SPILT AV STELLAN SKARSGÅRD) SER OVER WILLS LØSNING I FILMEN *Good Will Hunting*. KILDE: MIRAMAX PICTURES

La oss starte med to enkle problemer: Anta at vi har n forskjellige objekter.

- På hvor mange forskjellige måter kan vi ordne disse objektene etter hverandre?
- På hvor mange forskjellige måter kan vi plukke ut k (der $k \leq n$) av disse objektene? Her spiller altså ikke rekkefølgen av utplukkingen noen rolle.

Før vi ser på svarene, vil vi definere noen uttrykk som dukker opp i svarene, og for øvrig dukker opp i så mange sammenhenger i matematikk at de har fått sine egne symboler og betegnelser:

Definisjon 3.8.1: Fakultet

For et naturlig tall n definerer vi

$$0! = 1$$

og

$$n! = 1 \cdot 2 \cdot 3 \cdots n$$

hvis $n \geq 1$. Tegnet “!” leses som “fakultet”[a].

[a] *Factorial* på engelsk.

Ved første øyekast kan definisjonen av 0! virke underlig, men det har vist seg at det er lurt å definere $0! = 1$ på grunn av den naturlige måten dette uttrykket dukker opp i beregninger.

MERKNAD 3.8.2. Vi merker oss at $n!(n+1) = (n+1)!$ for alle naturlige tall n, og dette er en forenkling vi ofte får bruk for i beregninger.

Definisjon 3.8.3: Binomialkoeffisienter

La n og k være naturlige tall slik at $n \geq k$. Vi definerer

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

og leser dette som "n over k"[a]. Tall på formen $\binom{n}{k}$ kalles *binomialkoeffisienter.*

[a] *n choose k* på engelsk. Gitt den kombinatoriske tolkningen av $\binom{n}{k}$ i Setning 3.8.5, er det engelske navnet mye bedre enn det norske.

Legg merke til at siden vi har definert at $0! = 1$, så gir den siste definisjonen også mening for $k = 0$: vi har nemlig $\binom{n}{0} = \frac{n!}{0!(n-0)!} = \frac{n!}{n!} = 1$.

Vi vil nå besvare spørsmålene stilt ovenfor:

Setning 3.8.4

Det finnes $n!$ forskjellige måter å ordne n forskjellige objekter etter hverandre.

BEVIS

Vi starter med å velge objektet som skal stå først i rekken. Vi har selvsagt n forskjellige muligheter for dette valget. Når vi så skal velge neste objekt, har vi kun $n-1$ objekter å velge mellom, så vi har kun $n-1$ forskjellige muligheter. Antallet kombinasjoner for de første to objektene er $n(n-1)$. Objekt nummer tre kan velges mellom $n-2$ forskjellige objekter. Slik fortsetter vi. Når vi skal plukke ut det siste objektet, har vi kun én mulighet. Dette betyr at antallet måter vi kan ordne de n objektene på er

$$n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1 = n!$$

(ved Definisjon 3.8.1). $\square$

Setning 3.8.5

Det finnes $\binom{n}{k}$ forskjellige måter å plukke ut k objekter fra n forskjellige objekter.

Bevis

Vi går frem som i beviset for Setning 3.8.4 og plukker ut de k objektene én etter én. Det første objektet kan velges på n forskjellige måter, det neste på $n-1$, den neste deretter på $n-2$, og så videre helt til det k-te objektet, som kan velges på $n-k+1$ forskjellige måter. Vi får altså

$$n(n-1)(n-2)\cdots(n-k+2)(n-k+1)$$

måter å gjøre dette på. Nå har vi imidlertid ordnet de utplukkede objektene i en bestemt rekkefølge, men rekkefølgen skulle i utgangspunktet ikke ha noe å si. De k objektene vi har plukket ut kan plukkes ut i $k!$ forskjellige rekkefølger ved Setning 3.8.4, så vi får det samme utvalget av de k utvalgte objektene på $k!$ forskjellige måter. Antallet ulike valg av k objekter må derfor være

$$\frac{n(n-1)(n-2)\cdots(n-k+2)(n-k+1)}{k!}.$$

Multipliserer vi teller og nevner med $(n-k)! = 1\cdot 2\cdot 3\cdot(n-k)$, får vi svaret

$$\frac{n!}{k!(n-k)!} = \binom{n}{k}.$$

□

Setningen viser at binomialkoeffisienter er heltall, noen som overhodet ikke er opplagt fra definisjonen. Et alternativt bevis for dette gis i Oppgave 5.11.

Eksempel 3.8.6. Hvor mange forskjellige poker-hender på 5 kort finnes?

Løsning. Poker benytter en standard kortstokk på 52 kort. Antallet forskjellige kombinasjoner av 5 kort er ved Setning 3.8.5 gitt ved

$$\binom{52}{5} = \frac{52!}{5!47!} = 2\ 598\ 960.$$

□

Binomialteoremet. Grunnen til at tallene i Definisjon 3.8.3 har fått navnet *binomialkoeffisienter* er at disse tallene nettopp er de koeffisientene som dukker opp foran leddene når vi multipliserer ut $(x+y)^n$:

Teorem 3.8.7: Binomialteoremet

La R være en kommutativ ring og $n \in \mathbb{N}$. Da er koeffisienten til $x^{n-i}y^i$ når vi multipliserer ut $(x+y)^n$ lik binomialkoeffisienten $\binom{n}{i}$, for alle $i \in \{0, \dots, n\}$.

Altså sier teoremet at

$$\begin{aligned}
(x+y)^n &= \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \cdots + \binom{n}{i}x^{n-i}y^i + \\
&\quad \cdots + \binom{n}{n-2}x^2y^{n-2} + \binom{n}{n-1}x^1y^{n-1} + \binom{n}{n}y^n \\
&= x^n + nx^{n-1}y + \frac{n(n-1)}{2!}x^{n-2}y^2 + \cdots + \binom{n}{i}x^{n-i}y^i + \\
&\quad \cdots + \frac{n(n-1)}{2!}x^2y^{n-2} + nxy^{n-1} + y^n.
\end{aligned}$$

Det kan være instruktivt å først se på beviset i et spesialtilfelle:

EKSEMPEL 3.8.8. For å finne koeffisienten til xy^2 i $(x+y)^3$ regner vi ut ved hjelp av den distributive lov:

$$\begin{aligned}
(x+y)^3 &= (x+y)(x+y)(x+y) \\
&= xxx + xxy + xyx + \underline{xyy} + yxx + \underline{yxy} + \underline{yyx} + yyy \\
&= x^3 + 3x^2y + \underline{3xy^2} + y^3.
\end{aligned}$$

Vi har understreket alle leddene på formen xy^2. For å få et slikt ledd må vi velge én x og to y-er fra de tre faktorene $(x+y)$. Dette tilsvarer å velge de to faktorene hvor y skal tas ut, nemlig:

- andre og tredje faktor, som gir oss xyy;
- første og tredje faktor, som gir oss yxy;
- første og andre faktor, som gir oss yyx.

Dette er tre muligheter, som er lik antallet forskjellige måter å plukke ut 2 objekter fra 3. Ved Setning 3.8.5 er dette lik binomialkoeffisienten $\binom{3}{2}$, som nettopp er lik $\frac{3!}{2!(3-2)!} = \frac{1\cdot2\cdot3}{1\cdot2\cdot1} = 3$.

BEVIS FOR TEOREM 3.8.7
Når vi multipliserer ut

$$(x+y)^n = \underbrace{(x+y)\cdots(x+y)}_{n \text{ ganger}}$$

via den distributive lov får vi ledd på formen $x^{n-i}y^i$. Vi vil finne antallet slike ledd. For å få et slikt ledd må vi velge $n-i$ x-er og i y-er fra de n faktorene $(x+y)$. Dette tilsvarer å velge de i faktorene hvor y skal tas ut. Dette kan ved Setning 3.8.5 gjøres på $\binom{n}{i}$ forskjellige måter, som viser teoremet. $\square$

Binomialteoremet kan også vises ved induksjon, en teknikk vi skal se på i neste kapittel. Vi overlater induksjonsbeviset av binomialteoremet til Oppgave 5.14. Det kombinatoriske beviset gitt ovenfor har fordelen av å være kortere, men er kanskje litt vanskeligere å forstå.

Summetegn. Vi avslutter denne seksjonen med en omtale av summetegnet, som du kanskje har sett før i skolen, siden det er et redskap som er nyttig å bruke videre i boken og i matematikk og realfag generelt.

Vi så ovenfor at når vi skrev ut uttrykket for $(x+y)^n$ rett etter binomialteoremet, ble det temmelig langt og uoversiktlig. I matematikk får vi ofte bruk for å skrive ut slike lange summer der alle leddene har samme form, og da er det hendig å kunne bruke en kortere notasjon. Til dette bruker vi *summetegnet* $\sum$, etter den store greske bokstaven som heter “sigma”. (Den lille bokstaven betegnes med σ.)

La oss se på hvordan vi ville brukt summetegnet i tilfellet vi snakker om: her har alle ledd formen $\binom{n}{i}x^{n-i}y^i$, og vi ønsker å summere alle disse når i går fra 0 til n. Ved hjelp av summetegnet skriver vi dette som

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^i.$$

Her kalles i *summasjonsindeksen*[22] i summen, mens 0 og n er, henholdsvis, *nedre* og *øvre summasjonsgrense*.

La oss se på et annet eksempel:

$$1+3+5+7+\cdots+97+99.$$

Her er leddene alle oddetallene fra 1 til og med 99. Hvert av disse kan skrives som $2i+1$, og summen går fra $i=0$ til $i=49$. Da kan vi skrive dette som

$$1+3+5+7+\cdots+97+99 = \sum_{i=0}^{49}(2i+1).$$

Men vi kunne også skrevet leddene som $2i-1$ og latt summen gå fra $i=1$ til $i=50$ og skrevet:

$$1+3+5+7+\cdots+97+99 = \sum_{i=1}^{50}(2i-1).$$

Det generelle mønsteret er at dersom vi for hvert heltall i har et uttrykk a_i, så betyr $\displaystyle\sum_{i=m}^{n} a_i$ (der $m \leq n$) summen av alle disse uttrykkene fra og med $i=m$ og til og med $i=n$, med andre ord

$$\sum_{i=m}^{n} a_i = a_m + a_{m+1} + a_{m+2} + \cdots + a_{n-1} + a_n.$$

Her er altså i *summasjonsindeksen*, mens m og n er, henholdsvis, *nedre* og *øvre summasjonsgrense*.

[22] *Summation index* på engelsk.

Merk at det er helt tilfeldig at vi har valgt i som bokstav for summasjonsindeksen. Bokstavene n, j og k er like vanlige.

La oss se på et par eksempler til.

EKSEMPEL 3.8.9. Uttrykk summen $1 + x + x^2 + x^3 + \cdots + x^N$ ved hjelp av summetegnet, der $N \in \mathbb{N}$.

LØSNING. Leddene er på formen x^n der n går fra $n = 0$ (siden $x^0 = 1$ ved (62)) til $n = N$. Derfor er

$$1 + x + x^2 + x^3 + \cdots + x^N = \sum_{n=0}^{N} x^n.$$

(Her har vi brukt n som summasjonsindeks.) □

EKSEMPEL 3.8.10. Uttrykk summen $1 - x + x^2 - x^3 + \cdots + x^{99} - x^{100}$ ved hjelp av summetegnet.

LØSNING. Dette er samme eksemplet som det forrige, med $N = 100$, bortsett fra at leddene veksler fortegn. I slike tilfeller er det nyttig å huske at uttrykket $(-1)^n$ veksler mellom 1 og -1 etter som n er et partall eller oddetall. Derfor kan vi skrive leddene i uttrykket som $(-1)^n x^n$ og vi får:

$$1 - x + x^2 - x^3 + \cdots + x^{99} - x^{100} = \sum_{n=0}^{100} (-1)^n x^n.$$

□

Vi avslutter med tre nyttige regneregler for summetegn:

$$\sum_{i=m}^{n} a_i + \sum_{i=m}^{n} b_i = \sum_{i=m}^{n} (a_i + b_i). \tag{64}$$

$$\sum_{i=m}^{n} (ca_i) = c \sum_{i=m}^{n} a_i. \tag{65}$$

$$\sum_{i=m}^{n} a_i + \sum_{i=n+1}^{k} a_i = \sum_{i=m}^{k} a_i. \tag{66}$$

Reglene sier altså at vi kan

- slå sammen summer med de samme summasjonsgrensene;
- sette en felles faktor utenfor summetegnet;
- slå sammen to summer når nedre summasjonsgrense i én sum er 1 større enn øvre summasjonsgrense i den andre summen.

Det er enkelt å sjekke at disse reglene holder ved å skrive ut hva hver side i likhetene betyr. Vi overlater dette til Oppgave 3.35.

Oppgaver

Oppgaver til §3.1–§3.3

Oppgave 3.1. Regn ut $2 + 2 = S(S(0)) + S(S(0))$ ved å bruke (40) og (41), som i Eksempel 3.2.2.

"Even in this corner of the galaxy, Captain, two plus two equals four." Mr. Spock til Captain Kirk i episoden *The Conscience of the King* (1966) fra TV-serien *Star Trek*. (Kilde: Wikipedia, public domain)

Oppgave 3.2. Bevis Setning 3.2.6.

Oppgave 3.3. Bevis Setning 3.2.7.

Oppgave 3.4. Bevis Setning 3.2.9. (Hint: Vis først at vi har $(m + n)r = mr + nr$ for alle $m, n, r \in \mathbb{N}$ og bruk dette til å bevise (i).)

Oppgave 3.5. Vis at relasjonen R på $\mathbb{N}^2$ definert i (50) er en ekvivalensrelasjon. (Hint: du får bruk for Setning 3.2.11(i).)

Oppgave 3.6. Bevis Setning 3.3.4.

Oppgave 3.7. Verifiser de aritmetiske egenskapene til $\mathbb{Z}$ i Setning 3.3.5.

Oppgaver til §3.4–3.5

Oppgave 3.8. Vis at relasjonen (59) på $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$ er en ekvivalensrelasjon. (Hint: du får bruk for Setning 3.2.11(iv).)

Oppgave 3.9. Bevis Setning 3.4.2.

Oppgave 3.10. Verifiser de aritmetiske egenskapene til $\mathbb{Q}$ i Setning 3.4.3.

Oppgave 3.11. Bevis Setning 3.5.2, først på $\mathbb{Z}$, deretter på $\mathbb{Q}$. (Hint: du får bruk for (54) og (56).)

Oppgave 3.12. La R være en relasjon på en ikke-tom mengde X. Kan R samtidig være en ekvivalensrelasjon og en ordensrelasjon?

Oppgave 3.13. Bevis Setning 3.5.3. (Hint: du får bruk for (56).)

Oppgave 3.14. Fullfør beviset på Setning 3.5.4 ved å vise at $M = \mathbb{N}$.

Oppgaver til §3.6

Oppgave 3.15. La R være en kommutativ ring. Bevis at identitetselementet 1 er entydig. Med andre ord, vis at dersom 1 og $1'$ begge oppfyller aksiom (M3), da er $1 = 1'$. (Se også Merknad 3.6.4.)

Oppgave 3.16. Bevis resten av Setning 3.6.3.

Oppgave 3.17. Verifisér ordensrelasjonene i Eksempel 3.6.11.

Oppgave 3.18. Bevis (iii)-(v) i Setning 3.6.13.

Oppgave 3.19. Bevis Setning 3.6.16.
Utled også følgende konsekvens av trekantulikheten:

$$||a| - |b|| \leq |a - b| \quad \text{for alle } a, b \text{ i en ordnet kropp.}$$

Oppgaver til §3.7

Oppgave 3.20. (a) Finn feilen i følgende “bevis” av del (iii) av Setning 3.7.2.

"Bevis": Siden $a|b$, har vi $b = ka$ for $k \in \mathbb{N}$. Likeledes, siden $b|c$, har vi $c = kb$. Dermed har vi $c = bk = (ak)k = ak^2$. Det følger at $a|c$. $\square$

(b) Bevis Setning 3.7.2.

Oppgave 3.21. Finn primtallsfaktoriseringen av 45617.

Oppgave 3.22. Vis at dersom a og b er positive heltall uten felles primfaktorer (jf. Definisjon 3.7.8), og a deler bc, for et positivt heltall c, da må a dele c. (Hint: Bruk den entydige primtallsfaktoriseringen til a, b og c.)

Oppgave 3.23. Vis at $\sqrt{3}$ er irrasjonalt. (Hint: Mim etter beviset for Setning 1.3.5 og bruk *aritmetikkens fundamentalteorem* (3.7.7).)

Oppgave 3.24. Betrakt følgende relasjoner på mengden av positive heltall $\mathbb{Z}^+$:

$$\begin{aligned} R_1 &= \{(x, y) \mid x \text{ og } y \text{ har ingen felles primfaktorer}\} \\ R_2 &= \{(x, y) \mid x \text{ og } y \text{ har de samme primfaktorene}\} \end{aligned}$$

(jf. Definisjon 3.7.8).

(a) Avgjør om relasjonene er ekvivalensrelasjoner.
(b) I tilfellene som er ekvivalensrelasjoner: Hva er ekvivalensklassen til 1? Finn ekvivalensklassen til et vilkårlig heltall $n \geq 2$ uttrykt på enklest mulig måte.

Oppgaver til §3.8

Oppgave 3.25. Regn ut $(x+y)^6$.

Oppgave 3.26. Finn koeffisienten til x^8y^9 i $(3x+2y)^{17}$.

Oppgave 3.27. La $m, n \in \mathbb{N}$ med $m < n$. Vis at
(a) $\binom{m}{k} < \binom{n}{k}$ for alle $k \in \{1, \dots, m\}$;
(b) $\frac{1}{m^k}\binom{m}{k} < \frac{1}{n^k}\binom{n}{k} \leq \frac{1}{k!} \leq \frac{1}{2^{k-1}}$ for alle $k \in \{2, \dots, m\}$.

Oppgave 3.28. La $n \in \mathbb{N}$. Vis at

$$2^n = \sum_{k=0}^{n} \binom{n}{k}$$

ved å
(a) bruke binomialteoremet;
(b) bruke et kombinatorisk argument, samt at en mengde med n objekter har nøyaktig 2^n forskjellige delmengder (dette skal vi bevise i Eksempel 5.2.10).

Oppgave 3.29. La $n \in \mathbb{N}$. Vis at

$$3^n = \sum_{k=0}^{n} 2^k \binom{n}{k}.$$

Oppgave 3.30. La $n \in \mathbb{N}$. Vis at

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0.$$

Oppgave 3.31. La $m, n, \in \mathbb{N}$ slik at $r \leq m$ og $r \leq n$. Vis *Vandermondes identitet*

$$\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{r-k}\binom{n}{k},$$

oppkalt etter den franske matematikeren Alexandre-Théophile Vandermonde (1735–1796), selv om formelen allerede finnes i boken *Siyuan yujian* fra 1303 av den kinesiske matematikeren Zhu Shijie (1249–1314).

Utled som korollar at

$$\binom{2n}{n} = \sum_{k=0}^{n} \binom{n}{k}^2$$

for alle naturlige tall n.

Oppgave 3.32. Du har fått deg sommerjobb på et gatekjøkken som serverer tre retter: pølse, kebab og hamburger. Det kommer en gjeng på 13 personer som hver bestiller én av de tre rettene. Hvor mange forskjellige bestillinger er mulig fra en gruppe på 13 personer? (Her er vi selvsagt bare interessert i antallet retter av hver type, og ikke hvilken rett som bestilles av hver enkelt kunde.)

Oppgave 3.33. Hvor mange forskjellige mønstre kan vi oppnå ved å legge n kuler etter hverandre, hvorav k er røde og $n-k$ er hvite?

Oppgave 3.34. Hvor mange forskjellige "ord" (som ikke trenger å gi mening) kan vi lage ved å stokke om på bokstavene i ordet "kulturuke"? Sjekk om den norske lyrikeren Jan Erik Vold (1939-) har tatt med alle i diktet "Kulturuke" fra diktasamlingen *kykelipi* fra 1969; se

https://www.youtube.com/watch?v=OmRrpMzrYQw.

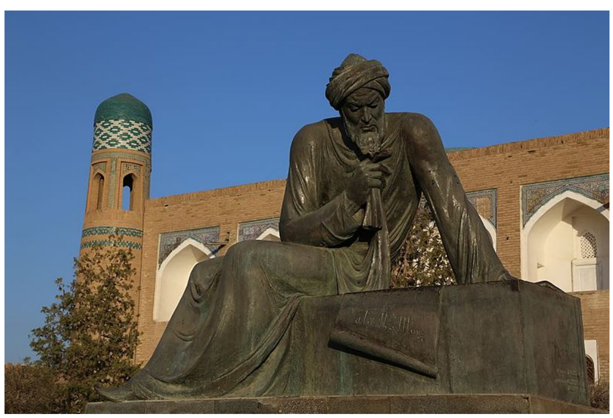
Oppgave 3.35. Sjekk (64)-(66).

KAPITTEL 4

Flere eksempler på ringer og kropper

Det er instruktivt, på det nåværende tidspunkt, å se på andre eksempler på ringer og kropper enn $\mathbb{Z}$ og $\mathbb{Q}$.

Studiet av ringer og kropper, samt en hel rekke andre algebraiske strukturer, er kjent som *abstrakt algebra* eller *moderne algebra*, for å skille det fra den mer elementære algebraen vi er kjent med fra skolen, som vi for eksempel finner i manipulasjoner av ligninger og i *Pythagoras' teorem.* Selve ordet *algebra* kommer fra det arabiske ordet *al-jabr*, som betyr å restaurere. Ordet kommer fra tittelen på boken *al-Kitab al-mukhtasar fi hisab al-jabr wa'l-muqabala*, som kan oversettes som "Den konsise boken om regning ved komplettering og balansering", av den persiske matematikeren Muhammad ibn Musa al-Khwarizmi (ca. 780–ca. 850). Foruten å inneholde den første generelle løsningen for annengradsligninger, er dette den første boken som introduserer og konsekvent bruker dagens velkjente algebraiske manipulasjoner som å multiplisere begge sider av en ligning med samme tall, og addere eller subtrahere begge sider av en ligning med samme tall. Det er dette som ligger bak ordene "restaurering", "komplettering" og "balansering". Derfor regnes al-Khwarizmi som faren til moderne algebra. For øvrig har navnet til al-Khwarizmi gitt opphavet til ordet "algoritme".

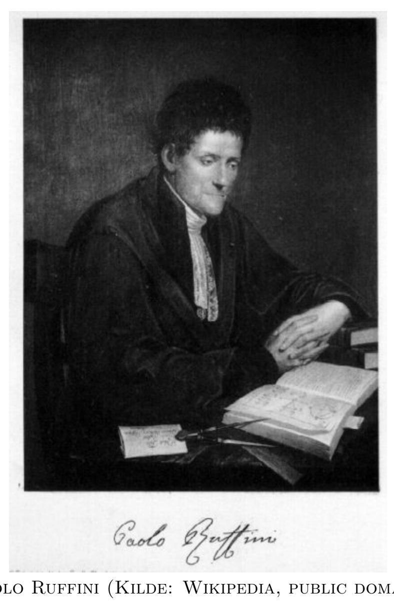


Statue av al-Khwarizmi i fødebyen Khiva i Uzbekistan (Kilde: Wikipedia, public domain.)

Metoder for å løse forskjellige typer ligninger av høyere grad spilte en viktig rolle i den videre utviklingen av algebra, og viktige bidrag ble gjort av de persiske matematikerne Omar Khayyam (1048–1131), Al-Karaji (ca. 953– ca. 1029) og Sharaf al-Din al-Tusi (ca. 1135– ca. 1213), de indiske matematikerne Mahāvīrā (800-tallet) og Bhāskara (ca. 1114–1185) og den kinesiske matematikeren Zhu Shijie (1249–1314). Araberne brakte algebraen til Europa, og den italienske matematikeren Fibonacci (ca. 1170–1245) bidro også til teorien om løsninger av tredjegradsligninger.

På 1500-tallet kom endelig de generelle løsningene på tredje- og fjerdegradsligninger, med bidrag av de italienske matematikerne Scipione del Ferro (1465-1526), Niccoló Fontana Tartaglia (1499-1557), Girolamo Cardano (1501-76), Ludovico Ferrari (1522-65) og Rafaele Bombelli (1526-1572). Som en konsekvens ble de komplekse tallene konstruert, noe vi skal lese mer om i Kapittel 8. Løsningsformler for ligninger av grad 3 og 4 ble først publisert i boken *Ars Magna* av Cardano i 1545. Med *løsningsformel* mener vi her et uttrykk som involverer koeffisientene i ligningen og endelig mange av de aritmetiske operasjonene addisjon, subtraksjon, multiplikasjon, divisjon, samt røtter. Dersom løsningene til en ligningstype kan skrives på en slik måte, sier vi at ligningene er *løsbare ved rotutdragning.* Selv om tredje- og fjerdegradsligninger er løsbare ved rotutdragning, er imidlertid løsningsformlene såpass lange at man ikke lærer dem på skolen.

I de etterfølgende århundrene konsentrerte mange matematikere seg på søken etter løsningsformler på ligninger av grad større eller lik 5. Det viste seg imidlertid at *et generelt polynom av grad* > 4 *ikke er løsbart ved rotutdragning.* Dette berømte resultatet kalles *Abel–Ruffini teoremet*, etter den italienske matematikeren Paolo Ruffini (1765–1822), som gav et ufullstedig bevis i 1799, og den norske matematikeren Niels Henrik Abel (1802–1829), som gav et komplett bevis i 1824. I denne forbindelsen utviklet Paolo Ruffini teorien om såkalte *permutasjonsgrupper* (se Eksempel 4.5.3), et av de viktigste eksemplene på den fundamentale algebraiske strukturen *gruppe* (se Definisjon 4.5.1).



Paolo Ruffini (Kilde: Wikipedia, public domain)

Niels Henrik Abel (Kilde: Wikipedia, public domain)

Abel døde bare 26 år gammel av tuberkolose og klarte dermed ikke å nå sitt mål om å beskrive nøyaktig hvilke ligninger som er løsbare ved rotutdragning. Teorien som beskriver nøyaktig hvilke betingelser koeffisientene i et polynom må oppfylle for at det skal være løsbart ved rotutdragning ble utviklet av den unge franske matematikeren Évariste Galois (1811–1832), og bærer i dag navnet *Galoisteori*. Denne vakre teorien forbinder løsbarheten til en ligning med symmetriegenskaper til koeffisientene, og *symmetrigrupper* spiller en sentral rolle (se Eksempel 4.5.4). Galois døde bare 20 år gammel i en duell, og skriblet ned mye av teorien i et brev til en god venn bare dager før duellen, siden han var sikker på å dø.

Évariste Galois (Kilde: Wikipedia, public domain)

Galoisteorien har en stor del av æren for at abstrakt algebra utviklet seg sterkt i løpet av 1800- og 1900-tallet parallelt med behovet for strengere rigorøsitet i matematikk. Viktige pionerer i arbeidet ble de tyske matematikerne Ernst Kummer (1810–1893), Leopold Kronecker (1823–1891), Richard Dedekind (1831–1916), Ferdinand Georg Frobenius (1849–1917), Felix Klein

(1849–1925), Ernst Steinitz (1871–1928) og David Hilbert (1862–1943), og den østerrikske matematikeren Emil Artin (1898–1962). Utviklingen nådde et høydepunkt med den tyske matematikeren Emmy Noether (1882–1935) og hennes skole ved Universitetet i Göttingen. Noether underviste ved dette universitetet allerede fra 1915 etter invitasjon av Klein og Hilbert, men fikk en betalt stilling først i 1923. De første årene underviste hun uten lønn og før 1919 hadde hun som kvinne ikke engang tillatelse til å undervise ved universitetet, så løsningen var at Hilbert måtte annonsere forelesningene i sitt navn, mens Noether holdt dem. Til tross for at Emmy Noether er en av de mest innflytelsesrike matematikerne noensinne, fikk hun aldri en professorstilling i Göttingen. Hun var i tillegg jøde og mistet dermed sin undervisningsstilling ved Universitetet i Göttingen da Hitler overtok makten i 1933. Hun emigrerte til USA, hvor hun fikk en stilling ved det tradisjonsrike kvinnecolleget Bryn Mayr i Pennsylvania. Dessverre døde hun brått og uventet etter en operasjon i 1935.

Emmy Noether (Kilde: Wikipedia, public domain)

Vi vil i resten av kapitlet se på en del eksempler på ringer og kropper, og avslutte så med en liten seksjon om andre viktige algebraiske strukturer, nemlig *grupper* og *vektorrom*. Disse vil alle spille sentrale roller i senere kurs, som for eksempel i kursene *MAT121–Lineær Algebra* og *MAT220–Algebra*. Bortsett fra *faktorteoremet* (Setning 4.1.9), som vi kommer til å få bruk for i beviset for *algebraens fundamentalteorem* 8.4.1 i §8.6, vil vi i denne boken ikke bruke noe annet fra dette kapitlet.

4.1. Polynomring i én variabel

La R være en kommutativ ring. Vi definerer $R[x]$ til å være mengden av alle polynomer i én variabel med koeffisienter i R, det vil si alle uttrykk på formen

$$P(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n, \quad n \in \mathbb{N}, \; a_0, \ldots, a_n \in R, \; n \in \mathbb{N}.$$

Her kaller vi $a_i$ene for *koeffisientene* i polynomet og vi pleier som oftest å droppe koeffisienter som er 0, slik at vi for eksempel skriver $2+3x+5x^3$ for $2+3x+0x^2+5x^3$. Elementet i ringen $R[x]$ betegnes gjerne kun med P.

Merk også at vi ved å sette inn et element $c \in R$ istedenfor x, får uttrykket

$$P(c) = a_0 + a_1c + a_2c^2 + \cdots + a_nc^n, \tag{67}$$

som er et element i R, siden alle $a_i \in R$. Altså definerer ethvert element $P \in R[x]$ en funksjon

$$\begin{array}{rcl} R & \longrightarrow & R \\ c & \mapsto & P(c) = a_0 + a_1c + a_2c^2 + \cdots + a_nc^n, \end{array} \tag{68}$$

som kalles *evalueringen*[1] *av* P.

Vi definerer operasjonene $+$ og $\cdot$ på denne mengden som vanlig addisjon og multiplikasjon av polynomer. Vi minner om hvordan disse er definert.

La $f(x) = a_0+a_1x+a_2x^2+\cdots+a_nx^n$ og $g(x) = b_0+b_1x+b_2x^2+\cdots+b_mx^m$ og anta for enkelhets skyld at $n \geq m$. Da er

$$\begin{aligned} f(x)+g(x) = \quad & (a_0+b_0)+(a_1+b_1)x+(a_2+b_2)x^2+\cdots+(a_m+b_m)x^m \\ & +a_{m+1}x^{m+1}+\cdots+a_nx^n \end{aligned}$$

og

$$f(x)\cdot g(x) = c_0 + c_1x + c_2x^2 + \cdots + c_{n+m}x^{n+m},$$

hvor

$$c_k = \sum_{i=0}^{k} a_ib_{k-i} = a_0b_k + a_1b_{k-1} + a_2b_{k-2} + \cdots + a_{k-2}b_2 + a_{k-1}b_1 + a_kb_0.$$

I praksis betyr dette altså at:

- addisjon fås ved å addere polynomene ledd for ledd;
- multiplikasjon fås ved å multiplisere ut polynomene ved å "late som" den distributive loven finnes, og så samle sammen alle ledd med samme potens av x.

Setning 4.1.1

La R være en kommutativ ring. Da er $R[x]$ med operasjonene $+$ og $\cdot$ definert ovenfor en kommutativ ring, men ikke en kropp. Det nøytrale elementet i ringen er det konstante polynomet 0, og det multiplikative identiteselementet er det konstante polynomet 1.

Beviset for Setning 4.1.1 er relativt rett frem, men er ganske langt og vi gir det ikke her. Det er imidlertid enkelt å se at $R[x]$ ikke er en kropp:

[1] *Evaluation* på engelsk.

Setning 4.1.2

Elementet $x \in R[x]$ har ingen multiplikativ invers. Spesielt er ikke $R[x]$ en kropp.

BEVIS
Vi argumenterer ved selvmotsigelse og antar at polynomet x har en multiplikativ invers, med andre ord at det finnes et polynom $P(x) \in R[x]$ slik at $xP(x) = 1$. Vi kan da sette inn $x = 0$ (eller "evaluere begge sider i 0") og få $0P(0) = 1$. Men venstresiden er 0 ved Setning 3.6.3(ix), slik at vi får $0 = 1$. Men dette motsier aksiom (M3) for en ring, som krever at $0 \neq 1$. □

Definisjon 4.1.3: Polynomring i én variabel

La R være en kommutativ ring. Vi kaller $R[x]$ med operasjonene $+$ og $\cdot$ *polynomringen i én variabel over R.*

Når $R = K$ er en kropp har polynomringen $K[x]$ mye til felles med ringen $\mathbb{Z}$. For eksempel finnes en divisjonsalgoritme i $K[x]$ som er helt tilsvarende *divisjonsalgoritmen* i $\mathbb{Z}$ (Setning 3.7.3). Før vi gir resultatet trenger vi en definisjon.

Definisjon 4.1.4: Graden til et polynom

Graden[a] til et polynom $P \in R[x]$ slik at $P \neq 0$ er eksponenten til den høyeste potensen av x i polynomet (med ikke-null koeffisient). Med andre ord, hvis

$$P(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n, \quad \text{med} \quad a_n \neq 0,$$

da er graden til P lik n. Vi betegner graden med $\deg P$.

Graden til det konstante polynomet 0 defineres til å være -1.

[a] *Degree* på engelsk.

MERKNAD 4.1.5. I noen lærebøker og matematiske tekster defineres graden til det konstante polynomet 0 til å være $-\infty$.

Vi merker at polynomer av grad 0 er polynomer på formen $P(x) = a_0$, for $a_0 \in R \setminus \{0\}$, det vil si de konstante polynomene forskjellig fra null.

Setning 4.1.6: Divisjonsalgoritmen for polynomer

La K være en kropp. Hvis $F, G \in K[x]$ og $G \neq 0$, da finnes entydige $Q, R \in K[x]$ slik at $\deg R < \deg G$ og $F = GQ + R$.

På samme måte som beviset for divisjonsalgoritmen i $\mathbb{Z}$ (Setning 3.7.3), bygger beviset for Setning 4.1.6 på *velordningsprinsippet* for de naturlige tallene, som vi skal komme til i §5.1. Beviset er temmelig likt beviset for Setning 3.7.3 og vil overlates til Oppgave 5.3.

Følgende definisjon er helt parallell til Definisjon 3.7.1:

Definisjon 4.1.7: Delelighet for polynomer

La K være en kropp og $F, G \in K[x]$ med $G \neq 0$, Vi sier at G *deler* F, ekvivalent at F *er delelig på* G, dersom det finnes en $Q \in K[x]$ slik at $F = GQ$.

I dette tilfellet sier vi også at G *er en faktor i* F .

Vi tar også med:

Definisjon 4.1.8: Rot

La K være en kropp og $P \in K[x]$. Vi sier at $a \in K$ er en *rot*[a] eller *nullpunkt*[b] i P dersom $P(a) = 0$.

[a] *Root* på engelsk.

[b] *Zero* på engelsk, som er mer vanlig enn *root*.

Et par nyttige konsekvenser av siste setning er:

Følgesetning 4.1.9: Faktorteoremet

La K være en kropp, $a \in K$ og $P \in K[x]$. Da er a en rot i P hvis og bare hvis P er delelig på $x - a$.

Bevis

Husk at "hvis og bare hvis" er det samme som en dobbel implikasjonspil "$\Longleftrightarrow$". Vi skal altså vise

$$a \text{ er en rot i } P \Longleftrightarrow x - a \text{ deler } P.$$

Vi viser først implikasjonen "$\Longleftarrow$" (det vi si "hvis"-delen).

Hvis $x - a$ deler P, da kan vi skrive $P(x) = (x - a)Q(x)$, for en $Q \in K[x]$. Setter vi inn $x = a$, får vi $P(a) = 0 \cdot Q(a) = 0$, slik at a er en rot i P.

Vi viser dernest implikasjonen "$\Longrightarrow$" (det vi si " bare hvis"-delen).

Ved divisjonsalgoritmen (Setning 4.1.6), med $F = P$ og $G = x - a$, kan vi skrive $P(x) = (x - a)Q(x) + R(x)$ for en $R \in K[x]$ som oppfyller $\deg R < \deg(x - a) = 1$. Altså er enten $\deg R = 0$, som betyr at R er et konstant polynom forskjellig fra null, eller $\deg R = -1$, som betyr at R er nullpolynomet. Uansett kan vi skrive $R(x) = c$ for en $c \in K$. Altså er $P(x) = (x - a)Q(x) + c$. Hvis a er en rot i P, får vi ved innsetting:

$$0 = P(a) = 0 \cdot Q(a) + c,$$

som gir at $c = 0$. Altså har vi $P(x) = (x - a)Q(x)$, som viser at $x - a$ deler P. □

Følgesetning 4.1.10

La K være en kropp og $P \in K[x] \setminus \{0\}$ med $\deg P = n$. Da har P høyst n røtter i K.

Bevis
Overlates til Oppgave 4.1. □

4.2. Funksjonsringer*

La X være en ikke-tom mengde og R en kommutativ ring. La $\mathcal{F}(X, R)$ være mengden av alle funksjoner $f : X \to R$. (For enkelhets skyld, kan du godt bare tenke deg at $X = R = \mathbb{Z}$, $\mathbb{Q}$ eller $\mathbb{R}$.) For alle $f, g \in F$, definér funksjonene

$$\begin{aligned} f + g : X &\longrightarrow R \\ x &\mapsto f(x) + g(x) \end{aligned}$$

og

$$\begin{aligned} f \cdot g : X &\longrightarrow R \\ x &\mapsto f(x) \cdot g(x). \end{aligned}$$

Setning 4.2.1

La X være en ikke-tom mengde og R en kommutativ ring. Da er $\mathcal{F}(X, R)$ med operasjonene $+$ og $\cdot$ definert ovenfor en kommutativ ring, med nøytralt element funksjonen som er identisk lik 0 og identitetselement funksjonen som er identisk lik 1.

Bevis
Overlates til Oppgave 4.2. □

Setning 4.2.2

Dersom X inneholder minst to elementer, er $\mathcal{F}(X, R)$ ikke en kropp.

Bevis
Velg nemlig en $a \in X$. Da er $X \setminus \{a\} \neq \emptyset$, siden vi antar at X inneholder minst to elementer, og vi kan definere to funksjoner $f, g : X \to R$ ved

$$f(x) = \begin{cases} 0, & x = a \\ 1, & x \neq a \end{cases} \quad \text{og} \quad g(x) = \begin{cases} 1, & x = a \\ 0, & x \neq a. \end{cases}$$

Da vil

$$f(x)g(x) = \begin{cases} 0 \cdot 1 = 0, & x = a \\ 1 \cdot 0 = 0, & x \neq a, \end{cases}$$

(hvor vi har brukt Setning 3.6.3(ix)). Dette betyr at funksjonen $fg = 0$ i ringen $\mathcal{F}(X, R)$, selv om $f \neq 0$ og $g \neq 0$. Med andre ord er egenskapen (x) i Setning 3.6.3 *ikke* oppfylt, som viser at $\mathcal{F}(X, R)$ ikke kan være en kropp. □

Dette viser samtidig at egenskapen (x) i Setning 3.6.3 ikke trenger å være oppfylt for kommutative ringer, jf. Merknad 3.6.7.

Vi overlater til Oppgave 4.3 å finne ut av hva som skjer hvis X består av kun ett element.

4.3. Ringen av heltall modulo m*

La $m \in \mathbb{Z}^+$ være et heltall. La $\equiv \pmod{m}$ være ekvivalensrelasjonen på $\mathbb{Z}$ fra Eksempel 2.4.14 og Merknad 2.4.15. Vi minner om hva det betyr:

$$x \equiv y \pmod{m} \iff m \text{ deler } |x - y| \iff x = y + km \text{ for en } k \in \mathbb{Z},$$

hvor høyre ekvivalens skyldes definisjonen av *deling* (Definisjon 3.7.1). Vi sier at x og y er *kongruente modulo m*.

Ekvivalensklassen til en vilkårlig $n \in \mathbb{Z}$ er

$$[n]_m = \{k \in \mathbb{Z} \mid m \text{ deler } |n - k|\} = \{n, n \pm m, n \pm 2m, n \pm 3m, \dots\},$$

og hver slik kaller vi en *kongruensklasse*. Hver kongruensklasse $[n]_m$ inneholder et entydig heltall n_0 som oppfyller $0 \leq n_0 \leq m - 1$. Dette heltallet n_0 er ikke noe annet enn resten vi får når n deles på m, jf. Setning 3.7.3. Derfor kan mengden av alle ekvivalensklasser, som vi betegner med $\mathbb{Z}_m$, beskrives som

$$\mathbb{Z}_m = \{[0]_m, [1]_m, [2]_m, \dots, [m-1]_m\}.$$

Vi overlater det til Oppgave 4.4 å vise at addisjon og multiplikasjon bevarer kongruenser, det vil si at

$$\begin{gathered} x \equiv y \pmod{m} \ \text{ og } \ z \equiv w \pmod{m} \\ \implies x + z \equiv y + w \pmod{m} \ \text{ og } \ xz \equiv yw \pmod{m} \end{gathered} \tag{69}$$

Vi kan dermed definere to veldefinerte operasjoner $+ : \mathbb{Z}_m \times \mathbb{Z}_m \to \mathbb{Z}_m$ og $\cdot : \mathbb{Z}_m \times \mathbb{Z}_m \to \mathbb{Z}_m$ ved $[x]_m + [y]_m = [x + y]_m$ og $[x]_m \cdot [y]_m = [xy]_m$. Operasjonene kalles *addisjon modulo m* og *multiplikasjon modulo m*, henholdsvis.

For å få litt intuisjon om dette, kan man bare tenke på at tilfellet $m = 12$ (eller $m = 24$ om man vil) ikke er annet enn den matematikken vi gjør når vi regner på klokken: Da vil for eksempel $[9]_{12} + [5]_{12} = [14]_{12} = [2]_{12}$ og $[3]_{12} \cdot [5]_{12} = [15]_{12} = [3]_{12}$. Vi regner altså ut på vanlig måte først; sluttsvaret blir resten vi får når svaret deles på 12. Slik virker også kongruensregning modulo m: vi adderer og multipliserer på vanlig måte i $\mathbb{Z}$ først, og tar resten av det svaret vi får når vi deler på m.

Setning 4.3.1

$\mathbb{Z}_m$ er en kommutativ ring.

BEVIS

Detaljene overlates til Oppgave 4.7. Det nøytrale elementet for addisjon er $[0]_m$ og det multiplikative identitetselementet er $[1]_m$; videre er $[-x]_m = [m-x]_m$ det additive inverselementet til $[x] \in \mathbb{Z}_m$. $\square$

Definisjon 4.3.2: Ringen av heltall modulo m

Vi kaller ringen $\mathbb{Z}_m$ for *ringen av heltall modulo m*.

Det er imidlertid ikke alltid slik at $\mathbb{Z}_m$ er en kropp: for eksempel har ikke $[2]_6$ noen multiplikativ invers i $\mathbb{Z}_6$; dette kan vi verifisere ved å skrive ned multiplikasjonstabellen for $[2]_6$ i $\mathbb{Z}_6$:

$$\begin{aligned}
[2]_6 \cdot [0]_6 &= [2 \cdot 0]_6 = [0]_6 \\
[2]_6 \cdot [1]_6 &= [2 \cdot 1]_6 = [2]_6 \\
[2]_6 \cdot [2]_6 &= [2 \cdot 2]_6 = [4]_6 \\
[2]_6 \cdot [3]_6 &= [2 \cdot 3]_6 = [6]_6 = [0]_6 \\
[2]_6 \cdot [4]_6 &= [2 \cdot 4]_6 = [8]_6 = [2]_6 \\
[2]_6 \cdot [5]_6 &= [2]_6 \cdot [5]_6 = [10]_6 = [4]_6
\end{aligned}$$

Vi ser at vi aldri får $[1]_6$ til svar. Vi ser også noe annet merkelig: de to elementene $[2]_6$ og $[3]_6$ i $\mathbb{Z}_6$ er forskjellig fra $[0]_6$ (siden 6 ikke deler noen av dem), men deres produkt er $[0]_6$. Dette er ikke mulig i en kropp, ved Setning 3.6.3(x), og er noe vi ikke har erfart tidligere i $\mathbb{Z}$ eller i $\mathbb{Q}$. Vi ser også at "forkortningsloven" i Setning 3.6.3(v) ikke gjelder: vi har for eksempel $[2]_6 \cdot [1]_6 = [2]_6 \cdot [4]_6$, men $[1]_6 \neq [4]_6$.

Det sentrale poenget i eksemplet ovenfor, som viser at $\mathbb{Z}_6$ ikke er en kropp, er at 6 ikke er et primtall, som følgende berømte resultat viser:

Setning 4.3.3

$\mathbb{Z}_m$ er en kropp hvis og bare hvis m er primtall.

BEVIS

Husk at "hvis og bare hvis" er det samme som en dobbel implikasjonspil "$\Longleftrightarrow$".

Vi viser først retningen "$\Longrightarrow$" ved et kontrapositivt bevis. Vi antar derfor at m ikke er primtall. Da er $m = ab$ for $a \geq 2$, $b \geq 2$ positive heltall. Da er $1 < a < m$ og $1 < b < m$, slik at $[a]_m \neq [0]_m$ og $[b]_m \neq [0]_m$ i $\mathbb{Z}_m$, mens $[a]_m \cdot [b]_m = [a \cdot b]_m = [m]_m = [0]_m$. Dette motsier Setning 3.6.3(x) for kropper. Altså er $\mathbb{Z}_m$ ikke en kropp hvis m ikke er primtall.

Vi viser dernest retningen “ $\Longleftarrow$”. Siden vi har vist at $\mathbb{Z}_m$ er en kommutativ ring (Setning 4.3.1), mangler det bare å vise at ethvert ikkenull element i $\mathbb{Z}_m$ har en multiplikativ invers når m er et primtall.

Husk at ethvert element i $\mathbb{Z}_m$ kan skrives som $[n]_m$, der $[n]_m$ er ekvivalensklassen $[n]_m$ til et heltall n med $0 \leq n < m$. Altså har vi

$$\mathbb{Z}_m = \{[0]_m, [1]_m, [2]_m \ldots, [m-1]_m\}.$$

La nå $[a]_m$ være et vilkårlig element i $\mathbb{Z}_m$ forskjellig fra $[0]_m$ og betrakt mengden:

$$\begin{aligned}\{[0]_m \cdot [a]_m, [1]_m \cdot [a]_m, [2]_m \cdot [a]_m, \ldots, [m-1]_m \cdot [a]_m\} \\ = \{[0 \cdot a]_m, [1 \cdot a]_m, [2a]_m, \ldots, [(m-1)a]_m\}.\end{aligned}$$

PÅSTAND 4.3.3.1. Alle elementene i siste mengde er forskjellige.

BEVIS FOR PÅSTAND. Vi argumenterer ved motsigelse. Hvis påstanden ikke er sann, finnes det forskjellige heltall $0 \leq k, l \leq m-1$ slik at $[ka]_m = [la]_m$. Vi kan ved symmetri anta $k > l$. Dette medfører at $k - l > 0$ og m deler $ka - la = (k-l)a$. Siden m er et primtall og $a < m$, har ikke a og m felles faktorer, og derfor må m dele $k - l$ (jf. Oppgave 3.22). Men dette er umulig, siden $0 < k - l < k < m$. □

Dermed består mengden $\{[0 \cdot a]_m, [1 \cdot a]_m, [2a]_m, \ldots, [(m-1)a]_m\}$ av forskjellige elementer, og antallet er m. Siden dette er en delmengde av $\mathbb{Z}_m$, som også har m elementer, har vi

$$\mathbb{Z}_m = \{[0 \cdot a]_m, [1 \cdot a]_m, [2a]_m, \ldots, [(m-1)a]_m\}.$$

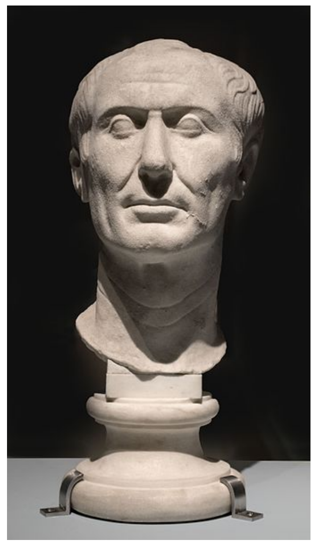
Men da må identitetselementet $[1]_m$ være med i mengden til høyre, det vil si at det må finnes et heltall $1 \leq n < m$ slik at $[1]_m = [na]_m$. Siden $[na]_m = [n]_m \cdot [a]_m$, har vi funnet at $[n]_m$ er den multiplikative inversen til $[a]_m$. □

Siden *kongruensregning* (det vil si regning i $\mathbb{Z}_m$) intuitivt handler om å “regne på vanlig måte i endelige sykler”, som i eksemplet med klokken ovenfor, er det ikke overraskende at slik regning forekommer i gamle kinesiske, indiske og arabiske kulturer, ofte knyttet til problemer som omhandler kalendre og astronomiske observasjoner, siden disse involverer gjentagende mønstre. Men vi finner også eksempler på rent matematiske problemer med kongruensregning i det kinesiske verket *Sunzi Suanjin*, av ukjent forfatter, skrevet i det tredje til femte århundret f.Kr. Den sveitsiske matematikeren Leonhard Euler (1707–1783) gjorde pionerende arbeid innenfor kongruensregning, og det moderne grunnlaget for kongruensregning ble lagt i 1801 i *Disquisitiones Arithmeticae*[2] av Carl Friedrich Gauss (1777–1855).

Koding av meldinger. I dag har kongruensregning mange anvendelser, for eksempel innenfor kryptografi og kodeteori. La oss se på et klassisk eksempel.

[2]Latin for *aritmetiske undersøkelser/diskusjoner.*

Eksempel 4.3.4 (**Caesars kode**). Ett av de tidligste eksemplene på kryptografi, eller koding av meldinger, finner vi hos den romerske keiseren Julius Caesar (100 f.Kr.–44 f.Kr.). Han kodet meldinger rett og slett ved å skifte hver bokstav tre plasser frem i alfabetet (slik at, eksempelvis A er sendt til D). For å dekode meldingen, måtte mottakeren skifte bokstavene tre plasser tilbake.



Julius Caesar, i marmorbyste datert 50–40 f.Kr., utstilt ved *Museo di antichità di Torino* (fra Wikipedia, Creative Commons Attribution 2.0 Generic license)

For å beskrive dette matematisk, la oss gå ut i fra at vi har et alfabet på m bokstaver. I Caesars tid hadde det latinske alfabetet 23 bokstaver, slik at $m = 23$ i dette tilfellet. Erstatt hver bokstav med et element i $\mathbb{Z}_m = \{[0]_m, [1]_m, \ldots, [m-1]_m\}$, i samme rekkefølge som bokstavene kommer i alfabetet (slik at, i Caesars tid, ville A representeres ved $[0]_{23}$, B med $[1]_{23}$, helt til Z, som representeres ved $[22]_{23}$). En mer generell versjon av Caesars kodingssystem som går ut på å flytte bokstavene k plasser frem i alfabetet kan beskrives ved funksjonen

$$\begin{aligned} f : \mathbb{Z}_m &\longrightarrow \mathbb{Z}_m \\ [n]_m &\mapsto [n+k]_m\,, \end{aligned}$$

slik at bokstaven representert ved $[n]_m$ blir byttet ut med bokstaven representert ved $[n+k]_m$. Denne prosessen kalles *koding*.

For å gjenkonstruere den opprinnelige meldingen, må mottakeren kjenne til k og anvende den inverse funksjonen $f^{-1} : \mathbb{Z}_m \to \mathbb{Z}_m$ på hvert symbol i den mottatte meldingen. Denne prosessen kalles *dekoding*. Tallet k kalles gjerne *nøkkelen*[3] til koden.

[3] *Key* på engelsk.

Den inverse funksjonen er selvsagt gitt ved

$$\begin{array}{rcl} f^{-1}:\mathbb{Z}_m & \longrightarrow & \mathbb{Z}_m \\ [n]_m & \mapsto & [n-k]_m \,. \end{array}$$

Som eksempel, anta at vi bruker dagens latinske alfabet på 26 bokstaver og at nøkkelen er 3 (slik at $m = 26$ og $k = 3$), og at vi mottar den kodede meldingen "PHHWBRXLQWKHSDUN". Hvis vi erstatter bokstavene med tall, får vi

$$15\ 7\ 7\ 22\ 1\ 17\ 23\ 11\ 16\ 22\ 10\ 7\ 18\ 3\ 20\ 13.$$

Anvender vi funksjonen f^{-1} på tallene får vi ut:

$$12\ 4\ 4\ 19\ 24\ 14\ 20\ 8\ 13\ 19\ 7\ 4\ 15\ 0\ 17\ 10,$$

og erstatter vi disse tallene med bokstaver, får vi den dekodede meldingen "MEETYOUINTHEPARK".

Vi kan generalisere Caesars kodingsmetode for å øke sikkerheten ved å velge en kodefunksjon på formen

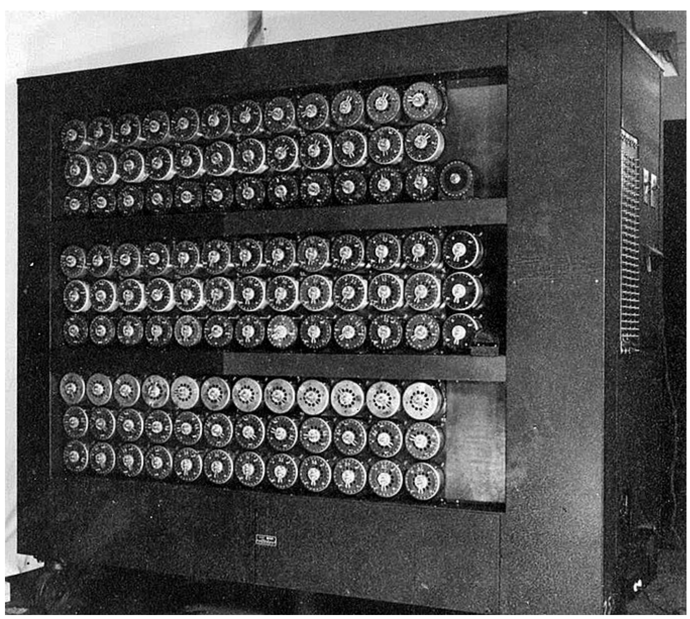
$$\begin{array}{rcl} f:\mathbb{Z}_m & \longrightarrow & \mathbb{Z}_m \\ [n]_m & \mapsto & [an+b]_m \,, \end{array}$$

hvor a og b er heltall. Men her må vi passe på; vi må velge a og b slik at funksjonen f er bijektiv, slik at meldingen kan dekodes. Ender vi opp med en ikke-bijektiv funksjon, vil jo forskjellige bokstaver sendes på samme bokstav i den kodede meldingen, slik at dekoding er umulig. Hvordan a og b må velges, skal du selv finne ut av i Oppgave 4.10(c).

I dag brukes matematikere til å utvikle og forsøke å knekke koder verden over. Et berømt eksempel er den britiske matematikeren og informatikeren Alan Turing (1912–1954), som under andre verdenskrig arbeidet ved *Government Communications Headquarters*, det britiske senteret for kryptoanalyse. I en periode var han sjef for enheten som hadde som oppgave å knekke den tyske marinekrypteringen, og innsatsen spilte en avgjørende rolle i å sikre de alliertes seier og forkorte krigen. I tillegg regnes Turing for å være faren til teoretisk informatikk og kunstig intelligens. Til tross for sine meritter, ble Turing aldri fullt anerkjent i sin levetid, siden han var homoseksuell og siden arbeidet under andre verdenskrig var hemmelighetsstemplet inntil nylig. Turing ble dømt for homoseksuelle handlinger i 1952, og valgte kjemisk kastrering i stedet for fengelsstraff. Han døde i 1954, trolig av selvmord. I 2009 kom Storbritannias statsminister Gordon Brown med en offentlig unnskyldning for måten Turing hadde blitt behandlet på, og i 2013 innvilget Dronning Elizabeth ham posthum benådning.

Alan Turing. Kilde: Wikipedia, public domain

Historien om Turing og hans innsats under andre verdenskrig har blitt (ikke helt historisk korrekt) filmatisert i *The Imitation Game* i 2014, av den norske regissøren Morten Tyldum, med Benedict Cumberbatch og Keira Knightley i hovedrollene, basert på 1983 biografien *Alan Turing: The Enigma*, av Andrew Hodges.



Maskinen benyttet av Alan Turing for å knekke koder under andre verdenskrig, kalt “bombe”. Kilde: Wikipedia, public domain

Delelighetsregler. Andre, morsomme og klassiske, anvendelser av kongruensregning, er utledningene av *delelighetsregler* for heltall. Disse reglene var nok mye mer kjent og i bruk før man hadde kalkulatorer, så det er nok sannsynlig at dine besteforeldre kjenner til disse reglene mer enn det du gjør.

Som et eksempel tar vi med regelen for delelighet på 3:

Setning 4.3.5: Delelighet på 3

Et positivt heltall er delelig på 3 hvis og bare hvis summen av dets siffre er delelig på 3.

Bevis

Husk igjen at "hvis og bare hvis" er det samme som en dobbel implikasjonspil "$\iff$".

Anta at heltallet er $N = s_n s_{n-1} \dots s_1 s_0$, skrevet i den vanlige representasjonen av heltall i 10-tallsystemet, jf. §3.2. Dette betyr at

$$N = 10^n s_n + 10^{n-1} s_{n-1} + \cdots + 10^2 s_2 + 10 s_1 + s_0.$$

Siden $10 \equiv 1 \pmod 3$ får vi, ved å bruke (69), at

$$N \equiv s_n + s_{n-1} + \cdots + s_1 + s_0 \pmod 3.$$

Dette betyr at N er delelig på 3 hvis og bare hvis $s_n + s_{n-1} + \cdots + s_1 + s_0$ er delelig på 3. □

Man kan i prinsippet utlede tilsvarende regler for delelighet for alle positive heltall, men de er mest kjent for de laveste tallene, se Oppgave 4.11.

Det norske personnummersystemet. Når noen blir født i Norge, blir vedkommende tildelt et 11-sifret tall, som kalles fødselsnummer. De seks første sifrene er fødselsdatoen. De siste fem kalles *personnummer*. De tre første sifrene i personnummeret kalles *individsifre*, hvorav det tredje angir kjønn: kvinner har partall, menn har oddetall. De første 9 sifrene er alltid tall i $\{0, \dots, 9\}$.

De to siste sifrene i personnumeret kalles *kontrollsifre* og er beregnet ut fra de foregående sifrene på følgende måte: Dersom de første ni sifrene er $x_1, x_2, \dots, x_9$, så er det tiende sifferet definert ved at det tilfredsstiller

$$x_{10} \equiv 8x_1 + 4x_2 + 5x_3 + 10x_4 + 3x_5 + 2x_6 + 7x_7 + 6x_8 + 9x_9 \pmod{11},$$

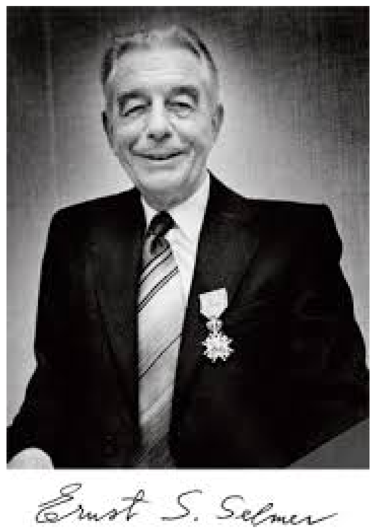
mens det ellevte sifferet er definert ved at det tilfredsstiller

$$x_{11} \equiv 6x_1 + 7x_2 + 8x_3 + 9x_4 + 4x_5 + 5x_6 + 6x_7 + 7x_8 + 8x_9 + 9x_{10} \pmod{11}.$$

Siden vi regner modulo 11, kunne vi i prinsippet få hva som helst fra og med 0 til og med 10 som kontrollsifre. For å unngå at vi får 10, passer man på å velge de tre sifrene i personnumeret slik at verken x_{10} eller x_{11} blir 10.

Det norske fødselsnummersystemet ble innført i 1964 og er utviklet av den norske matematikeren Ernst S. Selmer (1920–2006), daværende professor i matematikk ved Universitetet i Bergen. Koeffisientene foran $x_1, \dots x_9$ i systemet for kontrollsiffer er valgt slik at en i størst mulig grad skal fange opp de fleste former for skrivefeil og ombyttinger av tallene i personnummeret. For eksempel, som dere kan vise i Oppgave 4.12:

- alle enkeltfeil i et norsk personnummer kan oppdages og korrigeres selv om vi ikke vet hvor feilen er oppstått, og
- alle feil som skyldes ombytting av to siffer i et norsk personnummer kan oppdages.



ERNST SEJERSTED SELMER

For et annet eksempel på anvendelse av kongruensregning i forbindelse med kontrollsifre, se Oppgave 4.13.

4.4. Matriseringer*

La $m, n \in \mathbb{Z}^+$ og R være en kommutativ ring. (Du kan godt for enkelhets skyld la $R = \mathbb{Z}$, $\mathbb{Q}$ eller $\mathbb{R}$ i dette eksemplet.) En $(m \times n)$-*matrise*[4] er en samling av elementer fra R ordnet i m rader og n søyler, det vil si

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n-1} & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n-1} & a_{m,n} \end{bmatrix}.$$

En forkortet notasjon er $[a_{i,j}]_{1 \le i \le m, 1 \le j \le n}$ eller bare $[a_{i,j}]$. Elementene $a_{i,j}$ kalles *koeffisientene*[5] til matrisen. Mengden av alle slike matriser betegnes med $M_{m \times n}(R)$, som forenkles til $M_n(R)$ dersom $m = n$.

---

[4] *Matrix* på engelsk. Ja, det er her navnet på filmene kommer fra, siden programmering av datamaskiner bruker matriser! Ordet "matrise" kommer fra det latiske ordet *matrix*, som betyr "livmor" (ordet har samme stamme som *mater*, som betyr "mor"), og ble brukt i betydning "der hvor alt har opphav fra". I matematikk ble begrepet introdusert av den engelske matematikeren James Joseph Sylvester (1814–1897) i 1850.

[5] *Entries* på engelsk.

Vi kan definere summen av to $(m \times n)$-matriser $A = [a_{i,j}]$ og $B = [b_{i,j}]$ ved

$$A + B = [a_{i,j} + b_{i,j}].$$

Altså bare adderer vi koeffisientene i samme plass i matrisene. For eksempel er

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 5 \end{bmatrix} + \begin{bmatrix} 5 & -1 & 1 \\ 8 & 2 & -3 \end{bmatrix} = \begin{bmatrix} 6 & 1 & 0 \\ 8 & -1 & 2 \end{bmatrix}.$$

La $A = [a_{i,j}]_{1 \leq i \leq m, 1 \leq j \leq n}$ være en $(m \times n)$-matrise og $B = [a_{i,j}]_{1 \leq i \leq n, 1 \leq j \leq l}$ være en $(n \times l)$-matrise. Her er det essensielt at antallet søyler i A er likt antallet rader i B. Da kan vi definere produktet av matrisene ved:

$$A \cdot B = [c_{i,j}], \quad \text{med} \quad c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j} = a_{i,1} b_{1,j} + a_{i,2} b_{2,j} + \cdots + a_{i,n} b_{n,j}.$$

Med andre ord fremkommer koeffisientene $c_{i,j}$ i produktmatrisen $A \cdot B$ ved å multiplisere leddvis koeffisientene i den i-te rekken av A og den j-te rekken av B, og summere disse n produktene. (Ser vi på den i-te rekken av A og den j-te rekken av B som vektorer, er altså $c_{i,j}$ skalarproduktet av disse.) For eksempel er

$$\begin{aligned} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{bmatrix} &= \begin{bmatrix} 1 \cdot 7 + 2 \cdot 9 + 3 \cdot 11 & 1 \cdot 8 + 2 \cdot 10 + 3 \cdot 12 \\ 4 \cdot 7 + 5 \cdot 9 + 6 \cdot 11 & 4 \cdot 8 + 5 \cdot 10 + 6 \cdot 12 \end{bmatrix} \\ &= \begin{bmatrix} 58 & 64 \\ 139 & 154 \end{bmatrix}. \end{aligned}$$

For at begge operasjonene addisjon og multiplikasjon skal være definert, må vi altså begrense oss til tilfellet $m = n$, det vil si til mengden $M_n(R)$.

Merk at i tilfellet $m = n = 1$ får vi kun tilbake ringen R selv med sine operasjoner, så dette tilfellet er ikke interessant.

Setning 4.4.1

La $n \geq 2$. Da er $M_n(R)$ med operasjonene definert ovenfor en ikke-kommutativ ring.

Verifiseringen av alle ringaksiomene er temmelig kronglete for generell n, spesielt verifiseringen av at matrisemultiplikasjon er assosiativ (ringaksiom (M2)). Vi overlater det til Oppgave 4.16 å verifisere aksiomene i tilfellet $n = 2$. Det interessante er at produktet ikke er kommutativt: dette kan vi for eksempel se i tilfellet $n = 2$ ved å utføre følgende multiplikasjoner:

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 5 & 3 \end{bmatrix}$$

og

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 3 & 2 \end{bmatrix},$$

og det er lett å finne mange lignende eksempler for hver n.

Bakgrunnen for definisjonene på matriseoperasjonene. Definisjonene på matriseoperasjonene ovenfor kan nok virke litt underlige ved første øyekast. Det hele blir nok litt klarere hvis vi tar en titt på hvorfor matriser og deres operasjoner ble introdusert.

La oss betrakte et ligningssystem av m lineære ligninger i n ukjente variable $x_1, \dots, x_n \in R$. Dette kan vi skrive som

$$\begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & +\cdots+ & a_{1,n-1}x_{n-1} & + & a_{1,n}x_n & = & y_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & +\cdots+ & a_{2,n-1}x_{n-1} & + & a_{2,n}x_n & = & y_2 \\ \cdots & & \cdots & \cdots & \cdots & & \cdots & & \\ a_{m,1}x_1 & + & a_{m,2}x_2 & +\cdots+ & a_{m,n-1}x_{n-1} & + & a_{m,n}x_n & = & y_m \end{array},$$

med $a_{i,j}, y_i \in R$, for $1 \leq i \leq m$, $1 \leq j \leq n$ Ved hjelp av matrisemultiplikasjon, kan dette skrives som

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n-1} & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n-1} & a_{m,n} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_m \end{bmatrix}.$$

Addisjon av to matriser vil tilsvare å addere to ligningssystemer.

Matriser i forstand av "rader av koeffisienter" i ligningssystemer er funnet i kinesiske skrifter før Kristi fødsel, og ble utviklet videre av den kinesiske matematikeren Zhu Shijie (1249–1314) i boken *Siyuan yujian* fra 1303, som regnes for å være et mesterverk innen algebra. Metoden ble brakt til Europa av den italienske matematikeren Girolamo Cardano (1501–1576) i sitt verk *Ars Magna*[6] fra 1545.

Matrisemultiplikasjon ble innført av den franske matematikeren Jacques Philippe Marie Binet (1786–1856) i 1812 i studiet av såkalte *lineære avbildninger*, som vi nå skal forklare.

La $R = K$ være en kropp, for eksempel $K = \mathbb{R}$ eller $K = \mathbb{Q}$. Betrakt det kartesiske produktet

$$K^n = \underbrace{K \times \cdots \times K}_{n \text{ ganger}}.$$

Dette kan visualiseres som et "n-dimensjonalt koordinatsystem". Tilfellet $n = 1$ er (tall)linjen, tilfellet $n = 2$ er (det kartesiske) planet, tilfellet $n = 3$ er (det kartesiske) rommet, osv. Elementene i K^n betegnes med

$$\mathbf{a} = (a_1, \dots, a_n), \quad a_i \in K$$

og kalles gjerne *vektorer*. Vi merker at vi kan fremstille en vektor både som en $(1 \times n)$-matrise

$$\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$$

og som en $(n \times 1)$-matrise

$$\begin{bmatrix} a_1 \\ \cdots \\ a_n \end{bmatrix}.$$

[6] Latin for "Den store kunsten".

Vi kan addere vektorer ved å addere dem som matriser: Dersom $\mathbf{a} = (a_1, \ldots, a_n) \in K^n$ og $\mathbf{b} = (b_1, \ldots, b_n) \in K^n$, gir dette oss

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1, \ldots, a_n + b_n).$$

For enhver vektor $\mathbf{a} \in K^n$ og enhver $c \in K$ (som i denne sammenhengen gjerne kalles en *skalar*), kan vi definere *skalarmultiplikasjon* ved

$$c \cdot \mathbf{a} = c \cdot (a_1, \ldots, a_n) = (ca_1, \ldots, ca_n).$$

Ser vi på en vektor som en "pil" fra origo med en viss retning og lengde, så betyr dette at vi justerer ("skalerer") lengden på vektoren uten å endre retningen. (De to definerte operasjonene $+$ og $\cdot$ gjør K^n til et såkalt K-vektorrom, se Definisjon 4.5.5 og Eksempel 4.5.6.)

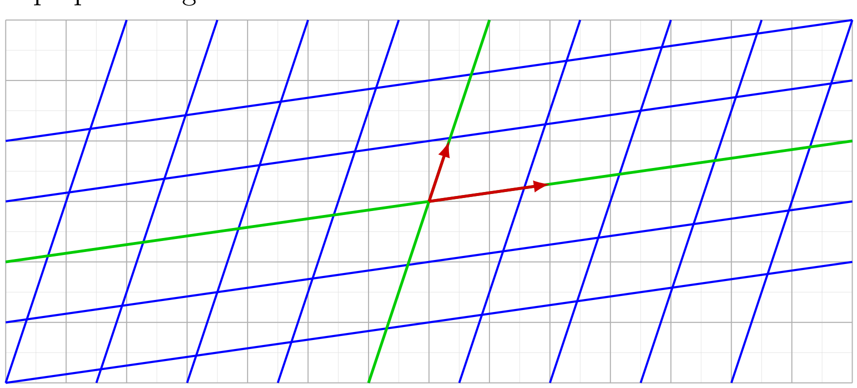
En funksjon eller en avbildning

$$T : K^n \longrightarrow K^m$$

kalles en *lineær transformasjon*[7] (eller *lineær funksjon/avbildning*) dersom

- $T(\mathbf{a} + \mathbf{b}) = T(\mathbf{a}) + T(\mathbf{b})$, for alle $\mathbf{a}, \mathbf{b} \in K^n$, og
- $T(c \cdot \mathbf{a}) = c \cdot T(\mathbf{a})$, for alle $\mathbf{a} \in K^n$, $c \in K$.

Den øverste egenskapen betyr at vi får samme resultat om de to vektorene $\mathbf{a}$ og $\mathbf{b}$ er først summert og så avbildet (som gir oss $T(\mathbf{a} + \mathbf{b})$), eller først avbildet og så summert (som gir oss $T(\mathbf{a}) + T(\mathbf{b})$). Den nederste egenskapen betyr at vi får samme resultat om vektoren $\mathbf{a}$ er først skalert med c og så avbildet (som gir oss $T(c\mathbf{a})$), eller først avbildet og så skalert med c (som gir oss $cT(\mathbf{a})$). En intuitiv måte å tenke på en linear tranformasjon er som en funksjon som "bevarer lineære objekter", det vil si som sender parallellogrammer på parallellogrammer. Følgende figur visualiserer hva som skjer med planet $\mathbb{R}^2$ etter en lineær transformasjon $T : \mathbb{R}^2 \to \mathbb{R}^2$: origo forblir uberørt, mens koordinataksene avbildes på rette linjer som ikke nødvendigvis skjærer hverandre vinkelrett og strekkes i forskjellig skala; enhetsrutene avbildes på parallellogrammer.



EFFEKTEN AV EN LINEÆR TRANSFORMASJON $\mathbb{R}^2 \to \mathbb{R}^2$

[7] *Linear transformation* på engelsk.

Et eksempel på en lineær transformasjon er projeksjonen fra rommet til kameraskjermen når man tar et bilde, hvilket blant annet forklarer hvorfor studiet av lineære transformasjoner har mange anvendelser.

Det er lett å se at en lineær tranformasjon $K \to K$ må være på formen $x \mapsto cx$ for en konstant $c \in K$. Det er også relativt enkelt å se at en lineær transformasjon $K^2 \to K^2$ må ha formen

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto \begin{bmatrix} a_{1,1}x_1 + a_{1,2}x_2 \\ a_{2,1}x_1 + a_{2,2}x_2 \end{bmatrix},$$

for konstanter $a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2} \in K$. Helt generelt kan man vise at en lineær transformasjon $K^n \to K^m$ må ha formen

$$\begin{bmatrix} x_1 \\ \cdots \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} a_{1,1}x_1 + \cdots + a_{1,n}x_n \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n \\ \cdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n \end{bmatrix},$$

der alle $a_{i,j} \in K$. Høyresiden kan vi skrive som matriseproduktet

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n-1} & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n-1} & a_{m,n} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}.$$

Altså kan enhver lineær transformasjon uttrykkes som

$$\begin{array}{rccc} T: & K^n & \longrightarrow & K^m \\ & \begin{bmatrix} x_1 \\ \cdots \\ x_n \end{bmatrix} & \mapsto & A \cdot \begin{bmatrix} x_1 \\ \cdots \\ x_n \end{bmatrix}, \end{array}$$

hvor A er en $(m \times n)$-matrise over K. Motsatt definerer enhver $(m \times n)$-matrise over K en lineær transformasjon $K^n \to K^m$.

Hvis $T : K^n \to K^m$ og $L : K^m \to K^l$ er to lineære transformasjoner gitt ved en $(m \times n)$-matrise A og en $(l \times m)$-matrise B, henholdsvis, da kan man vise at den sammensatte lineære transformasjonen $L \circ T : K^n \to K^l$ er definert ved $(l \times n)$-matrisen BA. Dette gir en fin geometrisk tolkning av matrisemultiplikasjon og forklarer grunnen til at den er definert slik.

Studiet av lineære transformasjoner og matriser kalles *lineær algebra* og er innholdet i kurset *MAT121–Lineær algebra*.

4.5. Andre algebraiske strukturer*

Det finnes mange andre algebraiske strukturer i tillegg til ringer og kropper. Vi skal se på et par av dem i denne seksjonen.

Grupper. Blant de enkleste algebraiske struktuene er de som kalles *grupper*, som er mengder med kun én operasjon som oppfyller visse aksiomer:

Definisjon 4.5.1: Gruppe

En *gruppe*[a] er en mengde G sammen med en funksjon

$$\begin{array}{rcl} *: G \times G & \longrightarrow & G \\ (a,b) & \mapsto & a*b \end{array}$$

som oppfyller følgende aksiomer:

(G1) $(a*b)*c = a*(b*c)$ for alle $a,b,c \in G$ (det vil si: operasjonen $*$ er *assosiativ*).

(G2) Det finnes et element $e \in G$ (kalt *identitet*) slik at $e*a = a*e$ for alle $a \in G$.

(G3) For enhver $a \in G$ finnes et element $a^{-1} \in G$ (kalt *inversen til* a), som oppfyller $a*a^{-1} = a^{-1}*a = e$.

Gruppen kalles *kommutativ* eller *abelsk*[b] dersom operasjonen $*$ er *kommutativ*, det vil si, $a*b = b*a$ for alle $a,b \in G$.

[a] *Group* på engelsk.
[b] *Abelian* på engelsk.

Navnet *abelsk* kommer selvsagt fra den norske matematikeren Niels Henrik Abel (1802–1829). Ofte bruker man symbolet $+$ for gruppeoperasjonen $*$ dersom gruppen er abelsk. En annen vanlig betegnelse på gruppeoperasjonen $*$ er $\cdot$. Siden en gruppe G kommer med sin operasjon $*$, er det vanlig å si at $(G,*)$ er en gruppe.

Eksempel 4.5.2. (i) Enhver kropp sammen med en av sine operasjoner $+$ og $\cdot$ er en abelsk gruppe.

(ii) Enhver ring med sin operasjon $+$ er en abelsk gruppe.

(iii) Mengden av $(n \times n)$-matriser med matrisemultiplikasjon er en ikke-abelsk gruppe, for $n \geq 2$.

Eksempel 4.5.3. La A være en ikke-tom mengde. En bijektiv funksjon $A \to A$ kalles en *permutasjon (av A)*[8]. Mengden S_A av alle permutasjoner av A, sammen med operasjonen *sammensetning av funksjoner* $\circ$ (Definisjon 2.2.12), danner en gruppe. Identitetselementet er identitetsfunksjonen $A \to A$ gitt ved $a \mapsto a$. Vi overlater til Oppgave 4.18(a) å verifisere dette.

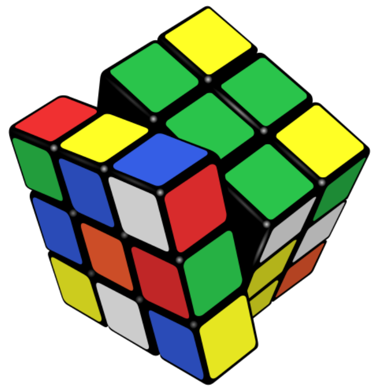
I denne forbindelsen kalles sammensetningen av funksjoner for *permutasjonsmultiplikasjon*. Gruppen S_A kalles *permutasjonsgruppen*[9] *til* A.

Hvis $A = \{1, \ldots, n\}$, det vil si A er en endelig mengde bestående av n elementer, da betegnes S_A med S_n eller $\mathrm{Sym}(n)$ og kalles *symmetrigruppen*

[8] *Permutation* på engelsk.
[9] *Permutation group* på engelsk.

til n elementer. Dette er en endelig gruppe med $n!$ elementer (jf. Oppgave 4.18(b)). Vi overlater til Oppgave 4.18(e)) å finne ut av for hvilke n gruppen S_n er abelsk.

Det finnes også eksempler på permutasjonsgrupper som består av permutasjoner av en mengde under visse regler. Et berømt slikt eksempel er permutasjonsgruppen til *Rubiks kube*.



Rubiks kube (Kilde: Wikipedia, Creative Commons Attribution-Share Alike 3.0 Unported license)

Eksempel 4.5.4. Grupper kan brukes til å uttrykke symmetrier til en geometrisk figur, slik som for eksempel et molekyl. Derfor har gruppeteori anvendelser i for ekempel *standardmodellen i partikkelfysikk*, *spesiell relativitet* og *molekylær kjemi*. *Symmetrigruppen* til et geometrisk objekt består av alle transformasjoner av objektet (slik som rotasjoner og speilinger) som bevarer objektet, sammen med operasjonen sammensetning av transformasjoner.

Som et første enkelt eksempel kan vi betrakte bokstaven A tegnet inn i planet, hvor den står på x-aksen mens y-aksen går gjennom toppunktet til bokstaven. Bokstaven er speilsymmetrisk om y-aksen, men ikke om
x-aksen. Transformasjonen s som speiler bokstaven om y-aksen er derfor en symmetri til bokstaven, mens den tilsvarende speilingen om x-aksen er det ikke. I tillegg er også identitetstransformasjonen e som bevarer bokstaven i sin opprinnelige stilling også en symmetri. Det finnes ingen flere symmetrier.

Betegner vi sammensetningen av transformasjonene som $*$, og skriver den første transformasjonen vi utfører til høyre, betyr for eksempel $e*s$ transformasjonen som består først av en speiling om y-aksen og deretter identitetstransformasjonen. Da er $e*e=e$ og $e*s=s*e=s$, som viser at e ikke overraskende er identitetselementet i symmetrigruppen. Dessuten, siden to speilinger om y-aksen bringer bokstaven A tilbake til sin opprinnelige stilling, har vi at $s*s=e$. Vi har dermed regnet ut hele multiplikasjonstabellen til gruppen og ser at den er abelsk. Kaller vi den for G, ser vi at det finnes en bijeksjon

$$\begin{array}{rcl} \varphi : G & \longrightarrow & \mathbb{Z}_2 \\ e & \mapsto & [0]_2 \\ s & \mapsto & [1]_2 \end{array}$$

(hvor φ er den greske bokstaven "fi"). Ser vi på $\mathbb{Z}_2$ som en gruppe med operasjonen $+$, ser vi at φ bevarer gruppeoperasjonene på den måten at

$$\begin{array}{ll} \varphi(e * s) & = \varphi(s) = [1]_2 = [0]_2 + [1]_2 = \quad \varphi(e) + \varphi(s) \\ \varphi(e * e) & = \varphi(e) = [0]_2 = [0]_2 + [0]_2 = \quad \varphi(e) + \varphi(e) \\ \varphi(s * s) & = \varphi(e) = [0]_2 = [1]_2 + [1]_2 = \quad \varphi(s) + \varphi(s) \end{array}$$

Vi sier derfor at gruppen G er *isomorf* til $\mathbb{Z}_2$ og kaller φ en *gruppeisomorfi*. La oss for ordens skyld skrive ned definisjonen:

To grupper G_1 og G_2, med operasjoner $*_1$ og $*_2$, henholdsvis, sies å være isomorfe[10] *dersom det finnes en bijeksjon $\varphi : G_1 \to G_2$ slik at*

$$\varphi(x *_1 y) = \varphi(x) *_2 \varphi(y).$$

Avbildningen φ kalles i dette tilfellet en *(gruppe)isomorfi*[11] Dette betyr i praksis at vi kan jobbe med de to gruppene som om de var like, siden objektene er i én-til-én korrespondanse på en slik måte at operasjonene forblir de samme. Eneste forskjellen, for å si det slik, er at objektene har forskjellige "navn".

La oss nå betrakte symmetriene til bokstaven H. Denne er invariant både under speiling om x-aksen og om y-aksen. La oss kalle disse transformasjonene s_x og s_y, henholdsvis. Dessuten er H selvsagt også invariant under identitetstransformasjonen e. Vi har dessuten én symmetri til, nemlig rotasjon av bokstaven 180° om en tenkt z-akse vinkelrett på xy-planet. La oss kalle denne transformasjonen r. Betegner vi igjen sammensetningen av transformasjonene som $*$, er det lett å se følgende:

- Dersom samme transformasjon utføres to ganger etter hverandre, kanselleres effekten og bokstaven kommer tilbake til sin opprinnelige posisjon. Dette betyr at $e * e = s_x * s_x = s_y * s_y = r * r = e$.
- Dersom de to speilingene s_x og s_y utføres etter hverandre, i hvilken som helst rekkefølge, har vi rotert bokstaven 180°. Altså har vi $s_x * s_y = s_y * s_x = r$.
- En speiling og en rotasjon etter hverandre tilsvarer den gjenværende speilingen, det vil si $s_x * r = r * s_x = s_y$ og $s_y * r = r * s_y = s_x$.

Det er lett å verifisere nå at mengden $\{e, s_x, s_y, t\}$ oppfyller aksiomene (G1)-(G3) til en gruppe. Vi har faktisk funnet hele multiplikasjonstabellen til symmetrigruppen til bokstaven H. Denne kalles *den dihedrale gruppen med fire elementer* eller *Kleins Vierergruppe*, etter den tyske matematikeren Felix

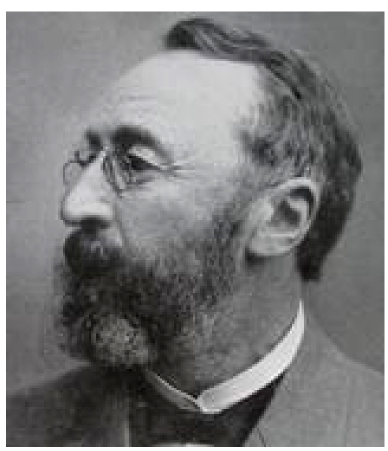
[10] *Isomorphic* på engelsk.
[11] *Isomorphism (of groups)* på engelsk.

Klein (1849–1925), og betegnes med D_2, V eller K_4. Multiplikasjonstabellen viser at gruppen er abelsk.

Regulære polygoner har mange symmetrier som kan uttrykkes ved større dihedrale grupper. Se Oppgave 4.20 for et slikt eksempel.

I tillegg til Abel, så har også de norske matematikerne Sophus Lie (1842–1899) og Peter Ludwig Sylow (1832–1918) gjort fundamentale bidrag til gruppeteorien. Spesielt har Sophus Lies livsverk en viktig plass i moderne matematikk, hvor begrepene *Lie-grupper* og *Lie-algebraer* oppkalt etter ham spiller viktige roller i mange grener av både teoretisk og anvendt matematikk, og ikke minst fysikk.

Sophus Lie (Kilde: Nasjonalbiblioteket)



Peter Ludwig Sylow (Kilde: Wikipedia, public domain)

Vektorrom. Et annet veldig viktig eksempel på en algebraisk struktur er *vektorrom*, som dere vil møte på i kurset *MAT121–Lineær algebra.* Dette er grovt sagt en gruppe sammen med en tilleggsoperasjon som er *skalarmultiplikasjon* med elementer fra en kropp:

Definisjon 4.5.5: Vektorrom

La K være en kropp. Et *vektorrom over* K eller K*-vektorrom*[a] er en mengde V sammen med to funksjoner

$$+ : V \times V \longrightarrow V$$

(kalt *vektoraddisjon*) og

$$\cdot : K \times V \longrightarrow V$$

(kalt *skalarmultiplikasjon*) som oppfyller følgende aksiomer:

(VA1) Vektoraddisjon er kommutativ: $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$ for alle $\mathbf{x}, \mathbf{y} \in V$;
(VA2) Vektoraddisjon er assosiativ: $(\mathbf{x}+\mathbf{y})+\mathbf{z} = \mathbf{x}+(\mathbf{y}+\mathbf{z})$ for alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$;
(VA3) V inneholder et element $\mathbf{0}$ (kalt *nullvektor*) som oppfyller $\mathbf{0}+\mathbf{x} = \mathbf{x}$ for alle $\mathbf{x} \in V$;
(VA4) Til ethvert element $\mathbf{x} \in V$ finnes et element $-\mathbf{x} \in V$ slik at $\mathbf{x}+(-\mathbf{x}) = \mathbf{0}$;
(SM1) $a \cdot b\mathbf{x} = (ab) \cdot \mathbf{x}$ for alle $a, b \in K$, $\mathbf{x} \in V$;
(SM2) $1 \cdot \mathbf{x} = \mathbf{x}$ for alle $\mathbf{x} \in V$ (hvor 1 er den multiplikative identiteten i K);
(D1) $(a+b) \cdot \mathbf{x} = a \cdot \mathbf{x} + b \cdot \mathbf{x}$ for alle $a, b \in K$, $\mathbf{x} \in V$;
(D2) $a \cdot (\mathbf{x}+\mathbf{y}) = a \cdot \mathbf{x} + b \cdot \mathbf{y}$ for alle $a \in K$, $\mathbf{x}, \mathbf{y} \in V$.

Elementene i V kalles *vektorer*, mens elementene i K kalles *skalarer*.

[a] K*-vector space* på engelsk.

Merk at aksiomene (VA1)-(VA4) sier at $(V, +)$ er en abelsk gruppe. Aksiomene (SM1)-(SM2) gir naturlige egenskaper som skalarmultiplikasjon skal oppfylle, sett i lys av kroppsoperasjonen multiplikasjon, mens aksiomene (D1)–(D2) gir naturlige distributive lover.

Standardeksemplet på et vektorrom er kjent fra skolen:

EKSEMPEL 4.5.6. Betrakt det kartesiske produktet $\mathbb{R}^n$. For to elementer $\mathbf{x} = (x_1, \ldots, x_n)$ og $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ og $c \in \mathbb{R}$ definerer vi:

- $\mathbf{x}+\mathbf{y} = (x_1+y_1, \ldots, x_n+y_n)$ (vektoraddisjon);
- $c\mathbf{x} = (cx_1, \ldots, cx_n)$ (skalarmultiplikasjon).

Det er lett å se at dette gjør $\mathbb{R}^n$ til et $\mathbb{R}$-vektorrom. Nullvektoren er selvsagt $(0, \ldots, 0)$, mens $-(x_1, \ldots, x_n) = (-x_1, \ldots, -x_n)$.

Mer generelt viser samme konstruksjon med de opplagte operasjonene at det kartesiske produktet K^n er et K-vektorrom for alle $n \in \mathbb{Z}^+$, når K er en kropp.

Andre kjente eksempler som vi allerede har sett er:

EKSEMPEL 4.5.7. La K være en kropp og betrakt mengden $\mathcal{F}(X,K)$ av funksjoner $X \to K$ med operasjonene $+$ og $\cdot$ definert på den opplagte måten som

$$\begin{array}{rcl} f+g : X & \longrightarrow & K \\ x & \mapsto & f(x)+g(x) \end{array}$$

for $f, g \in V$ og

$$\begin{array}{rcl} c \cdot f : X & \longrightarrow & K \\ x & \mapsto & cf(x). \end{array}$$

for $c \in K$, $f \in V$. Da er $\mathcal{F}(X,K)$ et K-vektorrom.

EKSEMPEL 4.5.8. La K være en kropp og betrakt polynomringen $K[x]$ med de vanlige operasjonene $+$ og $\cdot$ definert ovenfor. Da induserer $\cdot$ også en skalarmultiplikasjon

$$\begin{array}{rcl} K \times K[x] & \longrightarrow & K[x] \\ (c,P) & \mapsto & cP, \end{array}$$

og dette gjør $K[x]$ til et K-vektorrom.

For andre eksempler, se Oppgavene 4.17, 8.8, 10.64 og 10.70.

Oppgaver

Oppgaver til §4.1–4.2

OPPGAVE 4.1. Bevis Følgesetning 4.1.10.

OPPGAVE 4.2. Bevis Setning 4.2.1.

OPPGAVE 4.3. Hva kan du si om $\mathcal{F}(X,R)$ hvis X består av kun ett element?

Oppgaver til §4.3

OPPGAVE 4.4. Verifisér (69).

OPPGAVE 4.5. Vi så i beviset for Setning 4.3.3 at $\mathbb{Z}_m$ ikke oppfyller egenskapen (x) i Setning 3.6.3. Hvilke andre egenskaper i Setning 3.6.3 er ikke oppfylt av $\mathbb{Z}_m$?

Oppgave 4.6. Avgjør om det er riktig at $175 \equiv 22 \pmod{17}$.

Oppgave 4.7. Bevis Setning 4.3.1.

Oppgave 4.8. Skriv ned alle elementer og alle resultatetene av addisjon og multiplikasjon i $\mathbb{Z}_m$ for $m \leq 4$. Avgjør hvilke av disse er kropper, uten å bruke Setning 4.3.3.

Oppgave 4.9. Hva er det minste antallet elementer en kropp kan ha? (Hint: Oppgave 4.8 kan være til hjelp.)

Oppgave 4.10. Vi viser til Eksempel 4.3.4.

(a) En melding har blitt kryptert ved å bruke funksjonen $f : \mathbb{Z}_{26} \to \mathbb{Z}_{26}$ definert ved $f([x]_{26}) = [x+5]_{26}$. Hvis den kodede meldingen er $JCFHY$, hva er den dekodede meldingen?

(b) Forklar hvorfor $g([x]_{26}) = [2x+3]_{26}$ ikke ville vært en god kodefunksjon.

(c) Finn betingelsene på de tre heltallene a, b, m for at funksjonen $f : \mathbb{Z}_m \to \mathbb{Z}_m$ definert ved $f([x]_m) = [ax+b]_m$ skal fungere som kodefunksjon.

Oppgave 4.11. Utled regler for delelighet for tallene 2, 5, 7 og 11 som for 3 i Setning 4.3.5. (Tilfellene 2 og 5 er nærmest opplagte; tilfellet 7 er det vanskeligste.)

Oppgave 4.12. (a) Vis at alle enkeltfeil i et norsk personnummer kan oppdages og korrigeres selv om vi ikke vet hvor feilen er oppstått.

(b) Vis at alle feil som skyldes ombytting av to siffer i et norsk personnummer kan oppdages.

Oppgave 4.13. Et ISBN-nummer (International Standard Book Number) er et 10-sifret positivt heltall $s_1 \dots s_{10}$ slik at $\sum_{i=1}^{10} is_i \equiv 0 \pmod{11}$. (Denne summen kalles gjerne *kontrollsummen*.)

Vis at dersom vi bytter plass på to sifre i et ISBN-nummer, da vil det resulterende nummeret ikke være et gyldig ISBN-nummer. (Hint: dersom vi bytter plass på sifrene s_i og s_j med $s_i \neq s_j$, er differansen i kontrollsummen av det nye og det gamle tallet lik $(is_j + js_i) - (is_i + js_j) \pmod{11}$.)

(Dette gir derfor en enkel måte å finne enkle feil i slikt et tall.)

Oppgave 4.14. To elementer a og b i en ring R sies å være *nulldivisorer*[12] dersom $a \neq 0$, $b \neq 0$, men $a \cdot b = 0$. Et eksempel vi har sett på side 140 er $[2]_6$ og $[3]_6$ i $\mathbb{Z}_6$, siden $[2]_6 \neq [0]_6$ og $[3]_6 \neq [0]_6$ og $[2]_6 \cdot [3]_6 = [6]_6 = [0]_6$.

En kommutativ ring uten nulldivisorer kalles et *integritesområde*[13].

Vis at ethvert integritetsområde som inneholder kun endelig mange elementer er en kropp. (Hint: mim etter beviset for Setning 4.3.3.)

Oppgave 4.15. Prosedyren med å utvide $\mathbb{Z}$ til $\mathbb{Q}$ kan generaliseres: La R være et integritetsområde (se Oppgave 4.14 for definisjonen).

[12] *Zero divisors* på engelsk.

[13] *Integral domain* på engelsk.

(a) Vis at
$$(a,b) \sim (c,d) \iff ad - bc = 0.$$
definerer en ekvivalensrelasjon på $R \times R \setminus \{0\}$.

(b) Betegn ekvivalensklassen til (a,b) med $\frac{a}{b}$ og la $K(R)$ være mengden av alle ekvivalensklassene (jf. Definisjon 2.4.11). Vis at operasjonene
- $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$
- $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$.

er veldefinert på $K(R)$, det vil si uavhengige av valg av representanter for ekvivalensklassene.

(c) Vis at $K(R)$ med de definerte operasjonene er en kropp som inneholder R og at operasjonene i R er kompatible med operasjonene i $K(R)$.

(d) Vis at $K(R)$ også kan defineres som den minste kroppen som inneholder R (med kompatible operasjoner).

Kroppen $K(R)$ kalles *kvotientkroppen* til R. Altså er $\mathbb{Q}$ kvotientkroppen til $\mathbb{Z}$.

Oppgaver til §4.4–§4.5

OPPGAVE 4.16. La R være en ring og betrakt mengden $M_2(R)$ av (2×2)-matriser med operasjonene $+$ og $\cdot$ som i Setning 4.4.1.

(a) Bevis Setning 4.4.1 i tilfellet $m = 2$, det vil si verifisér at $M_2(R)$ med disse operasjonene er en ikke-kommutativ ring.

(b) Finn et par andre eksempler som viser at matriseproduktet ikke er kommutativ.

(c) La nå $R = K$ være en kropp. La
$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(K).$$
Finn en nødvendig og tilstrekkelig betingelse på koeffisientene a, b, c og d for at A skal ha en multiplikativ invers A^{-1} og finn A^{-1} i disse tilfellene.

OPPGAVE 4.17. La K være en kropp. Vis at det finnes et K-vektorrom med bare ett element.

OPPGAVE 4.18. La A være en ikke-tom mengde. Vi viser til Definisjon 4.5.1 for definisjonen av en gruppe og til Eksempel 4.5.3 for definisjonen av mengden S_A av alle permutasjoner av A.

(a) Vis at S_A, med operasjonen *sammensetning av funksjoner* $\circ$ (Definisjon 2.2.12), danner en gruppe, med identitetselement lik identitetsfunksjonen $A \to A$ gitt ved $a \mapsto a$.

(b) Vis at dersom A er en endelig mengde med n elementer, da har S_A, som vi gjerne betegner med S_n, nøyaktig $n!$ elementer.

For et element $j : \{1, \ldots, n\} \to \{1, \ldots, n\}$ i S_n brukes den kompakte notasjonen

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ j(1) & j(2) & j(3) & \cdots & j(n) \end{pmatrix}.$$

For eksempel betegnes permutasjonen

$$\begin{aligned} \{1,2,3,4\} &\longrightarrow \{1,2,3,4\} \\ 1 &\mapsto 3 \\ 2 &\mapsto 2 \\ 3 &\mapsto 4 \\ 4 &\mapsto 1 \end{aligned}$$

med

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}.$$

(c) List opp alle elementene i S_2 og angi multiplikasjonstabellen. Er S_2 abelsk?
(d) List opp alle elementene i S_3 og angi multiplikasjonstabellen. Er S_3 abelsk?
(e) For hvilke n er S_n abelsk?

Oppgave 4.19. Er den dihedrale gruppen/Kleins Vierergruppe D_2 fra Eksempel 4.5.4 isomorf til $\mathbb{Z}_4$ (med gruppeoperasjon addisjon modulo 4)?

Oppgave 4.20. I denne oppgaven skal vi studere symmetrigruppen til en likesidet trekant.

En likesidet trekant forblir uforandret hvis den

- roteres $120°$, $240°$ eller $360° = 0°$ mot urvisesen om en akse vinkelrett på trekantens plan;
- speiles om midtnormalene til hver side.

Kall de tre rotasjonene r_1, r_2, r_0, henholdsvis, og de tre speilingene s_1, s_2, s_3. Som i Eksempel 4.5.4 betegner vi sammensetningen av transformasjonene med $*$, med den første operasjonen som utføres skrevet til høyre.

Vis at mengden $\{r_0, r_1, r_2, s_1, s_2, s_3\}$ sammen med operasjonen $*$ er en gruppe og finn hele multiplikasjonstabellen. Er gruppen abelsk?

Gruppen kalles *den dihedrale gruppen med seks elementer* og betegnes med D_3.

KAPITTEL 5

Velordningsprinsippet og induksjon

Vi skal i dette kapitlet se på et par veldig viktige teknikker som man ofte møter på i mange grener av matematikken og i andre fagfelt. Prinsippene er omformuleringer av induksjonsaksiomet (N5) i Peano-aksiomene på side 87 og kalles *velordningsprinsippet*[1] og *induksjonsprinsippet*[2]. Den ene varianten av induksjonsprinsippet er også pensum i MAT111.

De eldste sporene av induksjonslignende bevis kan man finne i Euklids bevis for at det finnes uendelig mange primtall (Teorem 3.7.10) og i metoden til den indiske matematikeren Bhāskara (1114–1185) for å løse ligninger på formen $ax^2 + bx + c = y$, kalt den *sykliske metoden.* I den velkjente boken om algebra *al-fakhri fi al-jabr wa al-muqabala* fra ca. år 1000, brukte den persiske matematikeren Al-Karaji (ca. 953–ca. 1029) induksjon implisitt for å beregne summer av heltall og bevise binomialteoremet (Teorem 3.8.7). Et lignende eksempel på implisitt bruk av induksjon finner vi i boken *Arithmeticorum libri duo* (1575) av den italienske matematikeren og astronomen Francesco Maurolico (1494–1575), hvor han viste at summen av alle oddetall opp til n er n^2 (Oppgave 5.6).

Den første rigorøse bruken av induksjon pleier man å tilegne den fransk-jødiske matematikeren, filosofen, legen og astronomen Levi ben Gershon (1288–1344) i verket *Maaseh Hoshev* fra 1321. Den første eksplisitte formuleringen av induksjonsprinsippet ble gitt av den franske matematikeren og fysikeren Blaise Pascal (1623–1662) i *Traité du triangle arithmétique* (1665), hvor han blant annet studerer binomialkoeffisienter (se Definisjon 3.8.3).

Vi innleder kapitlet med velordningsprinsippet, før vi går løs på de to variantene av induksjonsprinsippet (§5.2-5.3).

5.1. Velordningsprinsippet

En viktig egenskap ved de naturlige tallene som man kan utlede fra Peano-aksiomene på side 3.1 er følgende (jf. Oppgave 5.1(i)):

Velordningsprinsippet

Enhver ikke-tom delmengde A av $\mathbb{N}$ har et minste element, det vil si et element $x \in A$ som oppfyller $x \leq a$ for alle $a \in A$.

---

[1] *Principle of well-ordering* på engelsk.

[2] *Principle of induction* på engelsk.

Dette prinsippet kan brukes til å bevise påstander, som vist i følgende eksempel.

EKSEMPEL 5.1.2. I en *round-robin-turnering* møter hver spiller i turneringen enhver annen deltager nøyaktig én gang og en kamp kan ikke ende uavgjort. Vi sier at spillerne $S_1, S_2, \ldots, S_M$ former en *sykel* av lengde M, hvis S_1 slår S_2, S_2 slår S_3, …, S_{M-1} slår S_M og S_M slår S_1. Her er M et heltall slik at $M \geq 3$. Bruk velordningsprinsippet til å vise at dersom det finnes en sykel av lengde $N \geq 3$, da finnes også en sykel av lengde 3.

LØSNING. Vi argumenterer ved selvmotsigelse (jf. §1.3): Vi antar at det *ikke* finnes en sykel av lengde 3 og viser at vi kommer til en motsigelse. Siden vi får opplyst at det eksisterer en sykel av lengde M, så vil mengden

$$A = \{n \in \mathbb{N} \mid \text{det finnes en sykel av lengde } n\}$$

være ikke-tom, fordi $M \in A$. Ved velordningsprinsippet vil A inneholde et minste element m. Det følger at det finnes en sykel $S_1, S_2, \ldots, S_m$ av lengde m og ingen sykler av lengde mindre enn m finnes. Per antagelse er $m \geq 4$.

Betrakt S_1, S_2, S_3. Siden vi antar at det ikke finnes en sykel av lengde 3, må kampen mellom S_1 og S_3 ha endt med at S_1 slår S_3. Dermed kan vi ta bort S_2 fra sykelen og få en sykel $S_1, S_3, \ldots, S_m$ som har lengde $m-1$. Men dette motsier at den minste sykelen har lengde m. $\square$

Som et annet eksempel på anvendelse av *velordningsprinsippet* vil vi nå vise *divisjonsalgoritmen* fra §3.7. Vi gjentar resultatet:

Setning 3.7.3: Divisjonsalgoritmen

Hvis $a \in \mathbb{Z}$ og $d \in \mathbb{Z}^+$, da finnes entydige $q, r \in \mathbb{Z}$ slik at $0 \leq r < d$ og $a = dq + r$.

BEVIS

Vi viser kun eksistensen av q og r i setningen. Beviset for entydigheten overlates til Oppgave 5.2.

Vi betrakter først tilfellet $a \geq 0$. La $d \in \mathbb{N} \setminus \{0\}$. La S være mengden av alle naturlige tall n på formen $a - dq$ for en $q \in \mathbb{N}$. Skrevet formelt har vi altså

$$S = \{n \in \mathbb{N} \mid n = a - dq \text{ for en } q \in \mathbb{N}\}.$$

Siden $a = a - d \cdot 0$, har vi $a \in S$, slik at $S \neq \emptyset$. Ved velordningsprinsippet har S et minste element $r = a - dq_0$ for en $q_0 \in \mathbb{N}$. Hvis vi nå viser at $r < d$, har vi funnet $q = q_0$ og r som oppfyller betingelsene i setningen (bortsett fra entydigheten).

Vi antar dermed at $r \geq d$ og utleder en motsigelse. (Dette betyr at vi argumenterer ved selvmotsigelse, jf. §1.3.) Hvis $r \geq d$, er

$$0 \leq r - d = a - dq_0 - d = a - d(q_0 + 1),$$

som viser at $r - d \in S$. Men siden $r - d < r$, motsier dette at r er det minste elementet i S.

Vi betrakter så tilfellet $a < 0$. Da er $-a > 0$ og vi kan bruke det vi nettopp har vist til å skrive $-a = dq + r$ for heltall q, r slik at $0 \leq r < d$. Da er $a = d(-q) - r$. Hvis $r = 0$, har vi skrevet a på den ønskede formen. Hvis $r > 0$, skriver vi $a = -d(q+1) + (d-r)$. Siden $0 < d - r < d$, har vi skrevet a på den ønskede formen. □

5.2. Matematisk induksjon

Matematisk induksjon er en veldig viktig teknikk som man ofte møter på i mange grener av matematikken og i andre fagfelt. Prinsippet er en omformulering av induksjonsaksiomet (N5) i Peano-aksiomene.

Her er en oppsummering av teknikken:

Induksjonsprinsippet (versjon I)

La $P(n)$ være et utsagn som gir mening for ethvert heltall $n \geq n_0$, der n_0 er et fiksert heltall.

For å vise at $P(n)$ er sant for alle heltall $n \geq n_0$ er det nok å vise følgende to fakta:

(I) Utsagnet $P(n)$ er sant når $n = n_0$ (det vil si $P(n_0)$ er et sant utsagn).

(II) Hvis utsagnet $P(n)$ er sant når $n = k$, der k er et fast, men uspesifisert heltall slik at $k \geq n_0$, så er også $P(n)$ sant når $n = k + 1$ (det vil si at hvis $P(k)$ er sant, så er også $P(k+1)$ sant).

MERKNAD 5.2.2. I steg (II) tar vi ikke stilling til om utsagnet $P(k)$ virkelig er sant: vi bare viser at *om* $P(k)$ er sant, så er $P(k+1)$ sant også. Antagelsen om at $P(k)$ er sann kalles *induksjonshypotesen*[3] og hele trinn (II) kalles *induksjonstrinnet*[4].

Grunnen til at dette virker er nettopp induksjonsaksiomet (N5):

Setning 5.2.3

Induksjonsaksiomet (N5) er ekvivalent med induksjonsprinsippet (versjon I).

[3] *Induction hypothesis* på engelsk.

[4] *Induction step* på engelsk.

Bevis
Vi viser at induksjonsaksiomet (N5) impliserer induksjonsprinsippet (versjon I) og overlater det til Oppgave 5.4 å vise at induksjonsprinsippet (versjon I) impliserer induksjonsaksiomet (N5).

Betrakt delmengden

$$M = \{n \in \mathbb{N} \mid P(n_0 + n) \text{ er sant}\} \subset \mathbb{N}.$$

Antagelsen (I) i induksjonsprinsippet gir at $0 \in S$.

Anta at $n \in M$ for en $n \in \mathbb{N}$, det vil si at $P(n_0 + n)$ er sant for $n_0+n \geq n_0$. Da gir antagelsen (II) i induksjonsprinsippet at $P(n_0+n+1)$ er sant, det vil si at $n + 1 \in M$. Husk at vi viste i (42) at $n + 1 = S(n)$, det vil si *etterfølgeren* til n. Induksjonsaksiomet (N5) gir da at $M = \mathbb{N}$, som betyr at $P(n_0 + n)$ er sant for alle $n \geq 0$, ekvivalent at $P(n)$ er sant for alle $n \geq n_0$. □

Eksempler. Vi skal nå se på en del eksempler på bruk av induksjon, av forskjellige typer. Forsøk deg gjerne på problemene før du leser gjennom løsningene.

Eksempel 5.2.4. Vis at for ethvert positivt heltall n, er summen av alle heltallene fra 1 til n lik $n(n + 1)/2$.

Bevis ved induksjon. Vi skal vise at utsagnet

$$P(n): \quad 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2} \tag{70}$$

er riktig for alle heltall $n \geq 1$. Med summetegn (jf. side 124) kan vi skrive (70) som

$$P(n): \quad \sum_{j=1}^{n} j = \frac{n(n+1)}{2}.$$

Først observerer vi at $P(1)$ er riktig: Ved å sette $n = 1$ inn i (70), ser vi nemlig at venstresiden da er lik 1 og høyresiden er lik $\frac{1(1+1)}{2} = \frac{2}{2} = 1$. (Dette er trinn I.)

La oss nå *anta* at utsagnet $P(n)$ holder for $n = k$, der $k \geq 1$, det vil si vi antar at $P(k)$ holder, med andre ord at

$$1 + 2 + 3 + \cdots + k = \frac{k(k+1)}{2}. \tag{71}$$

(eller $\sum_{j=1}^{k} j = \frac{k(k+1)}{2}$). Dette er *induksjonshypotesen*. Vi vil bruke dette til å vise at $P(k + 1)$ holder (*induksjonstrinnet*). Da regner vi ut:

$$\begin{aligned}
1+2+3+\cdots+k+(k+1) &= (1+2+3+\cdots+k)+(k+1) \\
&\overset{(71)}{=} \frac{k(k+1)}{2}+k+1 \\
&= \frac{k(k+1)+2(k+1)}{2} \\
&= \frac{(k+1)(k+2)}{2}.
\end{aligned}$$

(Første linje kan skrives som $\sum_{j=1}^{k+1} j = \sum_{i=j}^{k} j + (k+1)$.) Dette er $P(n)$ for $n = k+1$. Vi har derfor vist at hvis $P(k)$ er sant, så er også $P(k+1)$ sant og vi er ferdig. □

Merk hvor vi brukte induksjonshypotesen i induksjonstrinnet.

Merknad 5.2.5. Det finnes et annet bevis for (70). I følge en anekdote, hadde den tyske matematikeren Johann Carl Friedrich Gauss (1777 – 1855), da han var åtte, en lærer som ga hele klassen i oppgave å addere alle positive heltall fra 1 til 100, for å holde elevene opptatt. Gauss kom raskt frem til det riktige svaret, $5050 = \frac{100\cdot 101}{2}$, ved først å finne frem til formelen (70). Den kom han frem til ved å legge merke til at 2 ganger summen kunne skrives som

$$\begin{array}{ccccccccccccccc}
1 & + & 2 & + & 3 & + & \cdots & + & (n-2) & + & (n-1) & + & n & + \\
n & + & (n-1) & + & (n-2) & + & \cdots & + & 3 & + & 2 & + & 1 &
\end{array}$$

Siden summen av hver kolonne er $n+1$ og det finnes n kolonner, er $2(1+2+\cdots+n) = n(n+1)$.

Eksempel 5.2.6. Bruk matematisk induksjon til å vise at

$$\sum_{j=0}^{n} j\cdot j! = (n+1)! - 1 \tag{72}$$

for alle naturlige tall n. (Vi minner om definisjonen av $n!$ i Definisjon 3.8.1.)

Bevis ved induksjon. For $n = 0$ er venstresiden i (72) lik $0\cdot 0! = 0$, og høyresiden er lik $(0+1)! - 1 = 1! - 1 = 1 - 1 = 0$, slik at formelen (72) stemmer for $n = 0$.

Anta nå at formelen stemmer for $n = k$, der k er et naturlig tall. Altså antar vi at

$$\sum_{j=0}^{k} j\cdot j! = (k+1)! - 1. \tag{73}$$

Da vil venstresiden i (72) for $n = k+1$ kunne skrives som:

$$\begin{aligned}
\sum_{j=0}^{k+1} j \cdot j! &= 0 \cdot 0! + 1 \cdot 1! + 2 \cdot 2! + 3 \cdot 3! + \cdots + (k+1) \cdot (k+1)! \\
&= [0 \cdot 0! + 1 \cdot 1! + 2 \cdot 2! + 3 \cdot 3! + \cdots + k \cdot k!] + (k+1) \cdot (k+1)! \\
&= \sum_{j=0}^{k} j \cdot j! + (k+1) \cdot (k+1)! \\
&\overset{(73)}{=} \Big((k+1)! - 1\Big) + (k+1) \cdot (k+1)! \\
&= (k+1)!\Big(1 + (k+1)\Big) - 1 \\
&= (k+1)!(k+2) - 1 \\
&= (k+2)! - 1
\end{aligned}$$

(hvor vi har brukt observasjonen i Merknad 3.8.2 i siste likhet) og dette er lik høyresiden i (72) med $n = k+1$.

Vi har derfor vist at dersom (72) holder for et naturlig tall $n = k$, da holder (72) også for $n = k+1$.

Vi har dermed vist at (72) holder for alle naturlige tall ved induksjon. □

Eksempel 5.2.7. Vis at $4n < 2^n$ for alle heltall $n \geq 5$.

Bevis ved induksjon. Vi skal vise at ulikheten

$$4n < 2^n \tag{74}$$

er riktig for alle heltall $n \geq 5$.

For $n = 5$ sier (74)

$$4 \cdot 5 < 2^5,$$

med andre ord

$$20 < 32,$$

som er riktig.

Anta nå at (74) holder for et heltall $n = k \geq 5$, det vil si anta

$$4k < 2^k, \quad \text{for et heltall } k \geq 5. \tag{75}$$

Vi ønsker ut fra dette å vise at (74) også holder for $n = k+1$, det vil si at

$$4(k+1) < 2^{k+1} \tag{76}$$

er sant. Siden venstresiden i siste ulikhet er $4k+4$, forsøker vi oss med å addere 4 på begge sider av ulikheten (75). Da får vi

$$4k + 4 < 2^k + 4. \tag{77}$$

Vi observerer at $4 = 2^2 \leq 2^k$, siden $k \geq 5$, slik at

$$2^k + 4 \leq 2^k + 2^k = 2 \cdot 2^k = 2^{k+1}. \tag{78}$$

Setter vi sammen (77) og (78), får vi:

$$4(k+1) < 2^k + 4 \leq 2^{k+1},$$

som gir oss den ønskede ulikheten (76).

Vi har dermed vist ulikheten (74) for alle $n \geq 5$ ved induksjon. □

EKSEMPEL 5.2.8. Vis at $8^n - 3^n$ er delelig på 5 for alle heltall $n \geq 1$.

BEVIS VED INDUKSJON. For $n = 1$ er $8^n - 3^n = 8 - 3 = 5$, som er delelig på 5. Anta nå at

$$8^k - 3^k \text{ er delelig på 5 for et heltall } k \geq 1. \tag{79}$$

Vi vil vise at også $8^{k+1} - 3^{k+1}$ er delelig på 5. Da skriver vi

$$\begin{aligned} 8^{k+1} - 3^{k+1} &= 8 \cdot 8^k - 3 \cdot 3^k \\ &= 8 \cdot 8^k - 8 \cdot 3^k + 8 \cdot 3^k - 3 \cdot 3^k \\ &= 8 \cdot (8^k - 3^k) + (8 - 3) \cdot 3^k \\ &= 8 \cdot (8^k - 3^k) + 5 \cdot 3^k. \end{aligned}$$

Induksjonshypotesen (79) gir at $8 \cdot (8^k - 3^k)$ er delelig på 5. Dessuten er også $5 \cdot 3^k$ delelig på 5. Men da er også summen $8 \cdot (8^k - 3^k) + 5 \cdot 3^k$ delelig på 5 ved Setning 3.7.2(i), slik at vi har vist at $8^{k+1} - 3^{k+1}$ er delelig på 5. Vi har dermed vist påstanden for alle $n \geq 1$ ved induksjon. □

MERKNAD 5.2.9. Hvordan vi kom på "trikset" med å legge til og trekke fra $8 \cdot 3^k$ i andre linje av kjeden av likheter i siste bevis? Vi ønsker å kunne bruke induksjonshypotesen (79), som involverer uttrykket $8^k - 3^k$. Vi ser at $8^{k+1} = 8 \cdot 8^k$, og for å fremskaffe uttrykket $8^k - 3^k$ trekker vi fra $8 \cdot 3^k$, og skriver $8 \cdot 8^k - 8 \cdot 3^k = 8 \cdot (8^k - 3^k)$. Selvsagt må vi da også legge til $8 \cdot 3^k$.

EKSEMPEL 5.2.10. Vis at en endelig mengde A med n elementer (jf. Definisjon 2.3.15) har nøyaktig 2^n forskjellige delmengder.

BEVIS VED INDUKSJON. Basistilfellet for induksjonen er $n = 0$, hvor $A = \emptyset$ per definisjon. Den tomme mengden har bare seg selv som delmengde, altså har $\emptyset$ nøyaktig $1 = 2^0$ delmengder, slik at utsagnet stemmer for $n = 0$.

Anta nå at påstanden stemmer for $n = k$, hvor $k \in \mathbb{N}$, det vil si at en endelig mengde med k elementer har nøyaktig 2^k forskjellige delmengder. La A være en endelig mengde med $k+1$ elementer. Siden $k + 1 \geq 1$, har A minst ett element. La oss plukke en $a \in A$. Da kan vi skrive

$A = (A \setminus \{a\}) \cup \{a\}$. Merk at $A \setminus \{a\}$ er endelig og inneholder k elementer. Ved induksjonshypotesen har A nøyaktig 2^k forskjellige delmengder.

En delmengde $B \subset A$ kan være av to typer:

(I) $a \notin B$,
(II) $a \in B$.

I tilfelle (I) er $B \subset A \setminus \{a\}$. Ved induksjonshypotesen finnes nøyaktig 2^k forskjellige slike delmengder B.

I tilfellet (II) er $B = (B \setminus \{a\}) \cup \{a\}$, der $B \setminus \{a\} \subset A \setminus \{a\}$. Ved induksjonshypotesen finnes nøyaktig 2^k forskjellige slike delmengder $B \setminus \{a\}$, og dermed nøyaktig 2^k forskjellige slike delmengder B.

Siden det er klart at delmengdene B som i (I) og (II) er forskjellige, har vi alt i alt $2 \cdot 2^k = 2^{k+1}$ delmengder av A. Dermed stemmer påstanden også for $n = k + 1$. Vi har dermed vist påstanden ved induksjon. □

Induksjon kan brukes til å generalisere mengdeidentitetene i tabellen på side 2.1 til å gjelde snitt og union av et vilkårlig endelig antall mengder. Vi viser et eksempel og overlater resten til Oppgave 5.13.

Eksempel 5.2.11. Vis at hvis $A_1, \ldots, A_n$ og C er mengder, der $n \geq 2$ er et heltall, da er

$$C \setminus (A_1 \cap \cdots \cap A_n) = (C \setminus A_1) \cup \cdots \cup (C \setminus A_n). \tag{80}$$

Bevis ved induksjon. Vi betrakter først basistilfellet $n = 2$. Da sier (80) at $C \setminus (A_1 \cap A_2) = (C \setminus A_1) \cup (C \setminus A_2)$, som stemmer ved Setning 2.1.14.

Anta nå at (80) stemmer for $n = k$, der $k \geq 2$ er et heltall, det vil si anta at

$$C \setminus (A_1 \cap \cdots \cap A_k) = (C \setminus A_1) \cup \cdots \cup (C \setminus A_k) \text{ for et heltall } k \geq 2. \tag{81}$$

Vi har

$$\begin{aligned} C \setminus (A_1 \cap \cdots \cap A_{k+1}) &\overset{(*)}{=} C \setminus \Big((A_1 \cap \cdots \cap A_k) \cap A_{k+1} \Big) \\ &\overset{(**)}{=} C \setminus (A_1 \cap \cdots \cap A_k) \cup C \setminus A_{k+1} \\ &\overset{(81)}{=} \Big(C \setminus A_1 \cup \cdots \cup C \setminus A_k \Big) \cup C \setminus A_{k+1} \\ &\overset{(*)}{=} C \setminus A_1 \cup \cdots \cup C \setminus A_{k+1}, \end{aligned}$$

hvor vi i $(*)$ har brukt *assosiativitesegenskapen* til snitt og union av mengder (jf. tabellen på side 41) og i $(**)$ har brukt Setning 2.1.14 med $A = (A_1 \cap \cdots \cap A_k)$ og $B = A_{k+1}$. Dette viser (80) for $n = k + 1$. Vi har derfor vist at (80) er sant for alle heltall $n \geq 2$ ved induksjon. □

Typiske feil i induksjonsbevis. En vanlig feil i induksjonsproblemer er at man blir så opphengt i induksjonstrinnet at man glemmer å vise påstanden for $n = n_0$. Følgende eksempel viser hvor galt det da kan gå, i det at man "viser" en påstand som er opplagt feil.

"Oppgave". *Vis at $n+1 < n$ for alle heltall $n \geq 1$.*

"Løsning". Vi antar ulikheten holder for $n = k$, det vil si $k+1 < k$. Adderer vi 1 på begge sider, får vi $k+1+1 < k+1$, det vil si $(k+1)+1 < (k+1)$, som er ulikheten vi skal vise for $n = k+1$. Vi har dermed vist at $n+1 < n$ for alle heltall $n \geq 1$ ved induksjon. □

Det morsomme her er at man har utført induksjonstrinnet riktig, og man har vist at "dersom påstanden er sann for k, så er den også sann for $k+1$". Problemet er at man ikke har vist at påstanden er riktig for noen k i det hele tatt!

En annen, mer subtil feil, er gjemt i følgende "løsning". Forsøk å finne feilen selv før du leser teksten etter den feilaktige løsningen.

"Oppgave". *Vis at en mengde av minst to linjer i planet, hvor ingen er innbyrdes parallelle, alltid har ett punkt til felles.*

"Løsning": Vi betrakter utsagnet, for ethvert heltall $n \geq 2$:

$$P(n): \quad n \text{ forskjellige linjer i planet, ingen innbyrdes parallelle} \\ \text{går alle gjennom et felles punkt.}$$

Dette utsagnet er opplagt sant for $n = 2$: to forskjellige ikke-parallelle linjer møtes i ett punkt. Altså holder $P(2)$.

Anta nå at utsagnet er sant for et heltall $n = k$, det vil si at $P(k)$ holder. For å fullføre induksjonstrinnet, må vi vise at utsagnet også holder for $n = k+1$, det vil si at $P(k+1)$ holder. Vi må altså vise at for enhver mengde av $k+1$ forskjellige linjer i planet, ingen innbyrdes parallelle, så går alle gjennom et felles punkt.

Betrakt derfor en mengde av $k+1$ forskjellige linjer i planet, kall dem $\ell_1, \ldots, \ell_{k+1}$, ingen innbyrdes parallelle. Ved induksjonshypotesen, det vil si det faktum at $P(k)$ holder, vil de første k av disse, $\ell_1, \ldots, \ell_k$, gå gjennom et felles punkt, kall det p. Likeledes, fremdeles ved induksjonshypotesen, vil de siste k linjene, $\ell_2, \ldots, \ell_{k+1}$, gå gjennom et felles punkt, kall det q. Vi vil nå vise at $p = q$: hvis $p \neq q$, ville alle linjene som inneholder begge punktene, nemlig $\ell_2, \ldots, \ell_k$, være samme linje, fordi det finnes en entydig linje gjennom to punkter. Dette motsier antagelsen om at linjene er forskjellige. Derfor er $p = q$, og alle $k+1$ linjene går gjennom dette punktet.

Vi har derfor vist at $P(k+1)$ holder hvis $P(k)$ holder.

Vi har dermed vist at $P(n)$ holder for alle heltall $n \geq 2$ ved induksjon.□

Klarer du å se hvor feilen ligger? Feilen ligger i selve induksjonstrinnet: Trinnet der vi viser at påstanden for $n = k$ medfører påstanden for $n = k+1$ (formelt sagt: $P(k) \Rightarrow P(k+1)$) fungerer ikke dersom $k = 2$: dersom vi har tre linjer, ℓ_1, ℓ_2, ℓ_3 og de to første, ℓ_1 og ℓ_2, går gjennom p og de to siste, ℓ_2 og ℓ_3, går gjennom q, så finnes kun én linje som går gjennom p og q, nemlig ℓ_2, og vi kan ikke konkludere at $p = q$.

Moralen er at når vi beviser noe ved induksjon, da må beviset for at $P(k) \Rightarrow P(k+1)$ gjelde *for enhver* $k \geq n_0$, der n_0 er det tallet der vi starter induksjonen, *også for* $k = n_0$.

Ellers er det en ikke så uvanlig feil å tro at man har vist en formel/påstand kun ved å ha sjekket den i de første tilfellene (spesielt skjer dette i tilfeller der man blir bedt om å først gjette en formel/påstand basert på noen enkelttilfeller, og deretter å *bevise* den). Følgende to velkjente eksempler viser at man ikke kan trekke generelle slutninger ut i fra (noen få) enkelttilfeller.

EKSEMPEL 5.2.12. ("The Monstrous Counterexample", fra [**Da**].) La oss betrakte påstanden

$$P(n): \quad 1141n^2 + 1 \text{ er } \textit{ikke} \text{ et kvadrattall.}$$

(Med et *kvadrattall* mener vi et naturlig tall som kan skrives som x^2, for et naturlig tall x.) Det er lett å sjekke at denne er riktig for de laveste positive heltallene. Faktisk kan man vise at påstanden er riktig for alle heltall n slik at

$$1 \leq n \leq 30\,693\,385\,322\,765\,657\,197\,397\,207.$$

Men det morsomme er at påstanden slår feil for neste heltall,

$$n = 30\,693\,385\,322\,765\,657\,197\,397\,208.$$

EKSEMPEL 5.2.13. Et *Fermat-tall*, oppkalt etter den berømte franske matematikeren (og juristen!) Pierre de Fermat (1607–1665), er et tall på formen

$$F_n = 2^{2^n} + 1, \quad \text{for en } n \in \mathbb{N}.$$

De første Fermat-tallene er

$$\begin{aligned} F_0 &= 3 \\ F_1 &= 5 \\ F_2 &= 17 \\ F_3 &= 257 \\ F_4 &= 65537 \\ F_5 &= 4294967297 \\ F_6 &= 18446744073709551617 \end{aligned}$$

Fermat sjekket at de fem første Fermat-tallene $F_0, \dots, F_4$ er primtall (se Definisjon 3.7.6) og la frem en *formodning* om at alle F_n var det. Den sveitsiske matematikeren Leonhard Euler (1707–1783) viste imidlertid i 1732 at F_5 ikke er det, siden $4294967297 = 641 \cdot 6700417$. (Fremdeles i dag kjenner man ingen andre Fermat-tall enn de fem første som er primtall. Man vet at ingen av F_n for $5 \leq n \leq 32$ er primtall, og det største Fermat-tallet man har vist *ikke* er et primtall er $F_{3329780}$.)

5.3. Fullstendig induksjon

I matematikken brukes også en litt annen variant av induksjonsprinsippet, noen ganger kalt *fullstendig induksjon*[5] eller *sterk induksjon*[6]. Denne varianten gjennomgås ikke i MAT111.

Induksjonsprinsippet (versjon II)

La $P(n)$ være et utsagn som gir mening for ethvert heltall $n \geq n_0$, der n_0 er et fiksert heltall.

For å vise at $P(n)$ er sann for alle heltall $n \geq n_0$ er det nok å vise følgende:

(I) $P(n_0)$ er sant.

(II) Hvis utsagnet $P(m)$ er sant for alle heltall m som oppfyller $n_0 \leq m \leq k$, da er også $P(k+1)$ sant.

Forskjellen er altså at vi antar noe tilsynelatende sterkere i induksjonshypotesen, nemlig ikke bare at $P(k)$ er sann, men at alle $P(n_0), \dots, P(k)$ er sanne.

Grunnen til at dette virker er igjen induksjonsaksiomet (N5):

Setning 5.3.2

Induksjonsaksiomet (N5) er ekvivalent med induksjonsprinsippet (versjon II).

Bevis

Vi viser at induksjonsaksiomet (N5) impliserer induksjonsprinsippet (versjon II) og overlater det til Oppgave 5.4 å vise at induksjonsprinsippet (versjon II) impliserer induksjonsaksiomet (N5).

Betrakt delmengden

$$M = \{n \in \mathbb{N} \mid P(n_0), P(n_0+1), \dots, P(n_0+n) \text{ er sant}\} \subset \mathbb{N}.$$

Antagelsen (I) i induksjonsprinsippet gir at $0 \in M$.

Anta at $n \in M$ for en $n \in \mathbb{N}$, det vil si at $P(n_0), P(n_0+1), \dots, P(n_0+n)$ alle er sanne. Da gir antagelsen (II) i induksjonsprinsippet at $P(n_0+n+1)$

[5] *Complete induction* på engelsk.
[6] *Strong induction* på engelsk.

er sant, det vil si $n+1 \in M$. Siden vi i (42) viste at $n+1 = S(n)$, *etterfølgeren* til n, gir induksjonsaksiomet (N5) da at $M = \mathbb{N}$, som betyr at $P(n_0 + n)$ er sant for alle $n \geq 0$, ekvivalent at $P(n)$ er sant for alle $n \geq n_0$. □

Setningene 5.2.3 og 5.3.2 til sammen viser at de to induksjonstypene er ekvivalente. Derfor er det egentlig ikke nødvendig å skille mellom disse to typene induksjon så mye som enkelte lærebøker gjør. Sistnevnte type brukes i tilfeller hvor vi ikke så enkelt klarer å redusere påstanden $P(k+1)$ til noe som involverer $P(k)$, men bare klarer å redusere den til $P(n)$ med enda lavere n.

Eksempler. Et typisk og klassisk eksempel på bruk av fullstendig induksjon er beviset for *aritmetikkens fundamentalteorem* (Teorem 3.7.7) fra §3.7, som vi ikke beviste der, men som vi nå skal bevise. Vi skriver teoremet opp igjen:

Teorem 3.7.7: Aritmetikkens fundamentalteorem

Ethvert heltall $n \geq 2$ kan skrives som et produkt av primtall, på en *entydig* måte (bortsett fra rekkefølgen av faktorene).

Bevis (ved (fullstendig) induksjon)

Påstanden vi skal vise er

$$P(n): \quad n \text{ kan skrives som et entydig produkt av primtall}$$

(bortsett fra rekkefølgen av faktorene) for alle heltall $n \geq 2$.

Første trinn er å observere at $P(2)$ er opplagt sann, siden 2 er et produkt med seg selv som eneste faktor, og 2 er primtall.

La nå $k \geq 2$ og anta at $P(m)$ er sann for alle $m \in \mathbb{Z}$ slik at $2 \leq m \leq k$, det vil si

$$(82) \qquad 2, 3, \ldots, k \text{ kan skrives som et entydig produkt av primtall.}$$

Vi sjekker $P(k+1)$.

Dersom $k+1$ er et primtall, er $P(k+1)$ sann (av samme grunn som at $P(2)$ var sann), siden $k+1$ er et produkt med én faktor, nemlig seg selv, som er primtall.

Dersom $k+1$ *ikke* er et primtall, da kan vi skrive det som et produkt $k+1 = ab$ av to heltall a og b som begge er ≥ 2 og $< k+1$. Induksjonshypotesen (82) gir at både a og b kan skrives som produkt av primtall:

$$\begin{aligned} a &= p_1 p_2 \cdots p_s \\ b &= q_1 q_2 \cdots q_t. \end{aligned}$$

Dermed får vi at

$$k+1 = ab = p_1 p_2 \cdots p_s q_1 q_2 \cdots q_t$$

også er et produkt av primtall. Dette betyr at $k+1$ kan skrives som et produkt av primtall, altså at *eksistensdelen* av utsagnet $P(k+1)$ er sant. Det gjenstår derfor å vise at primfaktorene i $k+1$ er entydige.

Vi utfører et bevis ved selvmotsigelse og antar derfor at

$$k+1 = p_1 \cdots p_m = q_1 \cdots q_n,$$

for heltall $m, n \geq 2$, er to ulike primtallsfaktoriseringer. Hvis en av p-ene var lik en av q-ene, si $p_i = q_j$, ville vi ved forkortning (Setning 3.6.3(v)) fått $\frac{k+1}{p_i} = \frac{k+1}{q_j}$, som ville vært et heltall mindre enn $k+1$ og større enn 1 med to forskjellige primtallsfaktoriseringer. Dette motsier induksjonshypotesen (82). Altså er alle p-ene forskjellig fra alle q-ene.

Vi kan anta at $p_1 < q_1$. (Hvis ikke, bytter vi rollen på p-ene og q-ene.) Betrakt

$$x = (q_1 - p_1) q_2 \cdots q_n. \tag{83}$$

Siden $0 < q_1 - p_1 < q_1$, har vi $1 < q_2 \leq x < k+1$. Dermed må x ha en entydig primtallsfaktorisering, igjen ved induksjonshypotesen (82).

Påstand 5.3.3.1. p_1 er en faktor i x.

Bevis for påstand. Vi omskriver

$$\begin{aligned} x &= (q_1 - p_1) q_2 \cdots q_n \\ &= q_1 q_2 \cdots q_n - p_1 q_2 \cdots q_n \\ &= (k+1) - p_1 q_2 \cdots q_n \\ &= p_1 p_2 \cdots p_m - p_1 q_2 \cdots q_n \\ &= p_1 (p_2 \cdots p_m - q_2 \cdots q_n), \end{aligned}$$

som viser påstanden. □

Siden x har en entydig primtallfaktorisering og p_1 er forskjellig fra $q_2, \ldots, q_n$, ser vi fra (83) og påstanden at p_1 må være en faktor i $q_1 - p_1$. Dette betyr at det finnes et positivt heltall t slik at $q_1 - p_1 = tp_1$, som gir at $q_1 = (t+1)p_1$. Men siden $t+1 \geq 2$ og $p_1 \geq 2$, betyr dette at q_1 ikke kan være primtall, en motsigelse. □

Neste eksempel er av litt mer "praktisk karakter" og er en "klassiker". Den er også spesielt interessant fordi vi i beviset trenger å vise noen flere spesialtilfeller i tillegg til første trinn i induksjonen.

Eksempel 5.3.4. Vis at enhver porto på 12 cents og høyere kan oppnås ved å kombinere frimerker på 4 og 5 cents.

Bevis ved (fullstendig) induksjon. Påstanden vi skal vise kan matematisk formuleres som

$$P(n): \quad n \text{ kan skrives som en sum av 4ere og 5ere}$$

for alle heltall $n \geq 12$.

Vi observerer først at $P(12)$ er opplagt sann, siden $12 = 4+4+4 = 3\cdot 4$.

Vi observerer også, siden vi kommer til å trenge det senere i beviset, at påstandene $P(13)$, $P(14)$ og $P(15)$ holder:

$$\begin{aligned} P(13): && 13 &= 2\cdot 4+5 \\ P(14): && 14 &= 4+2\cdot 5 \\ P(15): && 15 &= 3\cdot 5. \end{aligned}$$

Anta nå at $P(n)$ holder for alle $n = 12, 13, 14, \ldots, k$, for et heltall $k \geq 12$. Vi vil nå vise at også $P(k+1)$ holder.

Dersom $k+1 \leq 15$, er vi ferdige, fordi vi er i ett av spesialtilfellene vi allerede har sjekket.

Dersom $k+1 \geq 16$, er $(k+1)-4 \geq 12$. (Her ligger nødvendigheten av å vise tilfellene $13, 14, 15$ for seg.) Ved induksjonshypotesen vet vi at $P((k+1)-4)$ holder, det vil si at vi kan skrive

$$(k+1)-4 = 4a+5b, \quad \text{for naturlige tall } a, b.$$

Men da vil

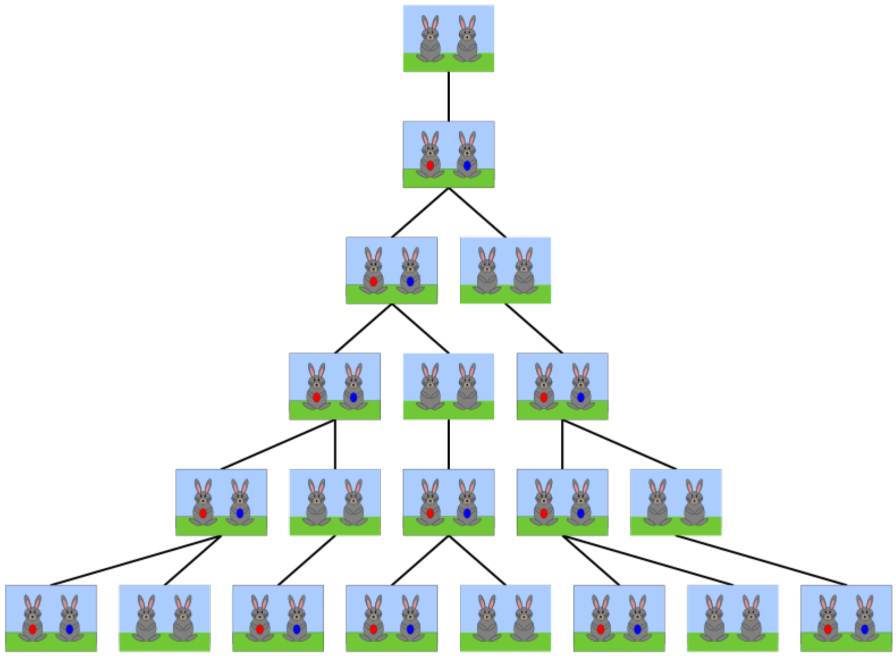
$$k+1 = 4a+5b+4 = 4(a+1)+5b,$$

som viser at $P(k+1)$ holder. Vi har dermed vist $P(n)$ for alle $n \geq 12$ ved induksjon. □

Typiske tilfeller der det er naturlig å bruke fullstendig induksjon er hvor størrelser "i ett trinn" er definert *rekursivt* ved hjelp av størrelser i flere foregående trinn. Et kjent slikt eksempel er *Fibonaccitallene*:

EKSEMPEL 5.3.5 (Fibonaccitallene). *Fibonaccitallene* har navnet fra den italienske matematikeren Leonardo Pisano (ca. 1170–1245), mest kjent som Fibonacci, som brukte dem i sin bok *Liber Abaci* (1202) for å beregne veksten av populasjoner av kaniner, selv om tallene tilsynelatende var kjent hos indiske matematikere flere hundre år før Kristus.

Problemet som Fibonacci betraktet er følgende: Et ungt par av kaniner, én av hvert kjønn, blir plassert på en øy. Et par av kaniner parer seg først når de er 1 måned gamle, og gir opphav til et nytt kaninpar når de er 2 måneder gamle. Etter denne alderen gir hvert par opphav til et nytt kaninpar hver måned, som vist på figuren:



KANINER SOM FORMERER SEG PÅ EN ØY. (KILDE: WIKIPEDIA, GNU FREE DOCUMENTATION LICENSE)

La F_n betegne antallet kaninpar på øyen mot slutten av måned n, når vi antar at kaniner aldri dør (en fullstendig urealistisk antagelse, men la gå!) Mot slutten av den første måneden er det ett par av kaniner, så $F_1 = 1$. Mot slutten av måned 2 har kaninene paret seg, men enda ikke født, slik at det fremdeles er kun ett par av kaniner på øyen. Altså er også $F_2 = 1$. Antallet kaninpar mot slutten av måned 3 er da lik 2, siden det opprinnelige paret nå har gitt opphav til et nytt kaninpar. Altså er $F_3 = 2$, som er lik $F_2 + F_1$. Dette er det generelle mønsteret: Mot slutten av måned n er antallet kaninpar lik summen av kaninparene vi hadde måneden tidligere, nemlig F_{n-1}, og summen av de nyfødte parene, som er F_{n-2}, siden hvert nytt kaninpar kommer fra et par kaniner som er mist to måneder gamle. Altså er $F_n = F_{n-1} + F_{n-2}$, for hver $n \geq 3$. Som oppsummering er det n-te fibonaccitallet F_n for enhver $n \in \mathbb{Z}^+$ definert *rekursivt* ved

$$F_1 = 1, \quad F_2 = 1, \quad F_n = F_{n-1} + F_{n-2} \text{ for hver } n \geq 3.$$

Noen ganger definerer man $F_0 = 0$, slik at fibonaccitallene F_n for enhver $n \in \mathbb{N}$ blir definert *rekursivt* ved

$$F_0 = 0, \quad F_1 = 1, \quad F_n = F_{n-1} + F_{n-2} \text{ for hver } n \geq 2.$$

En slik formel kalles gjerne en *rekursjonsformel*. Altså er hvert Fibonaccitall F_n for $n \geq 2$ summen av de to foregående.

La oss bruke fullstendig induksjon til å vise at

$$F_n \geq \left(\frac{3}{2}\right)^{n-2} \tag{84}$$

for alle $n \in \mathbb{Z}^+$, med andre ord at *veksten er eksponensiell*.

For $n = 1$ (basistilfellet) sier (84) at $F_1 \geq \frac{3}{2}^{-1} = \frac{2}{3}$, som er sant.

La nå $k \geq 1$ være et heltall. Vi antar nå at (84) er sant for alle heltall $n = j$ slik at $1 \leq j \leq k$ (dette er induksjonshypotesen).

Vi kan skrive $F_{k+1} = F_k + F_{k-1}$ (dette gir mening siden $k \geq 1$). Hvis $k \geq 2$, kan vi bruke induksjonshypotesen på både F_k og F_{k-1}, siden $k-1 \geq 1$. Da får vi

$$\begin{aligned} F_{k+1} &= F_k + F_{k-1} \geq \left(\frac{3}{2}\right)^{k-2} + \left(\frac{3}{2}\right)^{k-3} = \left(\frac{3}{2}\right)^{k-3}\left(\frac{3}{2}+1\right) \\ &= \left(\frac{3}{2}\right)^{k-3} \cdot \frac{5}{2} = \left(\frac{3}{2}\right)^{k-3} \cdot \frac{10}{4} \geq \left(\frac{3}{2}\right)^{k-3} \cdot \left(\frac{3}{2}\right)^{2} = \left(\frac{3}{2}\right)^{k-1}, \end{aligned}$$

som betyr at utsagnet (84) er sant også for $n = k + 1$. Vi må imidlertid huske at vi fremdeles må behandle tilfellet $k = 1$: da har vi $F_2 = 1 = \left(\frac{3}{2}\right)^0$, slik at (84) er sant også i dette tilfellet. (Som i forrige eksempel måtte vi altså sjekke et spesialtilfelle i tillegg til basistilfellet.)

Vi har dermed vist (84) for alle $n \in \mathbb{Z}^+$ ved (fullstendig) induksjon.

Oppgaver

OPPGAVE 5.1. (a) Utled *Velordningsprinsippet* fra Peanoaksiomene og ordensrelasjonen på $\mathbb{N}$. (Hint: argumentér ved selvmotsigelse og anta at det finnes en ikke-tom delmengde $A \subset \mathbb{N}$ som ikke har et minste element. Betrakt da mengden $M = \{n \in \mathbb{N} \mid i \notin A \text{ for alle } i \leq n\}$ og vis at $M = \mathbb{N}$. Du får bruk for Setning 3.5.4.)

(b) Vis at dersom vi erstatter aksiom (N4) i Peanoaksiomene med

(N4)' 0 er ikke etterfølgeren til et element i $\mathbb{N}$ og ethvert annet element i $\mathbb{N}$ er etterfølgeren til et element i $\mathbb{N}$ (det vil si: det finnes ikke noen $n \in \mathbb{N}$ slik at $S(n) = 0$, og for enhver $n \in \mathbb{N}$ slik at $n \neq 0$ finnes en $m \in \mathbb{N}$ slik at $S(m) = n$),

da er *Velordningsprinsippet* ekvivalent med *Induksjonsaksiomet* (N5).

OPPGAVE 5.2. Bevis entydighetsdelen av *divisjonsalgoritmen* (Setning 3.7.3).

OPPGAVE 5.3. Bevis Setning 4.1.6, ved å “mime etter” beviset for *divisjonsalgoritmen* (Setning 3.7.3). Start med å betrakte mengden

$$S = \{P \in K[x] \mid P = F - GQ \text{ for en } Q \in K[x]\}.$$

Hvis $0 \in S$, er du ferdig (ved å bruke at $\deg 0 = -1$). Hvis $0 \notin S$, la $R \in S$ være et polynom av minste grad (en slik minste grad eksisterer ved *velordningsprinsippet* i §5.1).

Oppgave 5.4. Vis at de to formene for induksjon hver impliserer induksjonsaksiomet (N5). Dette fullfører bevisene for Setningene 5.2.3 og 5.3.2.

Oppgave 5.5. Bruk matematisk induksjon til å vise at produktet av tre påfølgende positive heltall alltid er delelig på 6.

Oppgave 5.6. Vis at

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

for alle heltall $n \geq 1$.

Dette ble først bevist i boken *Arithmeticorum libri duo* (1575) av den italienske matematikeren og astronomen Francesco Maurolico (1494–1575).

Oppgave 5.7. Vis at

$$\sum_{j=0}^{n}(2j+1)^2 = 1^2+3^2+5^2+\cdots+(2n+1)^2 = \frac{(n+1)(2n+1)(2n+3)}{3}$$

for alle positive heltall n.

Oppgave 5.8. Vis at $2^n < n!$ for alle heltall $n \geq 4$.

Oppgave 5.9. La K være en ordnet kropp (jf. Definisjon 3.6.12) og $x \in K$ slik at $x > 1$ (henholdsvis $0 < x < 1$). Vis at da er $x^{n+1} > x^n$ (hhv. $x^{n+1} < x^n$) for alle $n \in \mathbb{N}$.

Vis lignende påstander med "$>$" erstattet med "$\geq$".

Vis også at dersom $0 < x_1 < x_2$ i K, da er også $x_1^n < x_2^n$ for alle $n \in \mathbb{Z}^+$.

Oppgave 5.10. La K være en kropp og $r \in K \setminus \{1\}$. Vis at for enhver $n \in \mathbb{Z}^+$ gjelder

$$\sum_{j=0}^{n} r^j = 1+r+r^2+\cdots+r^n = \frac{r^{n+1}-1}{r-1}.$$

(Hvis du ikke føler deg komfortabel med å la K være en vilkårlig kropp, kan du først utføre beviset for $K = \mathbb{Q}$, og deretter sjekke at alt fungerer hvis du erstatter $\mathbb{Q}$ med en vilkårlig kropp.)

Oppgave 5.11. Bruk induksjon til å vise at binomialkoeffisienter (jf. Definisjon 3.8.3) alltid er heltall.

Oppgave 5.12. Vis at ethvert heltall enten er et partall eller er et oddetall, som hevdet i Merknad 1.2.3. Ta utgangspunkt i Definisjonene 1.2.1 of 1.2.2. (Hint: Bruk velordningsprinsippet, induksjon eller divisjonsalgoritmen.)

Oppgave 5.13. Vis følgende identiteter for mengder:

(a) $A \cap (B_1 \cup \cdots \cup B_n) = (A \cap B_1) \cup \cdots \cup (A \cap B_n)$.
(b) $A \cup (B_1 \cap \cdots \cap B_n) = (A \cup B_1) \cap \cdots \cap (A \cup B_n)$.

(c) $C \setminus (A_1 \cup \cdots \cup A_n) = (C \setminus A_1) \cap \cdots \cap (C \setminus A_n)$.

Oppgave 5.14. (a) Vis *Pascals regel*

$$\binom{n+1}{i} = \binom{n}{i-1} + \binom{n}{i},$$

enten ved å bruke et kombinatorisk argument (det vil si et "telle-argument") som i §3.8) eller ved å simpelthen regne ut de to sidene av likhetstegnet og vise at de er like (eller begge deler!).

(b) Vis binomialteoremet (Teorem 3.8.7) ved induksjon på n. Du får bruk for Pascals regel fra (a).

Oppgave 5.15. Vis at ethvert positivt heltall n kan skrives som en sum av forskjellige potenser av 2, det vil si, som en sum av en delmengde av de naturlige tallene $2^0 = 1$, $2^1 = 2$, $2^2 = 4$, osv. (Hint: i induksjonstrinnet, betrakt hver for seg tilfellene der $k + 1$ er et oddetall og partall.)

Oppgave 5.16. Vis at enhver porto på $n \geq 8$ cents kan fås ved å kombinere frimerker på 3 og 5 cents.

Oppgave 5.17. La $a_1 = 2$, $a_2 = 9$ og $a_n = 2a_{n-1} + 3a_{n-2}$ for ethvert heltall $n \geq 3$. Vis at $a_n \leq 3^n$ for alle positive heltall n.

Oppgave 5.18. I en round-robin-biljardturnering er det n deltagere. (Enhver deltager spiller mot en annen deltager nøyaktig én gang og et spill kan ikke ende uavgjort.) Vis at uansett hvordan spillene slutter, vil det være mulig etter turneringen å lage en liste av alle deltagerne slik at hver deltager har slått neste deltager i listen i spillet der de møttes.

Oppgave 5.19. Vis følgende egenskaper til fibonaccitallene (definert i Eksempel 5.3.5):

(a) $F_{n+1} \leq 2F_n$ for alle $n \geq 1$.
(b) Fibonaccitallene F_{5k} er delelige på 5 for alle $k \geq 1$. (Hint: Vis først, ut i fra definisjonen, at $F_{n+5} = 5F_{n+1} + 3F_n$ for alle $n \geq 0$.)
(c) $F_{n+1}F_{n-1} = F_n^2 + (-1)^n$.

Oppgave 5.20. Finn feilen i følgende induksjonsbevis:

"Oppgave": *Vis at alle fibonaccitall (jf. Definisjonen i Oppgave 5.19) er partall*

"Løsning": Vi bruker sterk induksjon. Siden $F_0 = 0$, er påstanden sann for $n = 0$. Anta $F_0, F_1, \ldots, F_k$ er partall, for en $k \geq 0$. Da er

$$F_{k+1} = F_k + F_{k-1} = \text{partall} + \text{partall} = \text{partall},$$

slik at påstanden er sann for $n = k + 1$.

Vi har derfor vist at alle fibonaccitall er partall. $\square$

Oppgave 5.21. Finn feilen i følgende induksjonsbevis. Her uttrykker $\max\{x, y\}$ det største av tallene x og y.

“Oppgave”: *Vis at for ethvert positivt heltall n, dersom x og y er positive heltall slik at* $\max\{x, y\} = n$*, da er* $x = y$*.*

"Løsning": Vi beviser påstanden ved matematisk induksjon.

Påstanden er opplagt sann for $n = 1$: dersom x og y er positive heltall slik at $\max\{x, y\} = 1$, da må både $x = 1$ og $y = 1$.

Anta nå at påstanden er sann for et heltall $n = k$, det vil si at hvis $\max\{x, y\} = k$ for positive heltall x og y, da er $x = y$. For å fullføre induksjonstrinnet, må vi vise at påstanden også er sann for $n = k + 1$. Anta dermed at $\max\{x, y\} = k + 1$ for to positive heltall x og y. Da er $\max\{x - 1, y - 1\} = k$. Ved induksjonshypotesen er dermed $x - 1 = y - 1$, som medfører at $x = y$. Vi har dermed vist at påstanden holder for $n = k+1$, hvis den holder for $n = k$.

Vi har dermed vist at påstanden holder for alle positive heltall n ved induksjon. $\square$

Oppgave 5.22. I denne oppgaven skal vi bevise

Fermats lille teorem: *Hvis p er et primtall, da har vi for ethvert heltall n at $n^p \equiv n \pmod p$.*

Vi lar derfor p være et fiksert primtall gjennom hele oppgaven.

(a) Vis at hvis $a, b \in \mathbb{Z}$ slik at $\frac{pa}{b} \in \mathbb{Z}$ og p ikke deler b, da vil p dele $\frac{pa}{b}$.

(b) Bruk resultatet fra (a) til å vise at p deler $\binom{p}{k}$ for enhver $k \in \mathbb{Z}$ slik at $0 < k < p$.

(c) Bruk resultatet fra (b) og *binomialteoremet* 3.8.7 til å vise at for alle $x, y \in \mathbb{Z}$ gjelder
$$(x + y)^p \equiv x^p + y^p \pmod p.$$

(d) Bruk induksjon og resultatet fra (c) til å vise at
$$n^p \equiv n \pmod p$$
for alle positive $n \in \mathbb{Z}$. (Hint: $n = (n - 1) + 1$.)

(e) Forklar hvorfor formelen du fant i (d) også gjelder hvis $n \in \mathbb{Z}$ med $n \leq 0$.

Oppgave 5.23. Denne oppgaven omhandler det velkjente *sladreproblemet*[7]. Anta at det er n personer i en gruppe og at hver har kjennskap til en skandale som ingen andre i gruppen kjenner til. Personene kommuniserer med telefon, og hver gang to av personene kommuniserer, deler de informasjon om alle skandalene hver kjenner til. Hva er da det minste antallet telefonsamtaler $G(n)$ som trengs for at alle n personer kjenner til alle skandalene?

(a) Finn $G(1)$, $G(2)$, $G(3)$ og $G(4)$.

[7] *Gossip problem* på engelsk.

(b) Bruk induksjon til å vise at $G(n) \leq 2n-4$ for $n \geq 4$. (Hint: i induksjonstrinnet, la en ny person ringe en bestemt person i begynnelsen og slutten.)
(c) Vis at $G(n) = 2n - 4$ for $n \geq 4$. (Denne delen er svært vanskelig.)

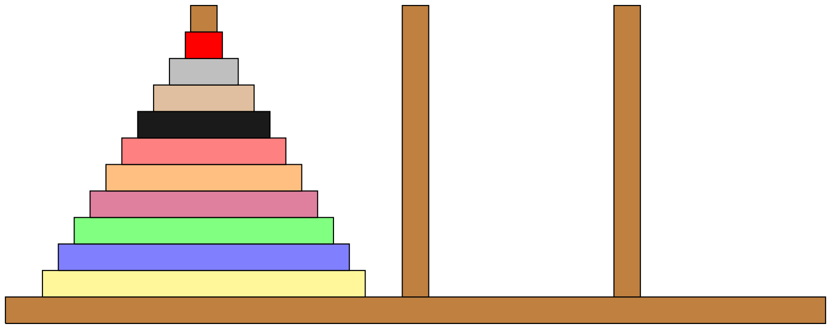
Oppgave 5.24. Denne oppgaven omhandler det velkjente *Josephus-problemet*. Det baserer seg på en skildring av historikeren Flavius Josephus (37–ca. 100), som var én av 41 jødiske opprørere i en hule som ble omringet av romerske soldater under den jødisk-romerske krigen i første århundre. Opprørerne foretrakk selvmord fremfor å bli fanget og bestemte seg for følgende prosedyre: de formet en sirkel, bestemte seg for en startposisjon og henrettet konsekutivt hver tredje mann som fremdeles var i live (på en slik måte at hver person ble henrettet av nestemann som skulle henrettes). Skildringen forteller at Josephus ble den siste i rekken og overga seg med nestsiste gjenlevende mann. (Noen historier forteller at dette var ren flaks, andre at Josephus, som ikke ønsket å dø, regnet ut hvor han skulle plassere seg i sirkelen.)

Det generelle *Josephus-problemet* er følgende: ordne n personer markert med $\{1, \ldots, n\}$ i en sirkel og henrett i tur og orden hver m-te gjenlevende person i sirkelen (start med m, $2m$, osv.) til det kun er én gjenlevende. Finn nummeret $J(n,m)$ på den siste gjenlevende. (I det opprinnelige historiske tilfellet er altså $n = 41$ og $m = 3$.)

La oss i resten av oppgaven betrakte tilfellet $m = 2$, hvor vi for enkelhets skyld setter $J(n) = J(n, 2)$.

(a) Finn $J(n)$ for $1 \leq n \leq 16$.
(b) Bruk resultatet fra (a) til å lage en formodning om hva $J(n)$ er for alle $n \in \mathbb{Z}^+$. (Hint: skriv $n = 2^l + k$, hvor $l, k \in \mathbb{N}$ og $k < 2^l$.)
(c) Vis at $J(n)$ tilfredsstiller *rekursjonsformlene* $J(2n) = 2J(n) - 1$ og $J(2n + 1) = 2J(n) + 1$ for $n \geq 1$ og $J(1) = 1$.
(d) Bevis formodningen i (b) ved induksjon og rekursjonsformlene fra (c).

Oppgave 5.25. *Tårnet i Hanoi* eller *Brahmas Tårn* er et matematisk spill som sies å ha blitt oppfunnet av den franske matematikeren Édouard Lucas (1842–1891) i 1883. Spillet består av tre pinner og en rekke runde skiver med et hull i midten. Skivene er av varierende bredde, og kan plasseres i en hvilken som helst av de tre pinnene. Spillet starter med alle de runde skivene plassert over en pinne, ordnet etter størrelse, med den minste øverst, som vist i figuren nedenunder.



Spillet går ut på å flytte alle skivene til en annen pinne, etter følgende regler:

- Bare én skive av gangen kan flyttes.
- Flyttingen foregår ved at den øverste skiven fra en av pinnene flyttes til en annen pinne og legges på toppen av andre skiver som allerede er der.
- Ingen skive kan plasseres over en mindre skive.

Hensikten med spillet er å få flyttet alle skivene fra en pinne til en annen med så få flyttinger som mulig.

I denne oppgaven skal vi regne ut minste antall flyttinger $H(n)$ når vi starter med $n \geq 1$ ringer.

(a) Vis at vi har $H(n) = 2H(n-1) + 1$ for $n \geq 2$. (Dette kalles en *rekursjonsformel*.)

(b) Bruk formelen i (a) og matematisk induksjon til å vise at

$$H(n) = 2^n - 1.$$

(c) Spillet tar utgangspunkt i følgende gamle legende, som finnes i flere varianter: Ved jordens begynnelse plasserte guden Brahma tre stolper i et tempel i Benares i India, verdens midtpunkt. På en av stolpene plasserte han 64 gullskiver, med den største nederst, og så ble skivene mindre og mindre oppover stolpen. Rundt år 3500 f.Kr. fikk munkene i byen i oppgave av guden å flytte alle ringene fra en stolpe til en annen ved å følge reglene gitt over. Når oppgaven var fullført skulle verden gå under og bli til støv.

Hvis vi antar at munkene klarer å flytte en skive i sekundet og aldri gjør noen feil, hvor lang tid vil det da ta før verden går under?

Om dere ønsker, kan dere spille på denne vevsiden

`http://www.superkids.com/aweb/tools/logic/towers/`

Oppgave 5.26. På en fjern øy bor det et lykkelig folkeslag. Øyen er så liten at alle kjenner alle. Beboerne er ekstremt intelligente og fornuftige, bortsett fra følgende: De regner det for å være veldig skamfullt å ha blå øyne. Hvis en beboer finner ut at han eller hun har blå øyne, vil han/hun øyeblikkelig forlate øyen og aldri komme tilbake.

Heldigvis finnes det ingen speil på øyen, ingen har kjennskap til genetikk og alle er altfor høflige til å nevne noens øyenfarge.

En dag ankommer en misjonær til øyen. Idet han holder på å dra, sier han til alle at han synes det er merkverdig at det finnes to forskjellige øyenfarger, blå og brune, i en såpass liten befolkning.

På førtitredje natt etter denne hendelsen, forlater mange beboere øyen.

Hvor mange beboere var blåøyde på øyen?

KAPITTEL 6

Rasjonale tall oppfyller ikke kompletthetsprinsippet

Matematiske problemer som involverer irrasjonale tall, det vil si tall som ikke er rasjonale, finnes i indiske skrifter allerede fra 800 f.Kr. Det er likevel grekerne som først systematisk studerte det vi i dag kaller rasjonale tall og som derfor krediteres for oppdagelsen av at det finnes tall utenom de rasjonale.

Som nevnt i introduksjonen til Kapittel 3 baserte grekerne sin geometri og mye av sin matematikk på teorien av *kommensurable linjestykker*, det vil si linjestykker hvor forholdet mellom lengdene kunne uttrykkes som et forhold mellom heltall. Med andre ord tenkte de seg at det alltid fantes et tredje, knøttlite linjestykke som gikk opp et helt antall ganger i to vilkårlig gitte linjestykker. Det kom dermed som et sjokk da de greske matematikerne ledet av Pythagoras (ca. 570 – ca. 495 f.Kr.) oppdaget at det fantes *inkommensurable linjestykker* trolig rundt år 450-500 f.Kr. Oppdagelsen er så omkranset av myter at det er umulig å vite akkurat hva som skjedde. En av mytene forteller at det var Hippasus fra Metapontum (ca. 530 – ca. 450 f.Kr.) som oppdaget at siden og diagonalen i et kvadrat er inkommensurable (som tilsvarer at $\sqrt{2}$ er irrasjonalt) eller at siden og diagonalen i en regulær femkant er det, og at han ble druknet av andre pythagoreere som straff for denne sjokkerende oppdagelsen.

Selv om eksistensen av inkommensurable linjestykker la grekernes matematikk tilsynelatende i grus, tok det ikke lang tid før Eudoxus av Knidos (ca. 390 – ca. 337 f.Kr.), en student av den velkjente filosofen Platon (ca. 425 – ca. 348 f. Kr.), løste problemet ved å bygge opp en ny teori der inkommensurable linjestykker kan tilnærmes med kommensurable. Denne teorien fant veien inn i Euklids *Elementer* fra ca. 300 f.Kr. Selv om denne teorien ikke er like rigorøs med hensyn på dagens krav, har den lagt mye av grunnlaget for dagens forståelse av rasjonale og irrasjonale tall. Det finnes imidlertid en vesentlig forskjell mellom denne teorien og den moderne teorien for reelle tall: mens vi i dag godtar irrasjonale tall på lik linje med de rasjonale, så skjelnet grekerne strengt mellom disse tallene.

Vi skal i dette kapitlet ta en nærmere, mer moderne, titt på de rasjonale tallene og hva som "mangler" ved dem. Det gjør vi først ved å innføre begrepet *skranker* for en mengde rasjonale tall (§6.1), og deretter ved å studere *følger* av rasjonale tall (§6.2-6.3).

6.1. Øvre og nedre skranker

Intuitivt kan vi tenke på en ordnet mengde (som for eksempel $\mathbb{Q}$) som en mengde der alle elementer kan ordnes i en linje der størrelsen øker mot høyre (eller motsatt), siden vi for alle par av elementer i mengden kan si hvilket av dem som er størst, ved *trikotomiegenskapen* til en ordensrelasjon (Definisjon 3.6.9 og Setning 3.5.2(i)).

Oppad begrensede mengder og øvre skranker. I en ordnet mengde kan vi snakke om delmengder som er *begrenset*:

Definisjon 6.1.1: Øvre skranke og oppad begrenset

La X være en ordnet mengde og $S \subset X$ en delmengde.

Vi sier at et element $M \in X$ er en *øvre skranke*[a] *for* S dersom $x \leq M$ for alle $x \in S$.

Vi sier at S er *oppad begrenset*[b] dersom en slik øvre skranke for S finnes.

[a] *Upper bound* på engelsk.
[b] *Bounded above* på engelsk.

Merk at definisjonen også fungerer for $S = X$, det vi si at vi kan snakke om hvorvidt "hele mengden" er oppad begrenset eller ikke.

Eksempel 6.1.2. Delmengden $\mathbb{N} \subset \mathbb{Q}$ er ikke oppad begrenset. Dette kan vises ved motsigelse: anta at det fantes en øvre skranke $M \in \mathbb{Q}$ for $\mathbb{N}$. Da er $M \geq 1$ og vi kan skrive $M = \frac{a}{b}$, med $a, b \in \mathbb{Z}^+$. Da er $a + 1 \in \mathbb{N} \subset \mathbb{Q}$ ved Peanoaksiom (N2) på side 87 og (42), og

$$a + 1 > a \geq \frac{a}{b} = M,$$

som motsier at M er en øvre skranke for $\mathbb{N}$.

Av samme grunn er heller ikke $\mathbb{Z}$ eller $\mathbb{Q}$ oppad begrenset.

La istedenfor $\mathbb{Z}^- \subset \mathbb{Z}$ være delmengden av alle negative heltall. Da vil -1 være en øvre skranke for $\mathbb{Z}^-$; faktisk vil ethvert heltall $n \geq -1$ være en øvre skranke. Derfor er $\mathbb{Z}^-$ oppad begrenset.

På samme måte er delmengden $\mathbb{Q}^- \subset \mathbb{Q}$ av alle negative rasjonale tall oppad begrenset. Her er 0 en øvre skranke; faktisk er ethvert rasjonalt tall $q \geq 0$ en øvre skranke.

Merknad 6.1.3. La $r \in \mathbb{Q}^{\geq 0}$. Siden $\mathbb{N}$ ikke er oppad begrenset, er delmengden $\{n \in \mathbb{N} \mid n \geq r\} \subset \mathbb{N}$ ikke-tom. Ved *velordningsprinsippet* har denne mengden et minste element. Vi betegner det med $\lceil r \rceil$ og kaller det *opprundingen*[1] av r. Altså er $\lceil r \rceil$ det minste heltallet n som oppfyller $n \geq r$.

[1] *Round-up* på engelsk.

Eksempelvis er $\lceil \frac{3}{2} \rceil = 2$ og $\lceil \frac{27}{4} \rceil = 7$.

EKSEMPEL 6.1.4. Betrakt delmengden

$$S = \{q \in \mathbb{Q} \mid q < 0 \ \text{ eller } \ q^2 < 2\}.$$

Denne mengden er ikke-tom, siden for eksempel $1 \in S$. Alle $q \in S$ må oppfylle at $q < 2$. For hvis $q \geq 2$, da vil

$$q^2 = q \cdot q \geq 2 \cdot q \geq 2 \cdot 2 = 4$$

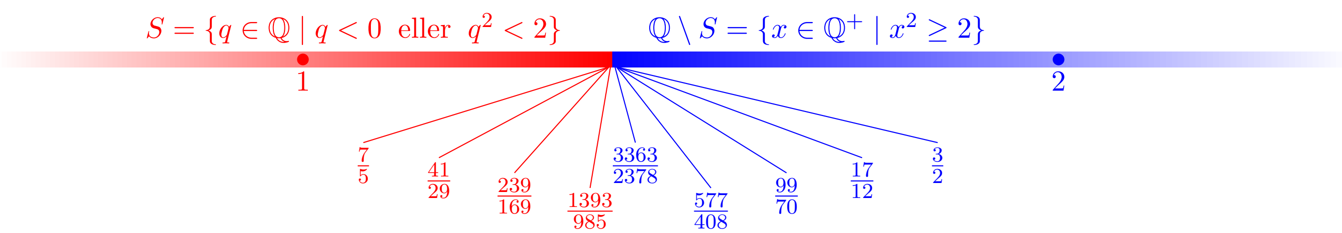
ved Setning 3.6.13 i begge ulikhetene, som medfører at $q \notin S$. (Man kan også bruke resultatet fra Oppgave 5.9 til å dedusere at $q^2 \geq 4$ om $q \geq 2$.) Det følger at 2 er en øvre skranke for S og S er oppad begrenset.

Vi kan finne øvre skranker for S som er mindre enn 2. Mer spesifikt kan vi vise at

(85) enhver $x \in \mathbb{Q}^+$ med $x^2 \geq 2$ er en øvre skranke for S.

Vi må altså vise at hvis $y \in S$, da er $y \leq x$. Anta dette ikke stemmer, det vil si $y > x$. Da er $y^2 > x^2 \geq 2$, som motsier at $y \in S$. Denne motsigelsen beviser (85).

Figuren nedenunder viser mengden S i rødt og *komplementet* $\mathbb{Q} \setminus S$ i blått. Ved (85) er alle tall i $\mathbb{Q} \setminus S$ en øvre skranke for S.



Den minste øvre skranken. Vi kan som sett ovenfor gjerne finne mange øvre skranker. Den minste slike, om den finnes, har et eget navn:

Definisjon 6.1.5: Minste øvre skranke

La X være en ordnet mengde og $S \subset X$ en oppad begrenset delmengde.

Et element $m \in X$ kalles en *minste øvre skranke*[a] *for* S eller *supremum*[b] *til* S dersom følgende er oppfylt:

(i) m er en øvre skranke for S (det vil si $m \geq x$ for alle $x \in S$);
(ii) hvis $m' \in X$ og $m' < m$, da er m' ikke en øvre skranke for S.

Vi betegner m med $\sup S$.

[a] *Least upper bound* på engelsk.
[b] Fra latin.

Selv om det ikke er klart om en slik minste øvre skranke finnes, er det klart fra del (ii) av definisjonen og *trikotomiegenskapen* til en ordensrelasjon (Definisjon 3.6.9) at det kan finnes høyst én minste øvre skranke.

EKSEMPEL 6.1.6. I Eksempel 6.1.2 er $\sup \mathbb{Z}^- = -1$ og $\sup \mathbb{Q}^- = 0$. Vi bemerker at $\sup \mathbb{Z}^- \in \mathbb{Z}^-$, mens $\sup \mathbb{Q}^- \notin \mathbb{Q}^-$. Supremum til en mengde kan altså både være et element i mengden og ikke.

EKSEMPEL 6.1.7. La

$$S = \left\{1 - \frac{1}{n} \,\middle|\, n \in \mathbb{Z}^+\right\} = \left\{0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5} \dots\right\} \subset \mathbb{Q}.$$

Vis at $\sup S = 1$.

LØSNING. Det er klart at 1 er en øvre skranke for S, det vil si at (i) i Definisjon 6.1.5 er oppfylt, siden alle $x \in S$ oppfyller $x < 1$. Vi må da vise at (ii) i Definisjon 6.1.5 er oppfylt. La derfor $q \in \mathbb{Q}$ være slik at $q < 1$. Vi hevder at det finnes en $n \in \mathbb{Z}^+$ slik at

$$1 - \frac{1}{n} > q. \tag{86}$$

Siden $1 - \frac{1}{n} \in S$, viser dette at q ikke er en øvre skranke for S, slik at (ii) i Definisjon 6.1.5 er oppfylt.

For å vise at det finnes en $n \in \mathbb{Z}^+$ slik at (86) er oppfylt, omformer vi denne til ekvivalente ulikheter

$$\begin{aligned} 1 - \frac{1}{n} &> q \\ 1 &> q + \frac{1}{n} \\ 1 - q &> \frac{1}{n} > 0 \\ n &> \frac{1}{1-q} \qquad (87) \end{aligned}$$

Her har vi brukt flere av egenskapene i Definisjon 3.6.12 og Setning 3.6.13 som vi overlater til Oppgave 6.4 å sjekke.

At det finnes et positivt heltall n som oppfyller (87) følger av at $\mathbb{N}$ ikke er oppad begrenset, jf. Eksempel 6.1.2. $\square$

Kompletthetsprinsippet. Mengder som har den spesielle egenskapen at alle oppad begrensede delmengder har en minste øvre skranke er såpass viktige at de har fått sitt eget navn:

Definisjon 6.1.8

En ordnet mengde X sies å ha *minste-øvre-skranke egenskapen*[a], eller å oppfylle *kompletthetsprinsippet*, dersom enhver ikke-tom oppad begrenset delmengde av X har en minste øvre skranke i X.

[a] *Least upper bound property* på engelsk.

Neste resultat er ekstremt viktig, fordi det forteller oss at $\mathbb{Q}$ mangler denne egenskapen:

Setning 6.1.9

Mengden $\mathbb{Q}$ oppfyller ikke kompletthetsprinsippet.

Bevis

Vi vil vise at delmengden $S = \{q \in \mathbb{Q} \mid q < 0 \;\text{ eller }\; q^2 < 2\}$ fra Eksempel 6.1.4, som vi viste er ikke-tom og oppad begrenset, ikke har en minste øvre skranke i $\mathbb{Q}$.

Vi argumenterer ved selvmotsigelse og antar dermed at $p = \sup S \in \mathbb{Q}$. Siden $1 \in S$, må $p \geq 1$. Videre må $p^2 \neq 2$ ved Setning 1.3.5. Definér

$$q = p - \frac{p^2 - 2}{p + 2} = \frac{2p + 2}{p + 2}.$$

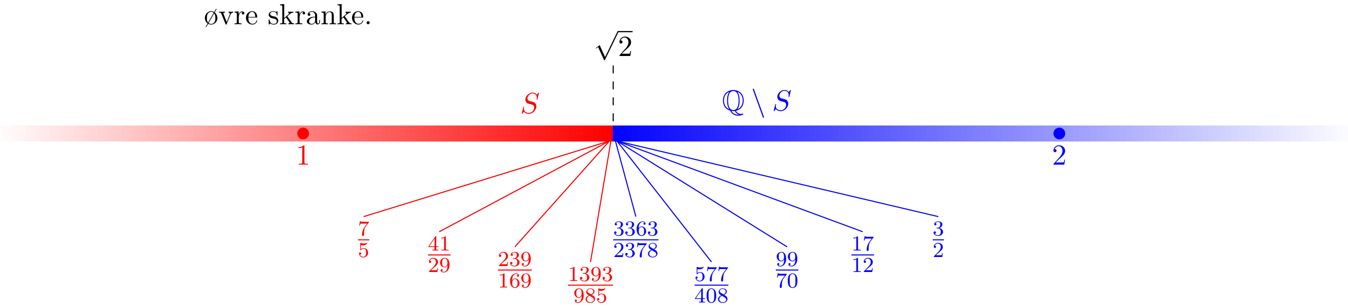
Da er $q \in \mathbb{Q}^+$ og

$$q^2 - 2 = \frac{2(p^2 - 2)}{(p + 2)^2}.$$

Hvis $p^2 < 2$, da er $p < q$ og $q^2 < 2$. Dette medfører at $q \in S$ og dermed at p ikke er en øvre skranke for S, en selvmotsigelse.

Hvis $p^2 > 2$, da er $p > q$ og $q^2 > 2$. Dette medfører at $q \notin S$ og dermed at q er en øvre skranke for S ved (85). Men siden $q < p$, motsier dette at p er den *minste* øvre skranken til S. $\square$

Denne setningen har følgende intuitive tolkning: de rasjonale tallene har "hull", fordi vi har mengder som er oppad begrenset uten en minste øvre skranke. Som vi vil se senere, så har vi ikke samme problem i de reelle tallene, for der vil mengden S fra siste bevis ha tallet $\sqrt{2}$ som en minste øvre skranke.



Mengden S i rødt; mengden $\mathbb{Q} \setminus S$ i blått. I $\mathbb{R}$ er $\sup S = \sqrt{2}$.

Som vi skal se i de to neste seksjonene, henger dette problemet med "hull" i de rasjonale tallene sammen med eksistensen av følger i $\mathbb{Q}$ som konvergerer mot "noe som ligger utenfor $\mathbb{Q}$".

Nedad begrensede mengder og nedre skranker. Før vi går videre er det naturlig å stille seg følgende spørsmål: Vi har sett på oppad begrensede mengder og eksistensen av minste øvre skranker for slike mengder. Hva med det "motsatte", nemlig *nedad begrensede mengder*, finnes det tilsvarende definisjoner og prinsipper for disse? Svaret er ja, og teorien for disse er helt "symmetrisk". Spesielt sier Teorem 6.1.14 nedenunder at prinsippene er ekvivalente, slik at det egentlig holder å utlede en teori kun for øvre skranker som vi har gjort hittil. For å være så fullstendige som mulig, vil vi i resten av seksjonen se på nedad begrensede mengder og nedre skranker.

Definisjon 6.1.10: Nedre skranke og nedad begrenset mengde

La X være en ordnet mengde og $S \subset X$ en delmengde.

Vi sier at et element $M \in X$ er en *nedre skranke*[a] *for* S dersom $x \geq M$ for alle $x \in S$.

Vi sier at S er *nedad begrenset*[b] dersom en slik nedre skranke for S finnes.

[a] *Lower bound* på engelsk.

[b] *Bounded below* på engelsk.

Definisjon 6.1.11: Begrenset mengde

Vi sier at en delmengde S av en ordnet mengde er *begrenset*[a] dersom den både er oppad og nedad begrenset.

[a] *Bounded* på engelsk.

Definisjon 6.1.12: Største nedre skranke

La X være en ordnet mengde og $S \subset X$ en nedad begrenset delmengde.

Et element $m \in X$ kalles en *største nedre skranke*[a] *for* S eller *infimum*[b] *til* S dersom følgende er oppfylt:

(i) m er en nedre skranke for S (det vil si $m \leq x$ for alle $x \in S$);

(ii) hvis $m' \in X$ og $m' > m$, da er m' ikke en nedre skranke for S.

Vi betegner m med $\inf S$.

[a] *Greatest lower bound* på engelsk.

[b] Fra Latin.

Definisjon 6.1.13: Største-nedre-skranke egenskapen

En ordnet mengde X sies å ha *største-nedre-skranke egenskapen*[a] dersom enhver ikke-tom nedad begrenset delmengde av X har en største nedre skranke i X.

[a] *Greatest lower bound property* på engelsk.

Teorem 6.1.14

En ordnet mengde X har minste-øvre-skranke egenskapen hvis og bare hvis den har største-nedre-skranke egenskapen.

Dette beviset er litt tungt og kan hoppes over uten tap for videre lesing.

Bevis for Teorem 6.1.14

Anta at X har minste-øvre-skranke egenskapen. For å vise at X også har største-nedre-skranke egenskapen, lar vi $S \subset X$ være en vilkårlig nedad begrenset mengde, og viser at S da har en største nedre skranke.

For å vise dette lar vi

$$N = \{x \in X \mid x \text{ er en nedre skranke for } S\}.$$

Vi vil vise at $\sup N$ eksisterer og at $\sup N = \inf S$.

Siden S er nedad begrenset, har S en nedre skranke. Dermed er $N \neq \emptyset$. Siden N består av alle $x \in X$ slik at $x \leq s$ for alle $s \in S$, ser vi at

$$\text{enhver } s \in S \text{ er en øvre skranke for } N. \tag{88}$$

Altså er N oppad begrenset. Siden X har minste-øvre-skranke egenskapen, finnes en minste øvre skranke $\sup N$ for N.

Hvis $y < \sup N$, da er y ikke en øvre skranke for N. Det følger av (88) at $y \notin S$. Dermed må $y \geq \sup N$ for alle $y \in S$. Altså er $\sup N$ en nedre skranke for S, hvilket vil si at $\sup N \in N$.

Hvis $y > \sup N$, er $y \notin N$, siden $\sup N$ er en øvre skranke for N.

Vi har derfor vist at

(i) $\sup N \in N$,
(ii) $y \notin N$, hvis $y > \sup N$,

som kan omformuleres, via definisjonen av N, som

(i) $\sup N$ er en nedre skranke for S,
(ii) y er ikke en nedre skranke for S, hvis $y > \sup N$.

Disse to egenskapene gir per definisjon at $\sup N$ er den største nedre skranken for S, altså $\sup N = \inf S$, som er det vi ønsket å vise.

Beviset for at X har minste-øvre-skranke egenskapen, hvis den har største-nedre-skranke egenskapen, er temmelig likt. □

6.2. Følger av rasjonale tall

For å illustrere hva som mangler med de rasjonale tallene $\mathbb{Q}$, kan vi også se på følger av rasjonale tall. Vi vil i denne seksjonen gjennomgå grunnleggende definisjoner og egenskaper, for så å komme tilbake til manglene i $\mathbb{Q}$ i neste seksjon.

Helt generelt er en følge av elementer i en mengde X en uendelig sekvens av elementer

$$a_1, a_2, a_3, a_4, \ldots$$

i mengden X, indeksert av de positive heltallene $\mathbb{Z}^+$. Selvsagt er det ikke så farlig om vi begynner følgen med indeksen 0 istedenfor 1, så vi kunne likegodt sagt at indekseringen er i $\mathbb{N}$, eller for den del kan vi starte på et hvilket som helst tall i $\mathbb{Z}$. Elementene a_n kan godt være like for forskjellige n, f.eks er

$$1, 1, 1, 1, 1, 1 \ldots \tag{89}$$

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots \tag{90}$$

$$1, -1, 1, -1, 1, -1, \ldots \tag{91}$$

alle eksempler på følger i $\mathbb{Q}$.

I matematikk ønsker vi imidlertid å være noe mer presise og gir følgende ekvivalente, men mer nøyaktige definisjon av en følge:

Definisjon 6.2.1: Følge

La X være en mengde. En *følge*[a] *i* X er en funksjon $f : \mathbb{Z}^+ \to X$. Hvis $f(n) = a_n$ for $n \in \mathbb{Z}^+$, betegner vi gjerne følgen som

$$a_1, a_2, a_3, \ldots,$$

eller kortere som $\{a_n\}$.

Elementene $a_n \in X$ kalles *leddene*[b] i/til følgen.

[a] *Sequence* på engelsk.

[b] *Terms* på engelsk.

Eksempel 6.2.2. (a) Følgen i (89) kan med notasjonen i denne definisjonen uttrykkes som funksjonen

$$\begin{array}{ccc} \mathbb{Z}^+ & \longrightarrow & \mathbb{Q} \\ n & \mapsto & 1 \end{array}$$

eller som (den konstante følgen) $\{1\}$.

(b) Følgen i (90) kan uttrykkes som funksjonen

$$\begin{array}{ccc} \mathbb{Z}^+ & \longrightarrow & \mathbb{Q} \\ n & \mapsto & \dfrac{1}{n} \end{array}$$

eller som $\{\frac{1}{n}\}$.

(c) Følgen i (91) kan uttrykkes som funksjonen

$$\begin{array}{ccc} \mathbb{Z}^+ & \longrightarrow & \mathbb{Q} \\ n & \mapsto & (-1)^{n+1} \end{array}$$

eller som $\{(-1)^{n+1}\}$.

Konvergens av rasjonale følger. Vi merker at leddene i følgen (90) nærmer seg tallet 0 etter hvert som n øker, mens leddene i følgen (91) ikke nærmer seg noe bestemt tall. Vi sier at følgen (90) *konvergerer (mot grenseverdien* 0*)*, mens følgen (91) *divergerer*.

Hva mener vi helt nøyaktig med at følgen (90) konvergerer mot 0? Jo, med det mener vi at vi kan få leddene $a_n = \frac{1}{n}$ til å være så nær vi vil 0 når vi kommer langt nok ut i følgen, det vil si ved å velge n stor nok. Avstanden mellom to tall er gitt ved differansen mellom det største og det minste tallet, det vil si ved absoluttverdien av differansen mellom tallene. Da betyr det at vi kan få $|a_n - 0| = |\frac{1}{n}| = \frac{1}{n}$ så liten vi vil ved å ta n stor nok. Ønsker vi for eksempel at $\frac{1}{n} < \frac{1}{10}$, kan vi sørge for at denne ulikheten gjelder ved å la $n > 10$. Ønsker vi for eksempel at $\frac{1}{n} < 5 \cdot 10^{-100}$, tar vi $n > \frac{10^{100}}{5}$. Det er klart hvordan det generelle bildet er: uansett hvor lite tall $\epsilon > 0$ vi får oppgitt (hvor ϵ står for den greske bokstaven "epsilon"), så kan vi sørge for at $\frac{1}{n} < \epsilon$ ved å velge $n > \frac{1}{\epsilon}$. Dette er bakgrunnen for følgende definisjon, hvor vi lar ϵ være et positivt rasjonalt tall, siden vi enda ikke har definert reelle tall:

Definisjon 6.2.3: Konvergens av følge

Følgen $\{a_n\}$ av rasjonale tall *konvergerer*[a] mot det rasjonale tallet L, dersom det for ethvert rasjonalt tall $\epsilon > 0$ finnes et tall $N \in \mathbb{Z}^+$ slik at $|a_n - L| < \epsilon$ for alle $n \geq N$.

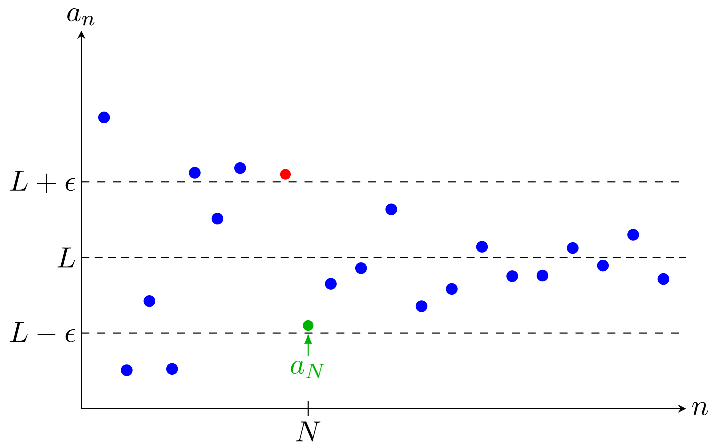
Vi skriver i så fall $\lim\limits_{n\to\infty} a_n = L$ eller "$a_n \to L$ når $n \to \infty$" og kaller L for *grensen*[b] til følgen.

Hvis følgen $\{a_n\}$ ikke konvergerer, sier vi at den *divergerer*[c].

[a] *Converge* på engelsk.
[b] *Limit* på engelsk.
[c] *Diverge* på engelsk.

Det er klart fra definisjonen at en N som fungerer for en gitt ϵ fungerer også for alle større ϵer. Kunsten er altså å finne N uansett hvor liten ϵ er. Selv om bokstaven ϵ kan brukes som symbol for alt mulig, er det blitt så vanlig å la ϵ betegne et "vilkårlig lite tall" at det er slik enhver matematiker vil bruke den. Dette er grunnlaget for at matematikere gjerne vil svare på spørsmål som "Vil du ha mer kaffe?" med "En ϵ, takk!", samtidig som det har gitt opphav til en del vitser som bare matematikere kan skjønne, som tittelen på denne boken! En annen slik vits, som bare matematikere synes er morsom, er "$\epsilon < 0$", også omtalt som "verdens korteste vits".

Hvis vi tegner grafen til funksjonen som definerer følgen, det vil si vi plotter alle punktene $(n, a_n) \subset \mathbb{Z}^+ \times \mathbb{Q}$ i planet der den horisontale aksen representerer $\mathbb{Z}^+$ og den vertikale aksen representerer $\mathbb{Q}$, kan konvergens visualiseres ved følgende figur: alle leddene a_n i følgen for $n \geq N$ havner i avstand mindre enn ϵ fra L. En annen ϵ ville gitt en annen N.



Konvergens av en følge

Merknad 6.2.4. Ulikheten $|a_n - L| < \epsilon$ er ekvivalent med ulikhetene $-\epsilon < a_n - L < \epsilon$ og $L - \epsilon < a_n < L + \epsilon$, og dette kommer vi til å bruke flere ganger.

Merknad 6.2.5. Ønsker man å komprimere teksten (jf. Merknad 2.1.4) kan man bruke symbolet $\forall$, som betyr "for alle" eller "for enhver", sammen med det allerede nevnte symbolet $\exists$ som betyr "eksisterer", og uttrykke "for ethvert rasjonalt tall $\epsilon > 0$ finnes et tall $N \in \mathbb{Z}^+$ slik at $|a_n - L| < \epsilon$ for alle $n \geq N$" ved

$$\forall \epsilon \in \mathbb{Q}^+ \;\; \exists N \in \mathbb{Z}^+ \; : \; |a_n - L| < \epsilon \;\; \forall n \geq N.$$

Før vi går i gang med å se på hvordan vi bruker denne definisjonen, er det viktig å spørre seg om en grense, slik den er definert, er *entydig*. Med andre ord, finnes det flere tall L som oppfyller betingelsene til å være en grense? (I så fall ville ikke definisjonen være svært nyttig.) Svaret er heldigvis nei, grenser er entydige:

Setning 6.2.6: Entydighet av grenser

Dersom en følge av rasjonale tall konvergerer mot L og L', da er $L = L'$.

Bevis
Siden beviset er litt teoretisk, er det best å vente med det til man har lest beviset for Setning 6.2.14. Beviset overlates derfor til Oppgave 6.8. $\square$

La oss bruke definisjonen først i tilfellene representert ved (89) og (90) ovenfor.

Hjelpesetning 6.2.7

Vi har

(a) $\lim_{n\to\infty} \frac{1}{n} = 0$,

(b) $\lim_{n\to\infty} c = c$, for enhver $c \in \mathbb{Q}$. (Leddene i følgen $\{c\}$ er altså konstant lik c, og vi kaller den derfor gjerne *den konstante følgen* $\{c\}$.)

Bevis
(a) (Vi repeterer det vi skrev før Definisjon 6.2.3.) Gitt $\epsilon \in \mathbb{Q}^+$. La N være et hvilket som helst heltall slik at $N > \frac{1}{\epsilon}$. (Et slikt heltall finnes siden $\mathbb{N}$ ikke er oppad begrenset, jf. Eksempel 6.1.2.) For eksempel kan vi la N være det minste heltallet slik at $N \geq \frac{1}{\epsilon} + 1$, altså $N = \left\lceil \frac{1}{\epsilon} + 1 \right\rceil$ (jf. Merknad 6.1.3). Hvis $n \geq N$, er $|\frac{1}{n} - 0| = \frac{1}{n} \leq \frac{1}{N} < \epsilon$. Dette viser at $\lim_{n\to\infty} \frac{1}{n} = 0$.

(b) Gitt $\epsilon \in \mathbb{Q}^+$. La $N = 1$. Hvis $n \geq N = 1$, er $|c - c| = 0 < \epsilon$. Dette viser at $\lim_{n\to\infty} c = c$. $\square$

Merknad 6.2.8. Som vi ser i del (a) av siste eksempel, vil tallet N generelt være avhengig av ϵ: jo mindre ϵ er, desto større må N være. Dette betyr, som allerede antydet i figuren på forrige side, at vi må gå lengre ut i følgen for å komme nærmere grensen L.

Avhengigheten av ϵ gjør at noen matematiske tekster betegner N i Definisjon 6.2.3 med $N(\epsilon)$. Avhengigheten betyr imidlertid ikke at N *må* variere når ϵ varierer: i del (b) av siste eksempel, der følgen er konstant, har vi for eksempel at $N = 1$ for alle ϵ.

La oss vise et mer komplisert eksempel der vi må tenke litt.

Eksempel 6.2.9. Vis at $\lim_{n\to\infty} \frac{4n}{3n+2} = \frac{4}{3}$.

Løsning. Gitt $\epsilon \in \mathbb{Q}^+$. Vi vil finne $N \in \mathbb{Z}^+$ slik at

$$\left|\frac{4n}{3n+2} - \frac{4}{3}\right| < \epsilon \text{ når } n \geq N. \tag{92}$$

Vi starter med å jobbe med ulikheten til venstre og omskriver først:

$$\left|\frac{4n}{3n+2} - \frac{4}{3}\right| = \left|\frac{4n \cdot 3 - 4(3n+2)}{3(3n+2)}\right| = \left|\frac{-8}{3(3n+2)}\right| = \frac{8}{3(3n+2)},$$

slik at ulikheten til venstre i (92) kan skrives som

$$\frac{8}{3(3n+2)} < \epsilon. \tag{93}$$

Vi isolerer n på den ene siden av ulikheten ved å omskrive som:

$$\begin{aligned} \frac{8}{3\epsilon} &< 3n+2 \\ \frac{8}{3\epsilon} - 2 &< 3n \\ \frac{8}{9\epsilon} - \frac{2}{3} &< n. \end{aligned}$$

Det vi har gjort viser at

$$\left|\frac{4n}{3n+2} - \frac{4}{3}\right| < \epsilon \Longleftrightarrow n > \frac{8}{9\epsilon} - \frac{2}{3}. \tag{94}$$

Vi lar derfor N være et hvilket som helst positivt heltall slik at $N > \frac{8}{9\epsilon} - \frac{2}{3}$, for eksempel kan vi la N være det minste heltallet slik at $N \geq \frac{8}{9\epsilon}$, med andre ord $N = \left\lceil \frac{8}{9\epsilon} \right\rceil$ (jf. Merknad 6.1.3). Da, hvis $n \geq N$, er $n \geq \frac{8}{9\epsilon} > \frac{8}{9\epsilon} - \frac{2}{3}$, slik at (94) viser at $\left|\frac{4n}{3n+2} - \frac{4}{3}\right| < \epsilon$.

Vi kan også gjøre et grovere overslag, siden vi ikke trenger å finne den minst mulige N vi trenger. Lar vi for eksempel $N = \left\lceil \frac{1}{\epsilon} \right\rceil$, det vil si N er det minste heltallet større eller lik $\frac{1}{\epsilon}$, har vi at hvis $n \geq N$, så er

$$n \geq \left\lceil \frac{1}{\epsilon} \right\rceil \geq \frac{1}{\epsilon} > \frac{8}{9\epsilon} > \frac{8}{9\epsilon} - \frac{2}{3},$$

slik at (94) igjen viser at $\left|\frac{4n}{3n+2} - \frac{4}{3}\right| < \epsilon$. □

Vi merker oss at det ikke er vanlig i lærebøker eller matematiske artikler å føre inn hele resonnementet som leder oss til den ønskede N. Vanligvis vil man i slike tilfeller starte med å oppgi den N man har funnet og vise at denne oppfyller den ønskede egenskapen. Vi viser hvordan dette gjøres:

LØSNING ("LÆREBOKVERSJON"). Gitt $\epsilon \in \mathbb{Q}^+$. La $N = \left\lceil \frac{1}{\epsilon} \right\rceil$. Hvis $n \geq N$, har vi $n \geq \frac{1}{\epsilon}$, og dermed

$$\begin{aligned}\left|\frac{4n}{3n+2} - \frac{4}{3}\right| &= \left|\frac{4n \cdot 3 - 4(3n+2)}{3(3n+2)}\right| = \left|\frac{8}{3(3n+2)}\right| \\ &= \frac{8}{3(3n+2)} < \frac{9}{3 \cdot 3n} = \frac{1}{n} \leq \epsilon.\end{aligned}$$

□

Ulempen med dette siste beviset er at leseren ikke får et innblikk i hvordan man kom på at $N = \left\lceil \frac{1}{\epsilon} \right\rceil$ ville fungere.

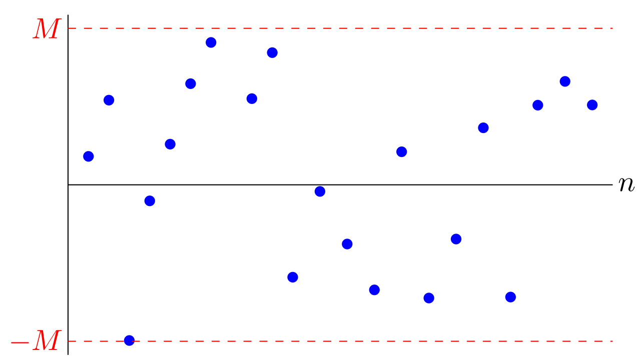
Vi tar med følgende resultat som viser en viktig konsekvens av konvergens, nemlig *begrensethet* av en følge. Vi trenger først en definisjon.

Definisjon 6.2.10: Begrenset følge

En følge $\{a_n\}$ av rasjonale tall er *begrenset*[a] dersom det finnes en $M \in \mathbb{Q}$ slik at $|a_n| \leq M$ for alle $n \in \mathbb{Z}^+$.

En følge er *ubegrenset*[b] dersom den ikke er begrenset.

[a] *Bounded* på engelsk.
[b] *Unbounded* på engelsk.



EN BEGRENSET FØLGE

Setning 6.2.11

En konvergent følge av rasjonale tall er begrenset.

BEVIS
Ideen bak beviset er enkel: siden følgen er konvergent, si mot a, vil alle ledd for store nok n havne i vilkårlig liten avstand fra a, for eksempel i avstand 1 fra a. Altså er absoluttverdien til alle slike ledd $\leq |a| + 1$. De resterende leddene er endelig mange og er derfor ikke noe problem. La oss skrive dette skikkelig:

La $\{a_n\}$ være en følge av rasjonale tall som konvergerer mot $a \in \mathbb{Q}$. Per definisjon av konvergens vet vi at det finnes en $N \in \mathbb{Z}^+$ slik at $|a_n - a| < 1$ når $n \geq N$. (Her har vi brukt definisjonen for konvergens med $\epsilon = 1$, men vi kunne ha brukt en hvilken som helst annen fast ϵ.) Trekantulikheten gir at $|a_n| = |a_n - a + a| \leq |a_n - a| + |a| < 1 + |a|$ for alle $n \geq N$. Da vil

$$|a_n| \leq \max\{|a_1|, \ldots, |a_{N-1}|, 1 + |a|\}$$

for alle $n \in \mathbb{Z}^+$, hvor "*max*" uttrykker det største av tallene i mengden. Dette viser at $\{a_n\}$ er begrenset. $\square$

MERKNAD 6.2.12. Noen spør seg kanskje hvorfor vi går veien om å finne N først. Kunne vi ikke droppet denne delen og konkludert med $|a_n| \leq \max\{|a_1|, |a_2|, |a_3|, \ldots\}$? Nei, for maksimum (=det største) av uendelig mange tall er ikke definert, mens maksimum av endelig mange tall er det. Derfor må vi først finne N.

EKSEMPEL 6.2.13. Følgen $\{n\}$ divergerer, siden den er ubegrenset, fordi $\mathbb{N}$ er det (Eksempel 6.1.2).

Regneregler for grenser. Det er ganske tungvint å bruke Definisjon 6.2.3 for å undersøke konvergens av hver eneste følge man treffer på. Heldigvis finnes det regneregler som kan brukes for å lette arbeidet:

Setning 6.2.14: Regneregler for grenser av følger

La $\{a_n\}$ og $\{b_n\}$ være konvergente følger av rasjonale tall. La $a = \lim_{n\to\infty} a_n$ og $b = \lim_{n\to\infty} b_n$. Da gjelder:

(i) $\lim_{n\to\infty} (a_n + b_n) = a + b$;

(ii) $\lim_{n\to\infty} (a_n - b_n) = a - b$;

(iii) $\lim_{n\to\infty} (ca_n) = ca$ for alle $c \in \mathbb{Q}$;

(iv) $\lim_{n\to\infty} (a_n b_n) = ab$;

(v) $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}$, dersom $b \neq 0$ og $b_n \neq 0$ for alle $n \in \mathbb{Z}^+$.

MERKNAD 6.2.15. Selv om reglene virker intuitivt opplagte, er det lurt å sette seg inn i argumentene som brukes i bevisene nedenunder, siden det er typiske argumenter som brukes i bevisstrategier på langt mer kompliserte setninger i matematikk. Legg også merke til hvordan *trekantulikheten* fra Setning 3.6.16(ii) brukes i bevisene nedenunder.

Bevis for Setning 6.2.14 (lang, resonnerende versjon)

(i) Gitt $\epsilon \in \mathbb{Q}^+$. Vi vil finne en $N \in \mathbb{Z}^+$ slik at

$$|(a_n + b_n) - (a + b)| < \epsilon \ \text{ når } \ n \geq N. \tag{95}$$

Det vi vet er at siden $a = \lim_{n\to\infty} a_n$ og $b = \lim_{n\to\infty} b_n$, så kan vi gjøre både $|a_n - a|$ og $|b_n - b|$ så små vi vil for stor nok n. Derfor bruker vi trekantulikheten til å skrive

$$|(a_n + b_n) - (a + b)| = |(a_n - a) + (b_n - b)| \leq |a_n - a| + |b_n - b|. \tag{96}$$

Vi merker derfor at vi får den ønskede ulikheten $|(a_n + b_n) - (a + b)| < \epsilon$ dersom for eksempel både $|a_n - a| < \frac{\epsilon}{2}$ og $|b_n - b| < \frac{\epsilon}{2}$.

Siden $\lim_{n\to\infty} a_n = a$, finnes en $N_1 \in \mathbb{Z}^+$ slik at $|a_n - a| < \frac{\epsilon}{2}$ når $n \geq N_1$. Likeledes, siden $\lim_{n\to\infty} b_n = b$, finnes en $N_2 \in \mathbb{Z}^+$ slik at $|b_n - b| < \frac{\epsilon}{2}$ når $n \geq N_2$. Velger vi N lik det største av tallene N_1 og N_2, som uttrykkes som $N = \max\{N_1, N_2\}$, da vil $n \geq N$ medføre at $n \geq N_1$ og $n \geq N_2$, slik at $|a_n - a| < \frac{\epsilon}{2}$ og $|b_n - b| < \frac{\epsilon}{2}$. Følgelig gir (96) at $|(a_n + b_n) - (a + b)| < \epsilon$. Vi har derfor vist (95) for $N = \max\{N_1, N_2\}$ og beviset er fullført.

(ii) Dette beviset er nesten helt likt det forrige, ved å bruke Merknad 3.6.17, og overlates til Oppgave 6.6.

(iii) Dette følger fra regel (iv) med $\{b_n\}$ lik den konstante følgen $\{c\}$.

(iv) Gitt $\epsilon \in \mathbb{Q}^+$. Vi vil finne en $N \in \mathbb{Z}^+$ slik at

$$|a_n b_n - ab| < \epsilon \ \text{ når } \ n \geq N. \tag{97}$$

Igjen vil vi jobbe med uttrykket $|a_n b_n - ab|$ og spalte det opp i deler som inneholder $|a_n - a|$ og $|b_n - b|$, siden vi vet at vi kan gjøre begge disse avstandene så små vi vil for stor nok n. Trikset er å legge til å trekke fra samme ledd $a_n b$ og deretter bruke trekantulikheten (Setning 3.6.16(ii)) og Setning 3.6.16(i):

$$\begin{aligned} |a_n b_n - ab| &= |a_n b_n - a_n b + a_n b - ab| \\ &\leq |a_n b_n - a_n b| + |a_n b - ab| \\ &= |a_n||b_n - b| + |b||a_n - a|. \end{aligned} \tag{98}$$

Vi merker derfor at vi får den ønskede ulikheten $|a_n b_n - ab| < \epsilon$ dersom vi klarer å få begge leddene $|a_n||b_n - b| < \frac{\epsilon}{2}$ og $|b||a_n - a| < \frac{\epsilon}{2}$.

Vi ser først på leddet $|b||a_n - a|$. Dersom $b = 0$ er det ingenting å vise. Vi kan derfor anta at $b \neq 0$. Siden $\lim_{n\to\infty} a_n = a$, finnes en $N_1 \in \mathbb{Z}^+$ slik at $|a_n - a| < \frac{\epsilon}{2|b|}$ når $n \geq N_1$. Dermed er $|b||a_n - a| < \frac{\epsilon}{2}$ når $n \geq N_1$. Dette gjelder opplagt også når $b = 0$ (med $N_1 = 1$).

Leddet $|a_n||b_n - b|$ er litt mer komplisert, siden $|a_n|$ ikke er en konstant. For at resonnementet ovenfor skal fungere må vi altså begrense $|a_n|$. Dette klarer vi ved Setning 6.2.11: siden $\{a_n\}$ er konvergent, finnes en $M \in \mathbb{Q}^+$ slik at $|a_n| \leq M$ for alle n. Siden $\lim_{n\to\infty} b_n = b$, finnes en $N_2 \in \mathbb{Z}^+$ slik at $|b_n - b| < \frac{\epsilon}{2M}$. Da vil altså $|a_n||b_n - b| < M \cdot \frac{\epsilon}{2M} = \frac{\epsilon}{2}$ for alle $n \geq N_2$.

Dermed velger vi $N = \max\{N_1, N_2\}$. Da vil $n \geq N$ medføre at både $|b||a_n - a| < \frac{\epsilon}{2}$ og $|a_n||b_n - b| < \frac{\epsilon}{2}$. Følgelig gir (98) at $|a_n b_n - ab| < \epsilon$. Vi har derfor vist (97) for $N = \max\{N_1, N_2\}$ og beviset er fullført.

(v) Siden $\frac{a_n}{b_n} = a_n \cdot \frac{1}{b_n}$, holder det ved (iv) å vise at $\lim\limits_{n\to\infty} \dfrac{1}{b_n} = \dfrac{1}{b}$, dersom $b \neq 0$.

Gitt $\epsilon \in \mathbb{Q}^+$. Vi vil finne en $N \in \mathbb{Z}^+$ slik at

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| < \epsilon \ \text{ når } \ n \geq N.$$

Vi omskriver

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \frac{|b - b_n|}{|b_n||b|} = \frac{1}{|b_n||b|} \cdot |b_n - b|.$$

Vi vet at vi kan gjøre $|b_n - b|$ så liten vi vil for stor nok n. Poenget må derfor være, som i beviset for (iii), å begrense $\frac{1}{|b_n||b|}$ oppad. Siden $|b|$ er en konstant, er problemet å begrense $\frac{1}{|b_n|}$ oppad. Dette er ekvivalent med å begrense $|b_n|$ nedad. Trekantulikheten gir oss at

$$|b| = |b - b_n + b_n| \leq |b - b_n| + |b_n|,$$

som gir $|b_n| \geq |b| - |b - b_n|$. Dersom for eksempel $|b - b_n| < \frac{|b|}{2}$, ser vi at $|b_n| > \frac{|b|}{2}$. Siden $\lim\limits_{n\to\infty} b_n = b$, vet vi at det finnes en $N_1 \in \mathbb{Z}^+$ slik at $|b_n - b| < \frac{|b|}{2}$ når $n \geq N_1$. Vi får altså at $|b_n| > \frac{|b|}{2}$ når $n \geq N_1$, og dermed at

$$\frac{1}{|b_n|} < \frac{2}{|b|}$$

når $n \geq N_1$. Det følger at

$$\begin{aligned} \left|\frac{1}{b_n} - \frac{1}{b}\right| &= \frac{1}{|b_n||b|} \cdot |b_n - b| = \frac{1}{|b_n|} \cdot \frac{1}{|b|} \cdot |b_n - b| \\ &< \frac{2}{|b|} \cdot \frac{1}{|b|} \cdot |b_n - b| = \frac{2}{b^2} \cdot |b_n - b| \end{aligned}$$

når $n \geq N_1$.

Siden $\lim\limits_{n\to\infty} b_n = b$, finnes en $N_2 \in \mathbb{Z}^+$ slik at $|b_n - b| < \frac{b^2\epsilon}{2}$ når $n \geq N_2$.

Dermed velger vi $N = \max\{N_1, N_2\}$. Da vil $n \geq N$ medføre at

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| < \frac{2}{b^2} \cdot |b_n - b| < \frac{2}{b^2} \cdot \frac{b^2\epsilon}{2} = \epsilon$$

og beviset er fullført. □

Beviset ovenfor er mye lengre enn bevis vi til vanlig finner i lærebøker og matematiske artikler, siden vi har vært opptatt av å vise hvordan vi resonnerer for å finne de N-ene som gir oss de riktige begrensningene. I tillegg har vi nevnt trekantulikheten hver gang vi har brukt den, men siden den er velkjent, vil man vanligvis ikke henvise til den, men ta det for gitt at leseren skjønner hva som foregår.

Det er instruktivt å se hvordan et slikt bevis til vanlig ville skrives:

Bevis for Setning 6.2.14 ("lærebokversjon")
(i) Gitt $\epsilon \in \mathbb{Q}^+$. Siden $\lim_{n\to\infty} a_n = a$, finnes en $N_1 \in \mathbb{Z}^+$ slik at $|a_n - a| < \frac{\epsilon}{2}$ når $n \geq N_1$. Likeledes, siden $\lim_{n\to\infty} b_n = b$, finnes en $N_2 \in \mathbb{Z}^+$ slik at $|b_n - b| < \frac{\epsilon}{2}$ når $n \geq N_2$. La $N = \max\{N_1, N_2\}$. Hvis $n \geq N$, er

$$|(a_n + b_n) - (a + b)| \leq |a_n - a| + |b_n - b| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

og beviset er fullført.

(ii) Dette beviset er nesten likt det forrige og overlates til leseren.

(iv) Gitt $\epsilon \in \mathbb{Q}^+$. Siden $\lim_{n\to\infty} a_n = a$, vil det finnes en $N_1 \in \mathbb{Z}^+$ slik at $|a_n - a| < \frac{\epsilon}{2|b|}$ når $n \geq N_1$, dersom $b \neq 0$. Dermed er $|b||a_n - a| < \frac{\epsilon}{2}$ når $n \geq N_1$, som også gjelder hvis $b = 0$.

Videre er $\{a_n\}$ begrenset ved Setning 6.2.11, og dermed finnes en $M \in \mathbb{Q}^+$ slik at $|a_n| \leq M$ for alle n. Siden $\lim_{n\to\infty} b_n = b$, finnes en $N_2 \in \mathbb{Z}^+$ slik at $|b_n - b| < \frac{\epsilon}{2M}$. Dermed er $|a_n||b_n - b| < M \cdot \frac{\epsilon}{2M} = \frac{\epsilon}{2}$ for alle $n \geq N_2$.

La $N = \max\{N_1, N_2\}$. Hvis $n \geq N$ vil

$$\begin{aligned} |a_n b_n - ab| &= |a_n b_n - a_n b + a_n b - ab| \leq |a_n b_n - a_n b| + |a_n b - ab| \\ &= |a_n||b_n - b| + |b||a_n - a| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \end{aligned}$$

og beviset er fullført.

(iii) Dette følger fra (iv) med $\{b_n\}$ lik den konstante følgen $\{c\}$.

(v) Det holder ved (iv) å vise at $\lim_{n\to\infty} \dfrac{1}{b_n} = \dfrac{1}{b}$, dersom $b \neq 0$.

Gitt $\epsilon \in \mathbb{Q}^+$. Siden $\lim_{n\to\infty} b_n = b$, finnes en $N_1 \in \mathbb{Z}^+$ slik at $|b_n - b| < \frac{|b|}{2}$ når $n \geq N_1$. Da $|b| = |b - b_n + b_n| \leq |b - b_n| + |b_n|$, vil $|b_n| \geq |b| - |b - b_n| > \frac{|b|}{2}$ når $n \geq N_1$. Likeledes finnes en $N_2 \in \mathbb{Z}^+$ slik at $|b_n - b| < \frac{b^2\epsilon}{2}$ når $n \geq N_2$.

La $N = \max\{N_1, N_2\}$. Hvis $n \geq N$, vil

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \frac{|b - b_n|}{|b_n||b|} < \frac{2}{|b|} \cdot \frac{1}{|b|} \cdot \frac{b^2\epsilon}{2} = \epsilon$$

og beviset er fullført. □

La oss se på hvordan vi kan anvende det siste resultatet sammen med grensene funnet i Hjelpesetning 6.2.7 for å finne grensen bevist i Eksempel 6.2.9 uten å bruke definisjonen av grense:

Eksempel 6.2.16. Finn $\lim_{n\to\infty} \dfrac{4n}{3n+2}$.

LØSNING. For å kunne bruke grensene i Hjelpesetning 6.2.7 omskriver vi:

$$\frac{4n}{3n+2} = \frac{4}{3+\frac{2}{n}}.$$

Ved Hjelpesetning 6.2.7 har vi $\lim_{n\to\infty} 2 = 2$ og $\lim_{n\to\infty} \frac{1}{n} = 0$. Dermed er $\lim_{n\to\infty} \frac{2}{n} = \lim_{n\to\infty} 2\cdot\frac{1}{n} = 2\cdot 0 = 0$ ved Setning 6.2.14(iv). Siden $\lim_{n\to\infty} 3 = 3$ ved Hjelpesetning 6.2.7, får vi $\lim_{n\to\infty}\left(3+\frac{2}{n}\right) = 3+0 = 3$ ved Setning 6.2.14(i). Videre, siden $\lim_{n\to\infty} 4 = 4$ ved Hjelpesetning 6.2.7, får vi

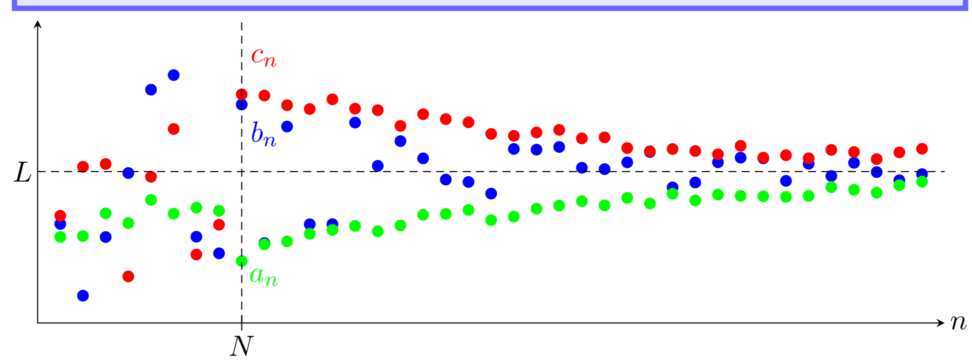
$$\lim_{n\to\infty}\frac{4n}{3n+2} = \lim_{n\to\infty}\frac{4}{3+\frac{2}{n}} = \frac{4}{3},$$

som stemmer med det vi fant ut i Eksempel 6.2.9. □

Vi tar også med følgende resultat, som vi kommer til å få bruk for senere.

Setning 6.2.17: Skvisesetningen for følger

Anta $\{a_n\}$, $\{b_n\}$ og $\{c_n\}$ er følger av rasjonale tall slik at $a_n \le b_n \le c_n$ for alle $n \ge N$, hvor $N \in \mathbb{Z}^+$. Hvis $\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = L$, da er $\lim_{n\to\infty} b_n = L$.



SKVISESETNING FOR FØLGER

BEVIS

Gitt $\epsilon \in \mathbb{Q}^+$. Vi vil vise at det finnes en $M \in \mathbb{Z}^+$ slik at

(99) $$|b_n - L| < \epsilon \text{ for alle } n \ge M.$$

Vi vet at $a_n \le b_n \le c_n$ for alle $n \ge N$, som medfører at

(100) $$a_n - L \le b_n - L \le c_n - L \text{ for alle } n \ge N.$$

Siden $\lim_{n\to\infty} a_n = L$, finnes en $N_1 \in \mathbb{Z}^+$ slik at $|a_n - L| < \epsilon$ for alle $n \ge N_1$. Spesielt er (jf. Merknad 6.2.4)

(101) $$-\epsilon < a_n - L \text{ for alle } n \ge N_1.$$

Siden $\lim_{n\to\infty} c_n = L$, finnes en $N_2 \in \mathbb{Z}^+$ slik at $|c_n - L| < \epsilon$ for alle $n \geq N_2$. Spesielt er (jf. Merknad 6.2.4)

$$c_n - L < \epsilon \quad \text{for alle} \quad n \geq N_2. \tag{102}$$

La $M = \max\{N, N_1, N_2\}$. Hvis $n \geq M$, gjelder alle (100)-(102), slik at

$$-\epsilon < a_n - L \leq b_n - L \leq c_n - L < \epsilon.$$

Spesielt er $-\epsilon < b_n - L < \epsilon$, som viser (99) (husk Merknad 6.2.4). □

EKSEMPEL 6.2.18. Vis at $\lim_{n\to\infty} \frac{c}{2^n} = 0$ for enhver $c \in \mathbb{Q}$.

LØSNING. Siden $4n < 2^n$ for alle heltall $n \geq 5$ ved Eksempel 5.2.7, er

$$0 < \frac{4}{2^n} < \frac{1}{n} \quad \text{for alle} \quad n \geq 5$$

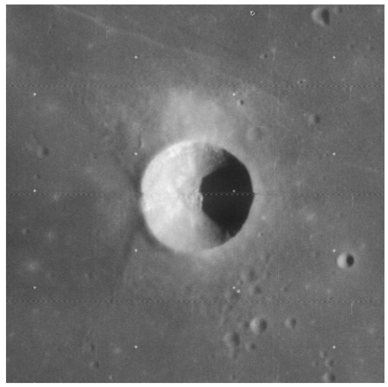
ved Setning 3.6.13(v). Siden $\lim_{n\to\infty} \frac{1}{n} = 0$ ved Hjelpesetning 6.2.7(a) og $\lim_{n\to\infty} 0 = 0$ ved Hjelpesetning 6.2.7(b), er $\lim_{n\to\infty} \frac{4}{2^n} = 0$ ved *skvisesetningen* (6.2.17). Ved Setning 6.2.14(iii) må vi derfor ha $\lim_{n\to\infty} \frac{c}{2^n} = \lim_{n\to\infty} \left(\frac{c}{4} \cdot \frac{4}{2^n}\right) = 0$ for alle $c \in \mathbb{Q}$. □

6.3. Cauchy-følger i $\mathbb{Q}$

Det er nå på tide å gå tilbake til det vi snakket om i begynnelsen av forrige seksjon, nemlig hvordan man kan bruke følger til å poengtere hva som mangler med de rasjonale tallene.

AUGUSTIN–LOUIS CAUCHY (KILDE: WIKIPEDIA, PUBLIC DOMAIN)

Til det trenger vi å studere en bestemt type følger som har den egenskapen at avstanden mellom leddene blir så liten vi vil når vi kommer langt nok ut i følgen. De har fått navnet sitt etter den franske matematikeren, fysikeren og ingeniøren (og baronen) Augustin–Louis Cauchy (1789–1857), en av grunnleggerne av den matematiske analysen. (Cauchy har forresten også fått oppkalt et krater på månen etter seg.)

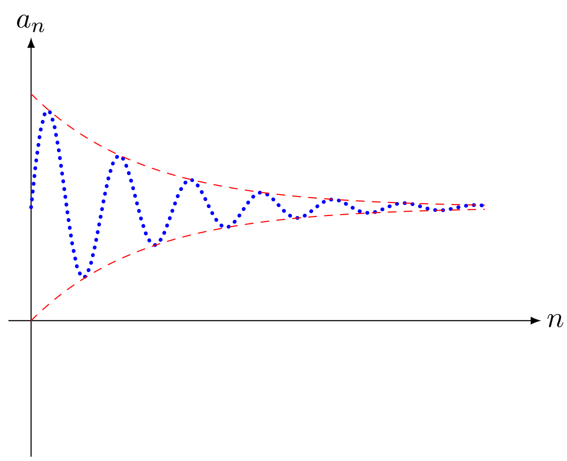


Cauchy-krateret på månen (Kilde: Wikipedia, public domain)

Definisjon 6.3.1

En følge $\{a_n\}$ av rasjonale tall kalles en *Cauchy-følge* dersom det for ethvert rasjonalt tall $\epsilon > 0$ finnes et tall $N \in \mathbb{Z}^+$ slik at $|a_n - a_m| < \epsilon$ for alle $m, n \geq N$.

Her har vi altså ikke noen grense involvert: det vi vet er at avstanden mellom leddene blir vilkårlig liten for stor nok n, som vist i følgende figur:



En Cauchy-følge

Merknad 6.3.2. Som før, jf. Merknad 6.2.8, vil N i siste definisjon kunne avhenge av ϵ.

Intuitivt skulle man kanskje tro at alle slike følger konvergerer, men dette skjer ikke i $\mathbb{Q}$. Som vi vil se senere, skjer det imidlertid i $\mathbb{R}$.

Det er imidlertid sant at alle konvergente følger er Cauchy-følger:

Setning 6.3.3

Enhver konvergent følge av rasjonale tall er en Cauchy-følge.

Bevis

Anta at $\{a_n\}$ konvergerer mot a. Gitt $\epsilon > 0$, vil vi finne en $N \in \mathbb{Z}^+$ slik at $|a_n - a_m| < \epsilon$ for alle $m, n \geq N$. Siden $\lim_{n\to\infty} a_n = a$ finnes en $N \in \mathbb{Z}^+$ slik at $|a_n - a| < \frac{\epsilon}{2}$ når $n \geq N$. Hvis $m, n \geq N$, har vi ved trekantulikheten

$$|a_n - a_m| = |(a_n - a) + (a - a_m)| \leq |a_n - a| + |a - a_m| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

og beviset er fullført. $\square$

Det motsatte gjelder ikke i $\mathbb{Q}$:

Setning 6.3.4

Det finnes Cauchy-følger av rasjonale tall som ikke konvergerer mot noe rasjonalt tall.

Bevis

Vi tar utgangspunkt i samme teknikk brukt i beviset for at $\mathbb{Q}$ ikke oppfyller kompletthetsprinsippet (Setning 6.1.9). Tilordningen av tallet q til tallet p i det beviset definerer en følge av rasjonale tall $\{p_n\}$ definert rekursivt ved

$$p_{n+1} = p_n - \frac{p_n^2 - 2}{p_n + 2} = 2 - \frac{2}{p_n + 2} \tag{103}$$

og ved å velge $p_1 \in \mathbb{Q}^+$ helt vilkårlig.

Det overlates til Oppgave 6.11 å vise at $\{p_n\}$ er veldefinert (det vi si, $p_n \neq -2$ for alle $n \in \mathbb{Z}^+$) og er en Cauchy-følge. Hvis $\{p_n\}$ har en grense $L \in \mathbb{Q}$, da vil

$$L = \lim_{n\to\infty} p_n = \lim_{n\to\infty} p_{n+1} = L,$$

siden leddene i følgen $\{p_{n+1}\}$ er de samme som leddene i $\{p_n\}$ forskjøvet med 1. Tar vi grenser på begge sider av (103) og bruker regnereglene i Setning 6.2.14, finner vi

$$L = 2 - \frac{2}{L+2},$$

og løser vi for L, finner vi

$$\begin{aligned} L(L+2) &= 2(L+2) - 2 \\ L^2 + 2L &= 2L + 2 \\ L^2 &= 2, \end{aligned}$$

men dette er umulig ved Setning 1.3.5. □

Dette siste eksemplet viser igjen problemet at de rasjonale tallene har "hull" og hvorfor vi trenger utvidelsen av $\mathbb{Q}$ til de reelle tallene $\mathbb{R}$: følgen $\{p_n\}$ er en Cauchy-følge, slik at leddene nærmer seg hverandre mer og mer, men følgen konvergerer ikke mot et tall i $\mathbb{Q}$, for dette tallet L måtte ha oppfylt $L^2 = 2$. Når vi først vet at $\mathbb{R}$ eksisterer, så vil $\{p_n\}$ konvergere mot tallet L som oppfyller $L^2 = 2$, nemlig $\sqrt{2}$. (Strengt tatt kunne også $L = -\sqrt{2}$ være en mulig grense, men siden alle leddene er positive, vil vi lett klare å eliminere den muligheten.) Vi skal vise dette i Eksemplene 7.5.6 og 7.6.2.

Vi skal i neste kapittel se på hvordan man kan "tette igjen hullene i $\mathbb{Q}$" og lage mengden $\mathbb{R}$ av de reelle tallene ved nettopp å bruke Cauchy-følger av rasjonale tall.

MERKNAD 6.3.5. Det finnes mange andre rekursivt definerte rasjonale følger som gir Cauchy-følger som har den egenskapen at *hvis* de konvergerer mot en grense L, da må L oppfylle $L^2 = 2$, for eksempel $p_{n+1} = \frac{p_n}{2} + \frac{1}{p_n}$. Denne følgen er et spesialtilfelle av en generell metode for å beregne tilnærmede verdier på kvadratrøtter kjent som *Herons metode*, som vi beskriver i neste eksempel.

EKSEMPEL 6.3.6 (Herons metode). Vi beskriver en metode for å tilnærme verdier på kvadratrøtter av tall fra verket *Metrika* fra år 60 skrevet av den gresk-egyptiske matematikeren og ingeniøren Heron av Alexandria (ca. 10–ca. 70).

Anta at vi vil finne (en tilnærmet verdi for) kvadratroten av et positivt heltall K, det vil si vi ønsker å finne (en tilnærmet verdi for) x slik at $x^2 = K$. Vi starter da med et rektangel med sider av lengde $l_1 = K$ og bredde $b_1 = 1$. Idéen er nå å trinnvis omdanne dette rektangelet til et kvadrat med samme areal. I hvert trinn lager vi et nytt rektangel som har lengde lik gjennomsnittet av sidelengdene i det forrige rektangelet, og bredde bestemt slik at arealet fortsatt er K.

Det betyr at rektangel nummer to i rekken har lengde

$$l_2 = \frac{l_1 + b_1}{2} = \frac{K+1}{2}$$

og bredde

$$b_2 = \frac{K}{l_2} = \frac{2K}{K+1}.$$

Hvis vi har konstruert n rektangler med sider av lengde $l_1, \ldots, l_n$ og bredde $b_1, \ldots, b_n$, så vil neste rektangel ha side av lengde

$$l_{n+1} = \frac{l_n + b_n}{2}$$

og bredde

$$b_{n+1} = \frac{K}{l_{n+1}}.$$

Siden rektanglene blir mer og mer lik et kvadrat av areal K for hvert trinn, vil man forvente at l_n og b_n nærmer seg mer og mer lengden L av siden til et kvadrat av areal K, det vil si et tall L slik at $L^2 = K$ (som er det vi i dag kaller kvadratroten av K).

Vi kan sette inn at $b_n = \frac{K}{l_n}$ i uttrykket for l_{n+1} ovenfor og oppnå følgen $\{l_n\}$ av rasjonale tall definitivt rekursivt ved

$$l_1 = K \text{ og } l_{n+1} = \frac{1}{2}\left(l_n + \frac{K}{l_n}\right). \tag{104}$$

Ved å argumentere som i beviset for Setning 6.3.4 kan man vise at *dersom* følgen konvergerer mot en grense L, da må $L^2 = K$ (overlates til Oppgave 6.12).

Oppgaver

Oppgaver til §6.1

Oppgave 6.1. La $S = \{q \in \mathbb{Q} \mid q < \frac{3}{2}\}$. Finn $\sup S$, dersom den eksisterer.

Oppgave 6.2. La S være en begrenset delmengde av en ordnet mengde. La M være en øvre skranke for S og m være en nedre skranke for S. Vis at $m \leq M$.

Oppgave 6.3. La X være en ordnet mengde som oppfyller kompletthets-prinsippet og $\emptyset \neq A \subset B \subset X$.

Vis at hvis B er oppad begrenset (henholdsvis nedad begrenset), da er $\sup A \leq \sup B$ (hhv. $\inf A \geq \inf B$).

Oppgave 6.4. Verifisér alle overgangene som bringer (86) til (87) i Eksempel 6.1.7 ved å henvise til egenskapene i Definisjon 3.6.12 og Setning 3.6.13.

Oppgave 6.5. Lag egne eksempler som illustrerer begrepene *nedad begrenset*, *nedre skranke* og *største nedre skranke* à la Eksemplene 6.1.2, 6.1.4 og 6.1.7, samt eksemplet i beviset for Setning 6.1.9, og som viser direkte at $\mathbb{Q}$ ei heller har største-nedre-skranke-egenskapen.

Oppgaver til §6.2–6.3

Oppgave 6.6. Bevis Setning 6.2.14(ii) ved å bruke Merknad 3.6.17.

Oppgave 6.7. Gi et (litt) annerledes bevis for Setning 6.2.14(iv) ved å bruke at

$$a_n b_n - ab = (a_n - a)(b_n - b) + a(b_n - b) + b(a_n - a).$$

Oppgave 6.8. Bevis Setning 6.2.6. (Hint: Anta at både L og L' oppfyller kriteriet for å være grense til følgen $\{a_n\}$. Dersom $L \neq L'$, la $\epsilon = \frac{|L-L'|}{2}$ og utled en motsigelse ved hjelp av trekantulikheten og ved å skrive $|L - L'| = |(L - a_n) - (L' - a_n)|$.)

Oppgave 6.9. Vis at $\lim_{n\to\infty} \dfrac{1}{n^2} = 0$ ved å

(a) bruke Definisjon 6.2.3;
(b) bruke Hjelpesetning 6.2.7(a) og regnereglene i Setning 6.2.14.

Oppgave 6.10. Vis at $\lim_{n\to\infty} \dfrac{2n^2}{n^2 - 1} = 2$ ved å

(a) bruke Definisjon 6.2.3;
(b) bruke Hjelpesetning 6.2.7(a) og regnereglene i Setning 6.2.14.

Oppgave 6.11. Betrakt følgen $\{p_n\}$ definert som i beviset for Setning 6.3.4, for et vilkårlig $p_1 \in \mathbb{Q}^+$.

(a) Vis at $p_n \in \mathbb{Q}^+$ for alle $n \in \mathbb{Z}^+$. (Spesielt viser dette at $\{p_n\}$ er veldefinert, siden $p_n \neq -2$.)
(b) Vis at $|p_{n+1} - p_n| \leq \frac{1}{2}|p_n - p_{n-1}|$ ved å bruke den rekursive formelen.
(c) Konkludér at $|p_{n+1} - p_n| \leq \frac{1}{2^{n-1}}|p_2 - p_1|$ for alle $n \in \mathbb{Z}^+$.
(d) Bruk resultatet i Oppgave 5.10 til å utlede at for alle $m \geq n \geq 2$ gjelder
$$|p_m - p_n| \leq \frac{1}{2^{n-2}}|p_2 - p_1|.$$
(e) Konkludér at $\{p_n\}$ er en Cauchy-følge.

Oppgave 6.12. Vis at følgen i Merknad 6.3.5 er en Cauchy-følge som ikke har noen grense i $\mathbb{Q}$.

Vis mer generelt at følgen i (104) generert av Herons metode er en Cauchy-følge, og at grensen L oppfyller $L^2 = K$, hvis følgen konvergerer.

Oppgave 6.13. Vis at en Cauchy-følge av rasjonale tall er begrenset.

Oppgave 6.14. Vis at dersom $\{p_n\}$ og $\{q_n\}$ er Cauchy-følger av rasjonale tall, da er også følgene $\{p_n + q_n\}$ og $\{p_n q_n\}$ Cauchy-følger. (Hint til andre del: skriv

$$|p_n q_n - p_m q_m| = |p_n q_n - p_n q_m + p_n q_m - p_m q_m| \leq |p_n||q_n - q_m| + |q_m||p_n - p_m|,$$

og bruk resultatet fra Oppgave 6.13.)

KAPITTEL 7

De reelle tallene

Som nevnt i innledningen til forrige kapittel innså allerede grekerne at det fantes tall som ikke var rasjonale. De fortsatte imidlertid med å skille strengt mellom det de kalte “tall”, som er det vi i dag kaller rasjonale tall, og “størrelser”, som er det vi i dag kaller irrasjonale tall. De siste ble bare godtatt som svar på geometriske problemer, for eksempel ville grekerne ikke starte med et linjestykke som hadde lengde $\sqrt{2}$. Arvtagerne etter grekerne ble de arabiske matematikerne, som smeltet sammen begrepene størrelser og tall og introduserte notasjonen for tall slik vi kjenner dem idag.

Den egyptiske matematikeren Abū Kāmil (ca. 850–ca. 930) regnes for å være den første som systematisk brukte og aksepterte irrasjonale tall som løsninger på ligninger og koeffisienter i ligninger. Den arabiske matematikeren Al-Uqlidisi regnes for å være den første som introduserte desimaltall; det skjedde i boken *Kitab al-Fusul fi al-Hisab al-Hindi* rundt 952. Verkene om desimaltall ble aldri oversatt til europeiske språk, og den flamske matematikeren, fysikeren og ingeniøren Simon Stevin (1548–1620) gjenoppdaget og introduserte desimaltall i Europa i et arbeid fra 1585.

Bernard Bolzano (Kilde: Wikipedia, public domain)

I løpet av 1700- og 1800-tallet utviklet matematikken seg raskt, spesielt studiet av funksjoner mellom mengder av reelle tall, og deres egenskaper som kontinuitet, deriverbarhet og integrerbarhet (det vi i dag kaller *kalkulus*[1]). Det vokste frem et stort behov for å rydde opp i uklarheter. En av de

[1]Fra latin *calculus*, som betyr “liten sten”, siden små stener brukes til å telle.

første som forsøkte å konstruere de reelle tallene fra de rasjonale var den bohemske matematikeren, logikeren, filosofen, teologen og katolske presten Bernard Bolzano (1781–1848) i første halvdel av 1800–tallet. Hans forsøk gikk ut på å definere reelle tall ved hjelp av følger av rasjonale tall, men han manglet de korrekte detaljene og publiserte ikke sine arbeider. Blant annet ga han et feil bevis på at Cauchy-følger konvergerer. En pionér i arbeidet med å fikse det teoretiske grunnlaget for kalkulus var den franske matematikeren og fysikeren Augustin-Louis Cauchy (1789–1857), som også behandlet reelle tall som grenser av følger av rasjonale tall, men dette var mer en observasjon enn en konstruksjon. Dessuten tok Cauchy for gitt at Cauchy-følger konvergerer.

Idéen av å bruke (Cauchy)-følger for å "tette igjen hullene" i de rasjonale tallene viste seg imidlertid å være god. En rigorøs konstruksjon av reelle tall på denne måten ble utviklet av av den franske matematikeren Charles Méray (1835–1911) i 1869 og den tyske matematikeren Georg Cantor (1845–1918) i 1871 uavhengig av hverandre, og de siste brikkene falt på plass gjennom et arbeid av Cantors kollega Eduard Heine (1821–1881) i 1872. Denne konstruksjonen skal vi se på i første seksjon i dette kapitlet. Deretter skal vi utlede alle kjente egenskaper til reelle tall ut i fra denne konstruksjonen i §7.2–7.3, i tillegg til egenskaper til følger av reelle tall (§7.4–7.6).

Vi må også nevne at den tyske matematikeren Karl Weierstrass (1815–1897) skal ha presentert en alternativ konstruksjon av de reelle tallene i sine forelesninger i Berlin i 1863. Han foretok en grundig analyse av *desimaltall* og viste at disse hadde de egenskapene man ønsket seg. Som vi skal se i §7.8 er desimaltall ikke noe annet enn følger av rasjonale tall, slik at denne tilnærmingen egentlig ikke er så forskjellig fra den til Méray–Cantor–Heine.

En tredje, alternativ konstruksjon av de reelle tallene ble publisert i 1872 av den tyske matematikeren Richard Dedekind (1831–1916). Vi vil gi et raskt overblikk over denne konstruksjonen i et appendiks (7.11) til kapitlet.

Richard Dedekind (Kilde: Wikipedia, public domain)

Vi vil dessuten se, i §7.9, at det finnes forskjellige typer uendeligheter i matematikk. De rasjonale tallene og de irrasjonale tallene er uendelig mange, men det finnes mange flere irrasjonale tall enn rasjonale tall. Forskjellen er at de rasjonale tallene kan, kanskje overraskende, "ordnes i en følge", mens de irrasjonale kan det ikke. Vi sier at de rasjonale tallene er *tellbare*, mens de irrasjonale er *ikke-tellbare*. I §7.10 vil vi kort studere såkalte *algebraiske tall*: dette er reelle tall som er røtter til polynomer med rasjonale koeffisienter, slik som $\sqrt{2}$, som er rot i $x^2 - 2$. Det var lenge et åpent spørsmål om det fantes reelle tall som ikke er algebraiske, såkalt *transcendentale*, før man på 1800-tallet slo fast at e of π er transcendentale, og ikke nok med det: de transcendentale tallene er ikke-tellbare, mens de algebraiske er tellbare.

7.1. Méray–Cantor–Heine-konstruksjonen av $\mathbb{R}$ fra $\mathbb{Q}$

Vi starter med et par definisjoner. Det norske navnet på ett av begrepene er unektelig morsomt. (For definisjonen av en *kropp*, se Definisjon 3.6.1.)

Definisjon 7.1.1: Underkropp og ordnet underkropp

Kroppen k sies å være en *underkropp*[a] av kroppen K dersom $k \subset K$ og kroppsoperasjonene på k er de samme som på K (mer formelt, dersom $+_k$ og $\cdot_k$ er operasjonene i k og $+_K$ og $\cdot_K$ er operasjonene i K, da er $a +_k b = a +_K b$ og $a \cdot_k b = a \cdot_K b$ for alle $a, b \in k$).

Dersom k og K er ordnede kropper, sier vi i tillegg at k er en *ordnet underkropp*[b] av kroppen K dersom ordensrelasjonen på k er den samme som på K (mer formelt, dersom $<_k$ er ordensrelasjonen på k og $<_K$ er ordensrelasjonen på K, da gjelder for alle $a, b \in k$ at $a <_k b$ hvis og bare hvis $a <_K b$).

[a] *Subfield* på engelsk.

[b] *Ordered subfield* på engelsk.

Målet i denne seksjonen er å bevise følgende:

Teorem 7.1.2: Eksistens av de reelle tall

Det finnes en ordnet kropp $\mathbb{R}$ som oppfyller kompletthetsprinsippet (ekvivalent: minste-øvre-skranke-egenskapen) og som inneholder $\mathbb{Q}$ som en ordnet underkropp.

Kroppen $\mathbb{R}$ kalles *kroppen av de reelle tallene*.

Idéen bak beviset er følgende: siden $\mathbb{Q}$ mangler grenser til Cauchy-følger, legger vi disse "grensene" til ved å representere dem ved Cauchy-følgene selv. I denne fremstillingen blir et rasjonalt tall $q \in \mathbb{Q}$ representert ved den *konstante* Cauchy-følgen $C_q = \{q, q, q, \ldots\}$, mens for eksempel det reelle tallet $\sqrt{2}$ blir representert ved Cauchy-følgen $\{p_n\}$ i beviset for Setning 6.3.4. Vi må imidlertid ta i betraktning at forskjellige Cauchy-følger kan gi samme grense; for eksempel gir forskjellige startpunkter p_1 i den sistnevnte følgen forskjellige Cauchy-følger som fremdeles representerer $\sqrt{2}$, og det samme gjelder følgene $\{p_n\}$ i Merknad 6.3.5. For å unngå denne tvetydigheten innfører vi en ekvivalensrelasjon hvor vi gjør alle Cauchy-følger ekvivalente hvis de har ledd som kommer vilkårlig nær hverandre langt nok ut:

Definisjon 7.1.3: Ekvivalente Cauchy-følger

To Cauchy-følger $\{p_n\}$ og $\{q_n\}$ i $\mathbb{Q}$ sies å være *ekvivalente* hvis $\lim_{n\to\infty} |p_n - q_n| = 0$.

Vi skriver $\{p_n\} \sim \{q_n\}$ når $\{p_n\}$ og $\{q_n\}$ er ekvivalente.

Vi overlater til Oppgave 7.1 å vise at $\sim$ virkelig definerer en ekvivalensrelasjon på mengden av alle Cauchy-følger i $\mathbb{Q}$.

Vi betegner ekvivalensklassen til Cauchy-følgen $\{p_n\}$ med $[p_n]$ og definerer mengden $\mathbb{R}$ til å være *mengden av alle ekvivalensklasser av Cauchy-følger i* $\mathbb{Q}$, det vil si, med notasjonen i Definisjon 2.4.11, har vi

$$\mathbb{R} = \{\text{Cauchy-følger i } \mathbb{Q}\}/\sim .$$

Vi merker at ekvivalensklassen $[C_q]$ til en *konstant* Cauchy-følge

$$C_q = \{q, q, q, q, q, \ldots\}$$

inneholder kun én konstant følge, nemlig følgen selv. Dette skyldes at det er klart fra definisjonen at $C_p \sim C_q$ hvis og bare hvis $p = q$. Med andre ord (jf. Observasjon 2.4.12) er $[C_p] = [C_q]$ hvis og bare hvis $p = q$. Spesielt betyr dette at funksjonen

$$\begin{aligned} j : \mathbb{Q} &\longrightarrow \mathbb{R} \\ q &\mapsto [C_q] \end{aligned} \tag{105}$$

er injektiv. Derfor kan vi identifisere $\mathbb{Q}$ med verdimengden til denne funksjonen, som er en delmengde av $\mathbb{R}$. Gjennom denne identifiseringen sier vi at $\mathbb{Q}$ er en delmengde av $\mathbb{R}$.

Nå har vi altså definert en ny mengde $\mathbb{R}$ som inneholder $\mathbb{Q}$. For å fullføre beviset for Teorem 7.1.2 må vi utvide kroppsoperasjonene $+$ og $\cdot$ og ordensrelasjonen $<$ fra $\mathbb{Q}$ til $\mathbb{R}$ og vise at disse gjør $\mathbb{R}$ til en ordnet kropp som oppfyller kompletthetsprinsippet.

Operasjonene addisjon og multiplikasjon på $\mathbb{R}$. Hvis $\{p_n\}$ og $\{q_n\}$ er Cauchy-følger av rasjonale tall, da er også følgene $\{p_n + q_n\}$ og $\{p_n q_n\}$ Cauchy-følger ved Oppgave 6.14. Dette betyr at for $[p_n], [q_n] \in \mathbb{R}$, er også $[p_n + q_n] \in \mathbb{R}$ og $[p_n] \cdot [q_n] = [p_n q_n] \in \mathbb{R}$. Vi kan derfor definere to operasjoner $+$ og $\cdot$ på $\mathbb{R}$ som er kompatible med de tilsvarende operasjonene på $\mathbb{Q}$:

$$[p_n] + [q_n] = [p_n + q_n] \ \text{ og } \ [p_n] \cdot [q_n] = [p_n \cdot q_n].$$

Vi må vise at disse operasjonene er veldefinert, det vil si at de er uavhengig av valg av representatene for ekvivalensklassene til Cauchy-følgene. Dette er tatt hånd om i følgende hjelpesetning:

Hjelpesetning 7.1.4

Hvis $\{p'_n\} \sim \{p_n\}$ og $\{q'_n\} \sim \{q_n\}$, da er $\{p'_n + q'_n\} \sim \{p_n + q_n\}$ og $\{p'_n \cdot q'_n\} \sim \{p_n \cdot q_n\}$.

Bevis

Per definisjon betyr $\{p'_n\} \sim \{p_n\}$ og $\{q'_n\} \sim \{q_n\}$ at $\lim_{n\to\infty} |p'_n - p_n| = 0$ og $\lim_{n\to\infty} |q'_n - q_n| = 0$. Ved trekantulikheten (Setning 3.6.16(ii)) har vi

$$0 \leq |(p'_n + q'_n) - (p_n + q_n)| \leq |p'_n - p_n| + |q'_n - q_n|$$

og dermed $\lim_{n\to\infty} |(p'_n + q'_n) - (p_n + q_n)| = 0$ fra Skvisesetningen (6.2.17). Det følger at $\{p'_n + q'_n\} \sim \{p_n + q_n\}$, som ønsket.

For å vise den resterende delen, skriver vi (ved Setning 3.6.16 igjen)

$$\begin{aligned} |p'_n q'_n - p_n q_n| &= |p'_n q'_n - p_n q'_n + p_n q'_n - p_n q_n| \\ &= |q'_n(p'_n - p_n) + p_n(q'_n - q_n)| \\ &\leq |q'_n||p'_n - p_n| + |p_n||q'_n - q_n|. \end{aligned}$$

Siden enhver Cauchy-følge av rasjonale tall er begrenset ved Oppgave 6.13, finnes $M \in \mathbb{Q}$ og $M' \in \mathbb{Q}$ slik at $|p_n| \leq M$ og $|p'_n| \leq M'$ for alle n. Da har vi

$$0 \leq |p'_n q'_n - p_n q_n| \leq M'|p'_n - p_n| + M|q'_n - q_n|.$$

Siden $\lim_{n\to\infty} \left(M'|p'_n - p_n| + M|q'_n - q_n|\right) = 0$ ved Setning 6.2.14, vil også $\lim_{n\to\infty} |p'_n q'_n - p_n q_n| = 0$ ved Skvisesetningen (6.2.17). Det følger dermed at $\{p'_n \cdot q'_n\} \sim \{p_n \cdot q_n\}$, som ønsket. $\square$

Hjelpesetning 7.1.5

Mengden $\mathbb{R}$ med operasjonene $+$ og $\cdot$ er en kropp med

- nøytralt element $0_\mathbb{R} = [C_0]$, ekvivalensklassen til den konstante Cauchy-følgen $C_0 = \{0, 0, 0, 0, \ldots\}$, og
- multiplikativt identiteselement $1_\mathbb{R} = [C_1]$, ekvivalensklassen til den konstante Cauchy-følgen $C_1 = \{1, 1, 1, 1, \ldots\}$.

Kroppen $\mathbb{R}$ inneholder $\mathbb{Q}$ som underkropp.

Bevis

Det overlates til Oppgave 7.3 å sjekke at operasjonene oppfyller kroppsaksiomene. Det additive inverselementet til et element $[p_n]$ er $[-p_n]$ og det multiplikative inverselementet til $[p_n] \neq [C_0]$ er $[1/p_n]$. Her må vi imidlertid være forsiktige på grunn av muligheten for at $p_n = 0$ for noen n. Det sentrale poenget er at siden $[p_n] \neq [C_0]$, så kan vi alltid finne en representant p_n for ekvivalensklassen slik at $p_n \neq 0$ for alle n, ved Oppgave 7.2(b).

Det er opplagt ut i fra definisjonen av operasjonene og identifiseringen av $\mathbb{Q}$ som delmengden av de konstante Cauchy-følgene i $\mathbb{R}$ at $\mathbb{Q}$ er en underkropp av $\mathbb{R}$. □

Ordensrelasjonen på $\mathbb{R}$. Vi vil nå definere en ordensrelasjon på $\mathbb{R}$ som utvider den kjente ordensrelasjonen på $\mathbb{Q}$. Vi starter med å definere, for to vilkårlige Cauchy-følger av rasjonale tall:

$$\begin{aligned} \{p_n\} < \{q_n\} \iff & \text{det finnes en } N \in \mathbb{Z}^+ \text{ og en } r \in \mathbb{Q}^+ \\ & \text{slik at } p_n + r < q_n \text{ for alle } n \geq N. \end{aligned} \tag{106}$$

For å bruke dette til å definere en ordensrelasjon på $\mathbb{R}$, må vi vise at siste egenskap er uavhengig av representantene for ekvivalensklassene:

Hjelpesetning 7.1.6

Anta at $\{p_n\} < \{q_n\}$. Hvis $\{p'_n\} \sim \{p_n\}$ og $\{q'_n\} \sim \{q_n\}$, da er $\{p'_n\} < \{q'_n\}$.

Bevis (lang, resonnerende versjon)

La oss skrive ned alle antagelsene vi har å rutte med.

Siden vi antar at $\{p_n\} < \{q_n\}$, vet vi at det finnes en $r \in \mathbb{Q}^+$ slik at

$$q_n > p_n + r \quad \text{for alle store nok } n. \tag{107}$$

Siden $\{p'_n\} \sim \{p_n\}$ og $\{q'_n\} \sim \{q_n\}$ kan vi gjøre differansene $|p_n - p'_n|$ og $|q_n - q'_n|$ så små vi vil for store nok n. Med andre ord, uansett gitt $\epsilon_1 > 0$ og $\epsilon_2 > 0$, har vi at

$$-\epsilon_1 < p_n - p'_n < \epsilon_1 \quad \text{for alle store nok } n \tag{108}$$

og

$$-\epsilon_2 < q_n - q'_n < \epsilon_2 \quad \text{for alle store nok } n \tag{109}$$

(husk Merknad 6.2.4). Ut fra dette ønsker vi å vise at $\{p'_n\} < \{q'_n\}$, som betyr at vi vil finne en $r' \in \mathbb{Q}^+$ slik at

$$q'_n > p'_n + r' \quad \text{for alle store nok } n. \tag{110}$$

For å utlede (110) ut i fra (107)-(109), skriver vi, for store nok n:

$$\begin{aligned} q'_n & \overset{(109)}{>} q_n - \epsilon_2 \overset{(107)}{>} (p_n + r) - \epsilon_2 \\ & \overset{(108)}{>} (p'_n - \epsilon_1) + r - \epsilon_2 = p'_n + (r - \epsilon_1 - \epsilon_2). \end{aligned}$$

Vi ser at vi oppnår (110) dersom $r - \epsilon_1 - \epsilon_2 > 0$, ved å la $r' = r - \epsilon_1 - \epsilon_2 > 0$. For eksempel fungerer dette for $\epsilon_1 = \epsilon_2 = \frac{r}{3}$. □

Vi skriver beviset ned i litt mer polert form, for å presentere det på den måten som er vanlig i matematiske tekster:

Bevis ("Lærebokversjon")

Siden vi antar at $\{p_n\} < \{q_n\}$, finnes det per definisjon en $N_1 \in \mathbb{Z}^+$ og en $r \in \mathbb{Q}^+$ slik at

$$q_n > p_n + r \quad \text{for alle} \quad n \geq N_1. \tag{111}$$

Siden $\{p'_n\} \sim \{p_n\}$ og $\{q'_n\} \sim \{q_n\}$ finnes $N_2 \in \mathbb{Z}^+$ og $N_3 \in \mathbb{Z}^+$ slik at

$$-\frac{r}{3} < p_n - p'_n < \frac{r}{3} \quad \text{for alle} \quad n \geq N_2 \tag{112}$$

og

$$-\frac{r}{3} < q_n - q'_n < \frac{r}{3} \quad \text{for alle} \quad n \geq N_3. \tag{113}$$

For $n \geq N = \max\{N_1, N_2, N_3\}$ har vi dermed ved (111)-(113) at

$$q'_n > q_n - \frac{r}{3} > (p_n + r) - \frac{r}{3} = p_n + \frac{2r}{3} > \left(p'_n - \frac{r}{3}\right) + \frac{2r}{3} = p'_n + \frac{r}{3},$$

som medfører $\{q'_n\} > \{p'_n\}$ per definisjon. □

Ved hjelpesetningen kan vi derfor definere, for alle $[p_n], [q_n] \in \mathbb{R}$:

$$[p_n] < [q_n] \iff \{p_n\} < \{q_n\} \text{ for hvilke som helst representanter } \{p_n\} \text{ og } \{q_n\}.$$

Hjelpesetning 7.1.7

Relasjonen $<$ er en ordensrelasjon på $\mathbb{R}$ som oppfyller

- $[p_n] < [q_n] \Longrightarrow [p_n] + [r_n] < [q_n] + [r_n]$,
- $[p_n], [q_n] > 0_{\mathbb{R}} \Longrightarrow [p_n] \cdot [q_n] > 0_{\mathbb{R}}$,

for alle $[p_n], [q_n], [r_n] \in \mathbb{R}$.

Bevis

Overlates til Oppgave 7.5. □

Dermed har vi gjort $\mathbb{R}$ til en ordnet kropp (jf. Definisjon 3.6.12) som inneholder $\mathbb{Q}$ som ordnet underkropp, siden det er klart at den nylig definerte ordensrelasjonen reduseres til den kjente på elementer i $\mathbb{Q}$ (som er representert ved konstante Cauchy-følger).

$\mathbb{R}$ **oppfyller kompletthetprinsippet.** Vi vil til slutt vise at $\mathbb{R}$ oppfyller komplettshetprinsippet, eller ekvivalent, minste-øvre-skranke egenskapen. Dette vil fullføre beviset for Teorem 7.1.2. Dette er definitivt den tyngste delen av konstruksjonen og kan derfor hoppes over ved første lesing.

Anta derfor at E er en ikke-tom oppad begrenset delmengde av $\mathbb{R}$. Vi ønsker å vise at E har en minste øvre skranke.

La $[u_n] \in \mathbb{R}$ være en øvre skranke. Siden $\{u_n\}$ er en Cauchy-følge, er den begrenset (ved Oppgave 6.13). Det finnes derfor en $U \in \mathbb{Q}$ slik at

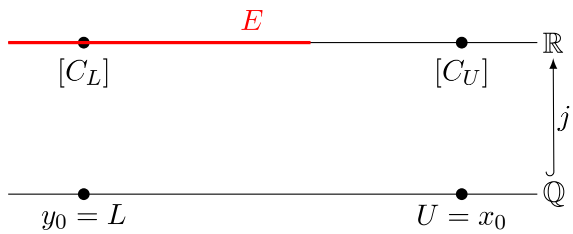
$$u_n \leq U - 1 \quad \text{for alle} \quad n.$$

Dermed er $[C_U] = [\{U, U, U, U, \ldots\}] > [u_n]$ i $\mathbb{R}$ og $[C_U]$ er også en øvre skranke for E. Poenget her er å bytte ut den opprinnelige øvre skranken med en øvre skranke som er representert ved en *konstant* følge.

Siden $E \neq \emptyset$, finnes ihvertfall ett element $[p_n] \in E$. Siden $\{p_n\}$ er begrenset (igjen siden det er en Cauchy-følge), finnes en $L \in \mathbb{Q}$ slik at

$$p_n \geq L + 1 \quad \text{for alle} \;\; n.$$

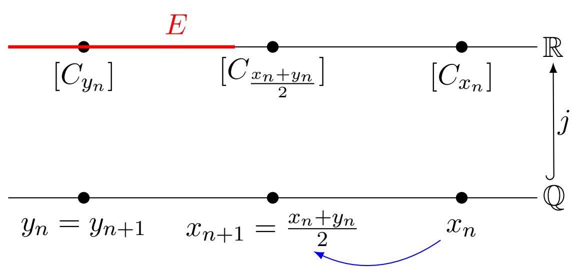
Dermed har vi $[C_L] = [\{L, L, L, L, \ldots\}] < [p_n]$ og $[C_L]$ er *ikke* en øvre skranke for E. Merk også at $L < U$ per konstruksjon. Vi setter $x_0 = U$ og $y_0 = L$. Følgende figur oppsummerer situasjonen, hvor avbildningen j er inklusjonen av $\mathbb{Q}$ inn i $\mathbb{R}$ gitt i (105):



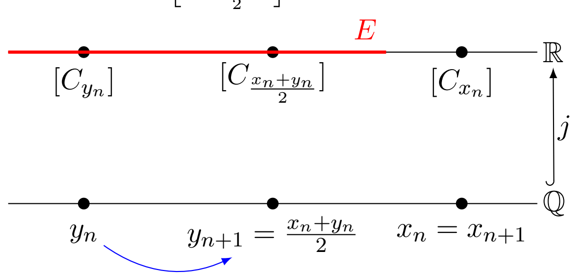
Vi definerer nå to følger $\{x_n\}$ og $\{y_n\}$ i $\mathbb{Q}$ rekursivt: Når vi har definert $x_0, \ldots, x_n$ og $y_0, \ldots y_n$, fortsetter vi på følgende måte:

(114)
$$\begin{cases} \text{Hvis } \left[C_{\frac{x_n+y_n}{2}}\right] \text{ er en øvre skranke for } E, \text{ la } x_{n+1} = \frac{x_n+y_n}{2} \text{ og } y_{n+1} = y_n. \\ \text{Hvis ikke, la } y_{n+1} = \frac{x_n+y_n}{2} \text{ og } x_{n+1} = x_n. \end{cases}$$

Følgende figurer oppsummerer situasjonen. I det øverste tilfellet, hvor $\left[C_{\frac{x_n+y_n}{2}}\right]$ er en øvre skranke for E, har vi:



I det nederste tilfellet, hvor $\left[C_{\frac{x_n+y_n}{2}}\right]$ ikke er en øvre skranke for E, har vi:



I hvert trinn i konstruksjonen av denne følgen vil altså én av x_n og y_n forbli den samme, mens den andre vil ta verdien til midtpunktet mellom de to. Måten vi velger hvilken av de to som tar verdien av midtpunktet gjør at begge følgene vil nærme seg mer og mer "det høyre endepunktet" på mengden E i figurene ovenfor. Det er nettopp dette som gjør at $\sup E$ ekisterer. I resten av beviset vil vi vise at

- $\{x_n\}$ og $\{y_n\}$ er ekvivalente Cauchy-følger. (Hjelpesetning 7.1.8); dette betyr at $[x_n] = [y_n]$ i $\mathbb{R}$.
- $\sup E = [x_n] = [y_n]$.

Vi starter med å merke at siden som sagt én av x_n og y_n forblir den samme i hvert trinn, mens den andre vil ta verdien til midtpunktet mellom de to, har vi $|x_{n+1} - y_{n+1}| = \frac{1}{2}|x_n - y_n|$. Dermed har vi

$$\begin{aligned} |x_n - y_n| &= \frac{1}{2}|x_{n-1} - y_{n-1}| = \frac{1}{2}\left(\frac{1}{2}|x_{n-2} - y_{n-2}|\right) = \cdots \\ &= \left(\frac{1}{2}\right)^n |x_0 - y_0| = \frac{|U-L|}{2^n}. \end{aligned}$$

Siden $\lim_{n\to\infty} \frac{|U-L|}{2^n} = 0$ ved Eksempel 6.2.18, har vi at

$$\lim_{n\to\infty} |x_n - y_n| = 0. \tag{115}$$

Siden $x_0 > y_0$, er det dessuten lett å se at

$$y_n \leq y_m \leq x_m \leq x_n \quad \text{for alle} \quad m \geq n \geq 0. \tag{116}$$

Hjelpesetning 7.1.8

Følgene $\{x_n\}$ og $\{y_n\}$ er ekvivalente Cauchy-følger.

BEVIS

At følgene er ekvivalente, hvis de er Cauchy, følger fra (115).

For å vise at følgene er Cauchy, la $m \geq n$. Ved trekantulikheten er

$$\begin{aligned} |x_m - x_n| &= |x_m - y_n + y_n - x_n| \leq |x_m - y_n| + |y_n - x_n| \\ &\overset{(116)}{=} (x_m - y_n) + (x_n - y_n) \overset{(116)}{\leq} (x_n - y_n) + (x_n - y_n) \\ &= 2(x_n - y_n). \end{aligned}$$

La $\epsilon > 0$. Ved (115) finnes en $N \in \mathbb{Z}^+$ slik at $x_n - y_n = |x_n - y_n| < \frac{\epsilon}{2}$ for $n \geq N$. Da er $|x_m - x_n| < \epsilon$ for $m \geq n \geq N$. Dette viser at $\{x_n\}$ en Cauchy-følge.

Beviset for at $\{y_n\}$ er en Cauchy-følge er helt tilsvarende. □

Hjelpesetningen viser at $[x_n] = [y_n] \in \mathbb{R}$.

Vi vil vise at $[x_n]$ er den minste øvre skranken til E.

For å gjøre det trenger vi følgende resultat (ta en titt på figurene på forrige side for å innse at resultatet er forventet):

Hjelpesetning 7.1.9

For enhver $k \geq 0$, er $[C_{x_k}] = [\{x_k, x_k, x_k, x_k, \ldots\}]$ en øvre skranke for E, mens $[C_{y_k}] = [\{y_k, y_k, y_k, y_k, \ldots\}]$ er det ikke.

Bevis

Vi beviser dette ved induksjon på k.

Ved konstruksjon er $[C_{x_0}] = [C_U]$ en øvre skranke for E.

Anta nå at $[C_{x_k}]$ er en øvre skranke for E, for $k \geq 0$. Hvis $x_{k+1} = x_k$, da er $[C_{x_{k+1}}] = [C_{x_k}]$ og dermed en øvre skranke for E. Hvis $x_{k+1} < x_k$, da gir den rekursive definisjonen av følgen (114) at

$$x_{k+1} = \frac{x_k + y_k}{2} \quad \text{og} \quad \left[C_{\frac{x_k+y_k}{2}}\right] \text{ er en øvre skranke for } E,$$

slik at $[C_{x_{k+1}}] = \left[C_{\frac{x_k+y_k}{2}}\right]$ er en øvre skranke for E.

Beviset hva angår $[C_{y_k}]$ er lignende (jf. Oppgave 7.6). □

Vi vil nå bevise at $[x_n]$ er en øvre skranke for E.

Vi argumenterer ved selvmotsigelse: Hvis $[x_n]$ ikke er en øvre skranke for E, vil det eksistere en $[q_n] \in E$ slik at $[q_n] > [x_n]$. Dette betyr at det finnes en $N_1 \in \mathbb{Z}^+$ og en $r \in \mathbb{Q}^+$ slik at

$$q_n > x_n + r \quad \text{for alle} \quad n \geq N_1. \tag{117}$$

Siden $\{x_n\}$ er Cauchy, finnes en $N_2 \in \mathbb{Z}^+$ slik at $|x_m - x_n| < \frac{1}{2}r$ for alle $n, m \geq N_2$; spesielt er

$$x_m - x_n < \frac{1}{2}r \quad \text{for alle} \quad n, m \geq N_2. \tag{118}$$

Dermed har vi, for $n, m \geq N = \max\{N_1, N_2\}$:

$$q_n \overset{(117)}{>} x_n + r \overset{(118)}{>} \left(x_m - \frac{1}{2}r\right) + r = x_m + \frac{1}{2}r.$$

Spesielt får vi, ved å sette $m = N$:

$$q_n > x_N + \frac{1}{2}r \quad \text{for alle} \quad n \geq N$$

og det følger at $[q_n] > [C_{x_N}] = [\{x_N, x_N, x_N, x_N, \ldots\}]$, som motsier Hjelpesetning 7.1.9, som sier at at $[C_{x_N}]$ er en øvre skranke for E.

Til slutt vil vi bevise at $[x_n]$ er en *minste* øvre skranke for E.

Igjen argumenterer vi ved selvmotsigelse: Hvis $[x_n]$ ikke er en minste øvre skranke for E, eksisterer en øvre skranke $[q_n]$ for E slik at $[q_n] < [x_n] = [y_n]$. Dette betyr at det finnes en $N_1 \in \mathbb{Z}^+$ og en $r \in \mathbb{Q}^+$ slik at

$$y_n > q_n + r \quad \text{for alle} \quad n \geq N_1. \tag{119}$$

Siden $\{y_n\}$ er Cauchy, finnes en $N_2 \in \mathbb{Z}^+$ slik at $|y_n - y_m| < \frac{1}{2}r$ for alle $n, m \geq N_2$; spesielt er

$$y_n - y_m < \frac{1}{2}r \quad \text{for alle} \quad n, m \geq N_2. \tag{120}$$

Dermed har vi, for $n, m \geq N = \max\{N_1, N_2\}$,

$$q_n \overset{(119)}{<} y_n - r \overset{(120)}{<} \left(y_m + \frac{1}{2}r\right) - r = y_m - \frac{1}{2}r.$$

Spesielt får vi, ved å sette $m = N$:

$$q_n < y_N - \frac{1}{2}r \text{ for alle } n \geq N,$$

og det følger at $[q_n] < [C_{y_N}] = [\{y_N, y_N, y_N, y_N, \ldots\}]$, som motsier Hjelpesetning 7.1.9, som sier at $[C_{y_N}]$ ikke er en øvre skranke for E.

Vi har derfor fullført beviset for at $[x_n]$ er en minste øvre skranke for E, og dermed vist at $\mathbb{R}$ oppfyller kompletthetsprinsippet. Beviset for Teorem 7.1.2 er dermed fullført.

MERKNAD 7.1.10. Siden Teorem 7.1.2 viser eksistensen av én kropp $\mathbb{R}$ som oppfyller de ønskede egenskapene, er det naturlig å spørre seg om det finnes flere slike kropper. Vi nevnte for eksempel i introduksjonen til kapitlet at både Weierstrass og Dedekind hadde alternative konstruksjoner for de reelle tallene. I konstruksjonen til Méray–Cantor–Heine som vi har sett på representeres de reelle tallene med ekvivalensklasser av Cauchy-følger, i Weierstrass' konstruksjon representeres de ved desimaltall (se §7.8), mens i Dedekinds konstruksjon representeres de av delmengder av rasjonale tall (se 7.11). Betyr dette at det finnes "flere typer" reelle tall?

Svaret er heldigvis at selv om "objektene" som representerer tallene er forskjellige i de forskjellige konstruksjonene, er de "like" på den måten at det finnes bijeksjoner mellom alle de resulterende mengdene slik at kroppsoperasjonene og ordensrelasjonen er bevart. Vi sier at de ordnede kroppene er *isomorfe*. La oss for ordens skyld skrive ned definisjonene:

To kropper K_1 og K_2, med kroppsoperasjoner $+_1, \cdot_1$ og $+_2, \cdot_2$, henholdsvis, sies å være isomorfe[2] *dersom det finnes en bijeksjon $\varphi : K_1 \to K_2$ slik at $\varphi(x +_1 y) = \varphi(x) +_2 \varphi(y)$ og $\varphi(x \cdot_1 y) = \varphi(x) \cdot_2 \varphi(y)$ for alle $x, y \in K_1$.*

Avbildningen φ (betegnet her med den greske bokstaven "fi") kalles i dette tilfellet en *(kropps)isomorfi*[3]. Dette betyr i praksis at vi kan jobbe med de to kroppene som om de var like, siden objektene er i én-til-én korrespondanse på en slik måte at kroppsoperasjonene forblir de samme. Eneste forskjellen, for å si det slik, er at objektene har forskjellige "navn".

Hvis kroppene K_1 og K_2 er i tillegg ordnede, med ordensrelasjoner $<_1$ og $<_2$, henholdsvis, sier vi at de er isomorfe som ordnede kropper *dersom i tillegg $x <_1 y \Leftrightarrow \varphi(x) <_2 \varphi(y)$ for alle $x, y \in K_1$.*

Et viktig resultat, som vi ikke skal bevise her, sier at

[2] *Isomorphic* på engelsk.

[3] *Isomorphism (of fields)* på engelsk.

To ordnede kropper med minste-øvre-skranke egenskapen er alltid isomorfe (som ordnede kropper).

Vi kan derfor trygt jobbe med $\mathbb{R}$, samt egenskapene vi utleder i de neste seksjonene, som om $\mathbb{R}$ er entydig. Vi sier at $\mathbb{R}$ er den eneste ordnede kroppen med minste-øvre-skranke egenskapen *opp til isomorfi*[4].

7.2. Noen egenskaper til de reelle tallene

Vi har allerede sett at mengden $\mathbb{R}$ av reelle tall oppfyller egenskapene:

- $\mathbb{R}$ er en ordnet kropp (det vil si $\mathbb{R}$ oppfyller aksiomene i Definisjonene 3.6.1, 3.6.9 og 3.6.12);
- $\mathbb{R}$ inneholder $\mathbb{Q}$ som ordnet underkropp (det vil si $\mathbb{Q}$ er en delmengde av $\mathbb{R}$ og operasjonene $+$ og $\cdot$ og ordensrelasjonen $<$ på $\mathbb{R}$ restrisert til $\mathbb{Q}$ er de samme som de vi har definert tidligere på $\mathbb{Q}$);
- $\mathbb{R}$ oppfyller kompletthetsprinsippet (Definisjon 6.1.8).

Det følger at $\mathbb{R}$ også oppfyller egenskapene i Setningene 3.6.3 og 3.6.13.

Husk at vi sier (jf. Definisjon 3.6.12) at x er positiv (henholdsvis negativ) hvis $x > 0$ (hhv. $x < 0$). Vi betegner mengden av *positive reelle tall* med $\mathbb{R}^+$ og mengden av *negative reelle tall* med $\mathbb{R}^-$, slik at

$$\mathbb{R} = \mathbb{R}^+ \cup \{0\} \cup \mathbb{R}^-.$$

Det er også vanlig å bruke notasjonene

$$\mathbb{R}^{\geq 0} = \mathbb{R}^+ \cup \{0\}$$

for mengden av de ikke-negative reelle tallene og

$$\mathbb{R}^{\leq 0} = \mathbb{R}^- \cup \{0\}$$

for mengden av de ikke-positive reelle tallene.

Vi kaller elementene i $\mathbb{R} \setminus \mathbb{Q}$ for *irrasjonale tall*.

Vi skal videre se på noen flere viktige egenskaper som $\mathbb{R}$ oppfyller.

Arkimedes' prinsipp og tetthet.

Setning 7.2.1: Arkimedes' prinsipp

Hvis $x, y \in \mathbb{R}^+$, da finnes en $n \in \mathbb{Z}^+$ slik at $nx > y$.

Bevis

Tanken er å undersøke alle tall nx for $n \in \mathbb{Z}^+$ og vise at vi før eller siden oppnår et tall større enn y.

La da

$$S = \{nx \mid n \in \mathbb{Z}^+\}.$$

Nå argumenterer vi ved selvmotsigelse: Anta at det vi skal vise ikke holder, det vil si at det *ikke* finnes noen $n \in \mathbb{Z}^+$ slik at $nx > y$. Da

[4] *Up to isomorphism* på engelsk.

ville y vært en øvre skranke for S. Da har S en minste øvre skranke, si $s = \sup S$. Siden $x > 0$, er $s - x < s$, slik at $s - x$ *ikke* kan være en øvre skranke for S (siden s er den *minste* slike). Altså vil det finnes et tall $mx \in S$, der $m \in \mathbb{Z}^+$, slik at $mx > s - x$. Men da er $(m+1)x > s$, som er umulig, siden $(m+1)x \in S$ og s er en øvre skranke for S. □

Setning 7.2.2: Tetthet

Ethvert åpent intervall i $\mathbb{R}$ inneholder både rasjonale og irrasjonale tall.

Bevis

Det holder selvsagt å vise setningen for åpne *endelige* intervaller. Disse har formen $(a,b) = \{x \in \mathbb{R} \mid a < x < b\}$ for reelle tall $a < b$, jf. §2.1.

La oss først vise at (a,b) inneholder et rasjonalt tall. Idéen er enkel: vi finner først ved Arkimedes' prinsipp en $n \in \mathbb{Z}^+$ slik at $n(b-a) > 1$, det vil si $b - a > \frac{1}{n}$. Så starter vi i origo og går mot intervallet (a,b) med skritt av lengde $\frac{1}{n}$: siden skrittlengden er mindre enn $b - a$, som er lengden til intervallet, vil vi før eller siden havne i intervallet, i et punkt på formen $\frac{m}{n}$, som er rasjonalt. Detaljene overlates til Oppgave 7.9.

Å finne et irrasjonalt tall i (a,b) overlates til Oppgave 7.10. □

Merknad 7.2.3. Setningen gir spesielt at det mellom to forskjellige reelle tall, spesielt to irrasjonale, alltid finnes et rasjonalt tall, faktisk uendelig mange rasjonale tall! Dette virker kanskje litt overraskende, siden vi er vant til at det finnes mange flere irrasjonale tall enn rasjonale. Og det er helt korrekt at det finnes "mange flere" irrasjonale tall enn rasjonale: det finnes riktignok uendelig mange av begge, men vi vil se i §7.9 at det finnes forskjellige typer uendeligheter som "måler" antallet elementer i en mengde, og de irrasjonale tallene er virkelig en mye større mengde enn de rasjonale tallene.

Avstandsbegrep (metrikk) på $\mathbb{R}$. Vi har allerede nevnt at vi definerer absoluttverdien til et reellt tall x som

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

Det er opplagt at absoluttverdien oppfyller

$$\begin{cases} |x| > 0 & \text{hvis } x \neq 0, \\ |0| = 0, \end{cases} \tag{121}$$

og vi har sett i Setning 3.6.16 at den oppfyller

$$|xy| = |x||y|, \tag{122}$$

$$|x+y| \leq |x| + |y| \quad \text{(Trekantulikheten)}. \tag{123}$$

Vi kan betrakte $|x|$ som avstanden mellom x og tallet 0. Mer generelt kan vi definere *avstanden* mellom $x, y \in \mathbb{R}$ som det reelle tallet $|x-y|$. Avstanden oppfyller følgende egenskaper, for alle $x, y, z \in \mathbb{R}$:

$$\begin{cases} |x-y| > 0 & \text{hvis } x \neq y, \\ |x-y| = 0, & \text{hvis } x = y \end{cases} \qquad \text{(Positivitet)} \tag{124}$$

$$|x-y| = |y-x| \qquad \text{(Symmetri)} \tag{125}$$

$$|x-z| \leq |x-y| + |y-z| \qquad \text{(Trekantulikhet)} \tag{126}$$

Her er egenskap (124) en konsekvens av (121), egenskap (126) en konsekvens av *trekantulikheten* (123) med x og y erstattet med $x-y$ og $y-z$, henholdsvis, mens egenskap (125) er opplagt. Spesielt har vi en fin geometrisk tolkning av trekantulikheten (126), som forklarer navnet: For $x, y, z \in \mathbb{R}$ sier (126) at "avstanden mellom to punkter er mindre eller lik summen av avstanden mellom de to punktene og et vilkårlig tredje punkt".

Vi har altså en "avstandsfunksjon"

$$\begin{array}{ccc} \mathbb{R} \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ (x, y) & \mapsto & |x-y|, \end{array}$$

som oppfyller egenskapene (124)-(126). Siden $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$ har vi en tilsvarende funksjon definert på $\mathbb{Q}$ og $\mathbb{Z}$. Å ha en slik avstandsfunksjon er veldig viktig, fordi den gjør oss i stand til å definere konvergens av følger, som vi allerede har sett i $\mathbb{Q}$, og som vi skal se i $\mathbb{R}$ i §7.4. Det gjør oss også i stand til å definere begrepene *kontinuitet* og *grenser*, som vi skal se i Kapittel 9. En mengde med et slikt avstandsbegrep har derfor fått et eget navn:

Definisjon 7.2.4: Metrikk og metrisk rom

La X være en mengde. En *metrikk*[a] på X er en funksjon $d : X \times X \to \mathbb{R}$ som oppfyller følgende aksiomer for alle $x, y, z \in X$:

(M1) $\begin{cases} d(x,y) > 0 & \text{hvis } x \neq y \\ d(x,y) = 0, & \text{hvis } x = y \end{cases}$ (Positivitet)

(M2) $d(x,y) = d(y,x)$ (Symmetri)

(M3) $d(x,z) \leq d(x,y) + d(y,z)$ (Trekantulikhet)

En mengde X sammen med en metrikk d på X kalles et *metrisk rom*[b].

[a] *Metric* på engelsk.

[b] *Metric space* på engelsk.

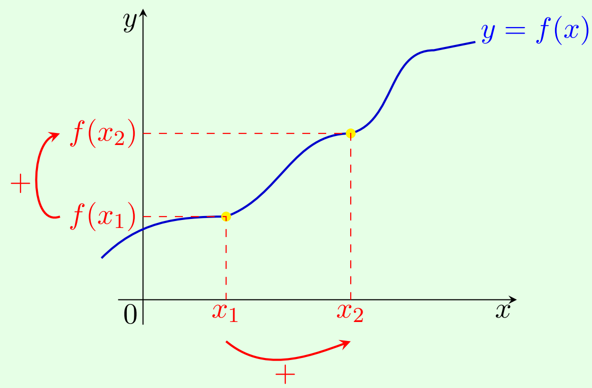
Betraktningene ovenfor viser at både $\mathbb{R}$, $\mathbb{Q}$ og $\mathbb{Z}$ er metriske rom (med det definerte avstandsbegrepet). Vi vil se i neste kapittel at også mengden av de komplekse tallene $\mathbb{C}$ er et metrisk rom, og vi vil se ytterligere eksempler i Kapittel 10.

Voksende og avtagende funksjoner. Vi avslutter seksjonen med en definisjon som skal være kjent fra skolen og som vi vil få bruk for ved flere anledninger. Siden vi har en ordensrelasjon $<$ på $\mathbb{R}$, er vi interessert i å sile ut funksjoner som enten bevarer eller snur ulikheter:

Definisjon 7.2.5: Voksende funksjon

La $f : D(f) \to \mathbb{R}$, med $D(f) \subset \mathbb{R}$, være en funksjon.

Vi sier at f er *voksende*[a] (henholdsvis *strengt voksende*[b]) dersom vi for alle $x_1 < x_2$ i $D(f)$ har at $f(x_1) \leq f(x_2)$ (hhv. $f(x_1) < f(x_2)$).



Strengt voksende funksjon

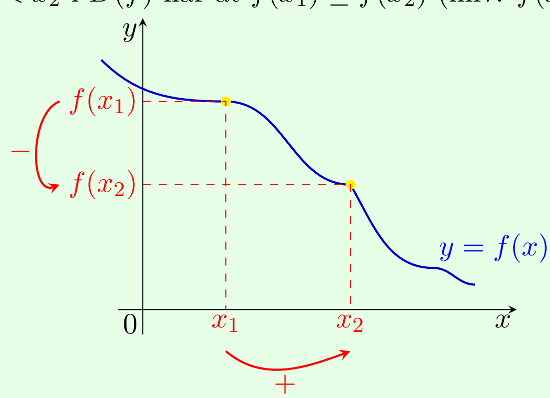
[a] *Increasing* på engelsk.

[b] *Strictly increasing* på engelsk.

Definisjon 7.2.6: Avtagende funksjon

La $f : D(f) \to \mathbb{R}$, med $D(f) \subset \mathbb{R}$, være en funksjon.

Vi sier at f er *avtagende*[a] (henholdsvis *strengt avtagende*[b]) dersom vi for alle $x_1 < x_2$ i $D(f)$ har at $f(x_1) \geq f(x_2)$ (hhv. $f(x_1) > f(x_2)$).



Strengt avtagende funksjon

[a] *Decreasing* på engelsk.

[b] *Strictly decreasing* på engelsk.

Definisjon 7.2.7: Monoton funksjon

La $f : D(f) \to \mathbb{R}$, med $D(f) \subset \mathbb{R}$, være en funksjon.

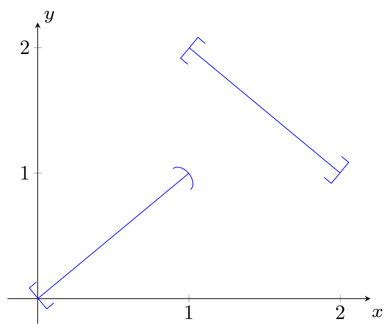
Vi sier at f er *monoton*[a] (henholdsvis *strengt monoton*[b]) dersom f er voksende eller avtagende (hhv. strengt voksende eller strengt avtagende).

[a] *Monotone* på engelsk.
[b] *Strictly monotone* på engelsk.

MERKNAD 7.2.8. En strengt monoton funksjon er automatisk injektiv (Definisjon 2.3.1). Det motsatte er ikke alltid sant, se for eksempel på funksjonen $f : [0, 2] \to \mathbb{R}$ gitt ved

$$f(x) = \begin{cases} x, & x \in [0, 1) \\ 3 - x, & x \in [1, 2], \end{cases}$$

som er injektiv, men ikke monoton.



GRAFEN TIL f I MERKNAD 7.2.8

Vi vil imidlertid se senere, i Setning 9.6.13, at det motsatte er sant for en *kontinuerlig* funksjon der $D(f)$ er et intervall.

7.3. Røtter og potenser av reelle tall

Eksistens av n-te røtter. Vi viser eksistensen av n-te røtter, noe vi er vant til er sant fra skolen.

Setning 7.3.1: Eksistens av røtter

Til enhver $x \in \mathbb{R}^{\geq 0}$ og enhver $n \in \mathbb{Z}^{+}$, finnes nøyaktig én $y \in \mathbb{R}^{\geq 0}$ slik at $y^n = x$.

Definisjon 7.3.2: n-te rot

Tallet y i forrige setning kalles *n-te roten av x* og betegnes med $\sqrt[n]{x}$ (forenklet til $\sqrt{x}$ når $n = 2$) eller $x^{\frac{1}{n}}$.

Bevis for Setning 7.3.1
La $n \in \mathbb{Z}^+$ og $x \in \mathbb{R}^{\geq 0}$ være gitt. Hvis $x = 0$, er det ikke noe å vise, siden det er klart fra egenskapen til kropper i Setning 3.6.3(x) at $y = 0$ er eneste tallet som oppfyller $0^n = 0$. Anta derfor at $x > 0$.

Entydigheten til n-te roten vises ved selvmotsigelse: hvis $y_1^n = y_2^n = x$ for to forskjellige $y_1, y_2 \in \mathbb{R}^{\geq 0}$, kan vi ved symmetri anta $y_1 < y_2$, og da er $y_1^n < y_2^n$ ved Oppgave 5.9, en motsigelse.

Det som gjenstår er å vise *eksistensen* av y. Idéen er temmelig enkel og intuitiv. Vi ser på mengden av alle tall t som oppfyller $t^n < x$, og viser at den minste øvre skranken y til denne mengden oppfyller $y^n = x$.

La derfor
$$S = \{t \in \mathbb{R}^+ \mid t^n < x\}.$$

Påstand 7.3.2.1. Mengden S er ikke-tom og oppad begrenset.

Bevis for påstand. La $u = \frac{x}{1+x}$. Da er $0 < u < 1$ og dermed er $u^n \leq u < x$ ved Oppgave 5.9. Derfor er $u \in S$, slik at $S \neq \emptyset$.

Hvis $s > 1 + x$, da er $s^n \geq s > x$ ved Oppgave 5.9, slik at $s \notin S$. Dermed er $1 + x$ er øvre skranke for S, slik at S er oppad begrenset. □

Siden $\mathbb{R}$ oppfyller kompletthetsprinsippet, har S en minste øvre skranke $y = \sup S > 0$, som vi vil vise oppfyller det ønskede kriteriet $y^n = x$.

For å vise denne likheten, vil vi vise at begge ulikhetene $y^n < x$ og $y^n > x$ vil føre til en selvmotsigelse. Ved trikotonimegenskapen til ordensrelasjonen (Definisjon 3.6.9) vil $y^n = x$ da være eneste mulighet.

Vi vil få bruk for ulikheten
$$a^n - b^n < (a - b)na^{n-1}, \quad \text{for } 0 < b < a, \tag{127}$$
som utledes i Oppgave 7.12(b).

Anta at $y^n < x$. Da er $\frac{x-y^n}{n(y+1)^{n-1}} > 0$. Sett
$$h = \min\left\{1, \frac{x - y^n}{n(y+1)^{n-1}}\right\},$$
hvor uttrykket min betyr "det minste av tallene". Merk at
$$0 < h \leq 1 \tag{128}$$
og $h \leq \frac{x-y^n}{n(y+1)^{n-1}}$, det vil si
$$hn(y + 1)^{n-1} \leq x - y^n. \tag{129}$$
Ved å bruke ulikheten (127) med $a = y + h$ og $b = y$, får vi:
$$(y + h)^n - y^n \overset{(127)}{<} hn(y + h)^{n-1} \overset{(128)}{\leq} hn(y + 1)^{n-1} \overset{(129)}{\leq} x - y^n.$$
Altså er $(y + h)^n < x$ og dermed $y + h \in S$. Men siden $y + h > y$, motsier dette at y er en øvre skranke for S.

Anta dernest at $y^n > x$. Sett

$$k = \frac{y^n - x}{ny^{n-1}} = \frac{y}{n} - \frac{x}{ny^{n-1}}. \tag{130}$$

Da er $0 < k < y$. Vi vil nå vise at $y - k$ er en øvre skranke for S, som motsier at y er en minste øvre skranke, siden $y - k < y$. Hvis $t \geq y - k$, får vi ved å bruke ulikheten (127) med $a = y$ og $b = y - k$:

$$y^n - t^n \leq y^n - (y-k)^n \overset{(127)}{<} kny^{n-1} \overset{(130)}{=} y^n - x.$$

Dette viser at $t^n > x$ og dermed at $t \notin S$. Det følger at $y - k$ er en øvre skranke for S, og beviset er fullført. $\square$

Før vi går videre, viser vi følgende velkjente resultat, som vi kommer til å få bruk for senere.

Hjelpesetning 7.3.3: Røtter bevarer orden

Dersom $0 < x_1 < x_2$, da er $\sqrt[n]{x_1} < \sqrt[n]{x_2}$.

Bevis

Vi argumenterer ved motsigelse. La $y_1 = \sqrt[n]{x_1}$ og $y_2 = \sqrt[n]{x_2}$. Anta at $y_1 \geq y_2$. Da vil $x_1 = y_1^n \geq y_2^n = x_2$ ved Oppgave 5.9, en motsigelse. $\square$

Merknad 7.3.4. Setning 7.3.1 kan også utvides:

- Dersom n er oddetall, kan man vise at *til enhver* $x \in \mathbb{R}$ *og enhver* $n \in \mathbb{Z}^+$, *finnes nøyaktig én* $y \in \mathbb{R}$ *slik at* $y^n = x$. For å gjøre det, bruker man at $y^n = x$ er ekvivalent med $(-y)^n = -x = |x|$, hvis $x < 0$. Altså er $y < 0$, hvis $x < 0$ og fremdeles entydig. Vi betegner y fremdeles med $\sqrt[n]{x}$ og merker at $\sqrt[n]{-x} = -\sqrt[n]{x}$.
- Dersom n er partall, kan man vise at *til enhver* $x \in \mathbb{R}^+$ *og enhver* $n \in \mathbb{Z}^+$, *finnes nøyaktig to* $y \in \mathbb{R}$ *slik at* $y^n = x$, nemlig $y = \pm \sqrt[n]{x}$.

Vi overlater detaljene til Oppgave 7.14.

Merknad 7.3.5. Anta n er et partall. Setning 7.3.1 sier at funksjonen f definert ved $f(x) = x^n$ har verdimengde $[0, \infty)$ og er injektiv (jf. Definisjon 2.3.1). Altså kan vi skrive

$$\begin{aligned} f : [0, \infty) &\to [0, \infty) \\ x &\to x^n, \end{aligned}$$

som er en bijektiv funksjon (jf. Definisjon 2.3.4). Den inverse funksjonen er, igjen ved Setning 7.3.1:

$$\begin{aligned} f^{-1} : [0, \infty) &\to [0, \infty) \\ x &\to \sqrt[n]{x}. \end{aligned}$$

Funksjonen f er strengt voksende ved Oppgave 5.9, mens f^{-1} er strengt voksende ved Hjelpesetning 7.3.3.

Anta n er et oddetall. På tilsvarende måte sier Setning 7.3.1 og Merknad 7.3.4 at funksjonen

$$\begin{aligned} g : \mathbb{R} &\to \mathbb{R} \\ x &\to x^n \end{aligned}$$

er bijektiv, med invers funksjon

$$\begin{aligned} g^{-1} : \mathbb{R} &\to \mathbb{R} \\ x &\to \sqrt[n]{x}. \end{aligned}$$

Begge funksjonene g og g^{-1} er strengt voksende, igjen ved Oppgave 5.9 og Hjelpesetning 7.3.3.

Følgesetning 7.3.6

Hvis $a, b \in \mathbb{R}^+$ og $n \in \mathbb{Z}^+$, da er $(ab)^{\frac{1}{n}} = a^{\frac{1}{n}} b^{\frac{1}{n}}$.

Bevis

La $A = a^{\frac{1}{n}}$ og $B = b^{\frac{1}{n}}$. Siden multiplikasjon er kommutativ (aksiom (M2) i Definisjon 3.6.1), har vi

$$ab = A^n B^n = (AB)^n.$$

Entydighetsegenskapen i Setning 7.3.1 viser derfor at

$$(ab)^{\frac{1}{n}} = AB = a^{\frac{1}{n}} b^{\frac{1}{n}},$$

som er det vi skulle vise. □

Rasjonale eksponenter. Setning 7.3.1 og dens korollar kan brukes til å definere potenser av positive reelle tall med *rasjonale* eksponenter:

Definisjon 7.3.7

For $m \in \mathbb{Z}$, $n \in \mathbb{Z}^+$ og $x \in \mathbb{R}^+$ definerer vi

$$x^{\frac{m}{n}} = (x^m)^{\frac{1}{n}}.$$

For å vise at dette gir mening, må vi vise at vi får samme resultat hvis vi bytter ut m og n med $m' \in \mathbb{Z}$, $n' \in \mathbb{Z}^+$ slik at $\frac{m}{n} = \frac{m'}{n'}$, det vil si at $(x^m)^{\frac{1}{n}} = \left(x^{m'}\right)^{\frac{1}{n'}}$. Dette overlates til Oppgave 7.15(a).

Reelle eksponenter. Vi kan definere x^r for alle $x \in \mathbb{R}^+$ og $r \in \mathbb{R}$ ved å gå frem på følgende måte. Vi betrakter mengden

$$\{x^t \mid t \in \mathbb{Q}, t \leq r\} \subset \mathbb{R}.$$

Man kan vise at $\sup\{x^t \mid t \in \mathbb{Q}, t \leq r\} = x^r$ når $r \in \mathbb{Q}$ (jf. Oppgave 7.15(b)). Dermed gir følgende definisjon mening:

Definisjon 7.3.8

For $r \in \mathbb{R}$ og $x \in \mathbb{R}^+$ definerer vi

$$x^r = \sup\{x^t \mid t \in \mathbb{Q}, t \leq r\}.$$

Følgende regler, som er kjent fra skolen, kan utledes fra definisjonen:

Setning 7.3.9: Regler for eksponenter

For alle $a, b \in \mathbb{R}^+$ og $r, s \in \mathbb{R}$ gjelder

(i) $a^r \cdot a^s = a^{r+s}$,

(ii) $\frac{a^r}{a^s} = a^{r-s}$,

(iii) $(a^r)^s = a^{rs}$,

(iv) $(ab)^r = a^r b^r$,

(v) $\left(\frac{a}{b}\right)^r = \frac{a^r}{b^r} = a^r b^{-r}$,

(vi) $1^r = 1$,

(vii) $a^0 = 1$,

(viii) hvis $r < s$, da er $\begin{cases} a^r < a^s, & \text{når } a > 1 \\ a^r > a^s, & \text{når } 0 < a < 1, \end{cases}$

(ix) hvis $a < b$, da er $\begin{cases} a^r < b^r, & \text{når } r > 0 \\ a^r > b^r, & \text{når } r < 0. \end{cases}$

Bevis

Merk at (vii) bare er en gjentagelse av konvensjonen (62). Resten overlates til Oppgave 7.15(c). Dersom $r, s \in \mathbb{Q}$, følger disse reglene av relativt enkel utregning (bruk Hjelpesetning 7.3.3 for å vise (ix)). Å vise at reglene utvides til å gjelde alle $r, s \in \mathbb{R}$ ved hjelp av Definisjon 7.3.8 er den vanskelige biten. $\square$

Merknad 7.3.10. La $r \in \mathbb{R}\backslash\{0\}$. Setning 7.3.9(iii) sier at *potensfunksjonene*

$$\begin{array}{ccc} \mathbb{R}^+ & \longrightarrow & \mathbb{R}^+ \\ x & \mapsto & x^r \end{array}$$

og

$$\begin{array}{ccc} \mathbb{R}^+ & \longrightarrow & \mathbb{R}^+ \\ x & \mapsto & x^{\frac{1}{r}} \end{array}$$

er inverse funksjoner. Ved Setning 7.3.9(ix) er begge strengt voksende for $r > 0$ og strengt avtagende for $r < 0$ (Definisjonene 7.2.5-7.2.6).

Merknad 7.3.11. Fiksér $a \in \mathbb{R}^+ \setminus \{1\}$. Da sier Setning 7.3.9(viii) at *eksponensialfunksjonen*

$$\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \mapsto & a^x \end{array}$$

er strengt voksende hvis $a > 1$ og strengt avtagende hvis $0 < a < 1$. Spesielt er funksjonen *injektiv* (jf. Merknad 7.2.8). Vi skal senere, i Setning 9.5.12, se at verdimengden til funksjonen er $\mathbb{R}^+$, det vil si at det til enhver $y \in \mathbb{R}^+$ finnes en $x \in \mathbb{R}$ slik at $a^x = y$, og denne x er entydig ved injektivitet. Tallet x kalles *logaritmen til y med grunntall (eller base) a* og betegnes med $\log_a y$ (Definisjon 9.5.13). Med andre ord har vi at

$$\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R}^+ \\ x & \mapsto & a^x \end{array}$$

og

$$\begin{array}{ccc} \mathbb{R}^+ & \longrightarrow & \mathbb{R} \\ x & \mapsto & \log_a x \end{array}$$

er inverse funksjoner.

7.4. Følger av reelle tall

Definisjon 6.2.1 av følge gjelder for en vilkårlig mengde X og gjelder derfor også for $X = \mathbb{R}$. Alle følger av rasjonale tall vi studerte i §6.2-6.3 er også eksempler på følger av reelle tall.

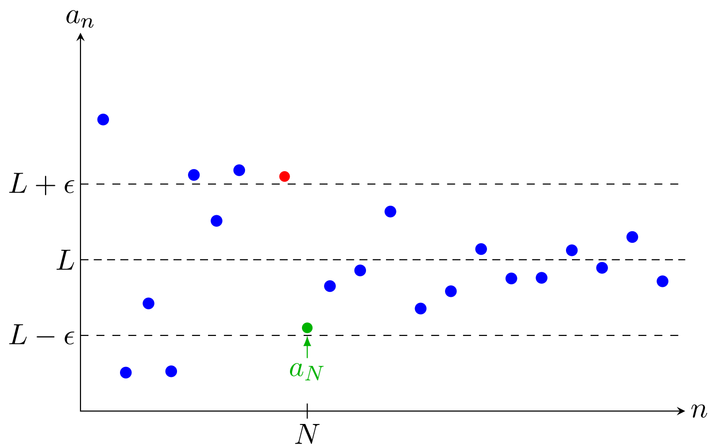
Definisjon 6.2.3 for konvergens av følger i $\mathbb{Q}$ kan også brukes som definisjon på konvergens av følger i $\mathbb{R}$ mot et reelt tall. Vi pleier likevel å gi definisjonen for konvergens i $\mathbb{R}$ ved hjelp av en ϵ som er et positivt *reelt* tall, og ikke bare rasjonalt. På grunn av tetthetsegenskapen til rasjonale tall (Setning 7.2.2) er dette en helt ekvivalent definisjon. Det samme gjelder Definisjon 6.3.1 for Cauchy-følger. Også Definisjon 6.2.10 kunne beholdes som definisjon for begrensethet av reelle tallfølger, men man pleier likevel å gi definisjonen med konstanten $M \in \mathbb{R}$. Vi gjengir for ordens skyld alle disse definisjonene på nytt for reelle følger:

Definisjon 7.4.1: Konvergens av følge

Følgen $\{a_n\}$ av reelle tall *konvergerer* mot det reelle tallet L, dersom det for ethvert reelt tall $\epsilon > 0$ finnes et tall $N \in \mathbb{Z}^+$ slik at $|a_n - L| < \epsilon$ for alle $n \geq N$.

Vi skriver i så fall $\lim\limits_{n\to\infty} a_n = L$ eller “$a_n \to L$ når $n \to \infty$” og kaller L for *grensen* til følgen.

Hvis følgen $\{a_n\}$ ikke konvergerer, sier vi at den *divergerer*.



KONVERGENS AV EN FØLGE: ALLE LEDD FRA OG MED a_N HAVNER I INTERVALLET $(L-\epsilon, L+\epsilon)$

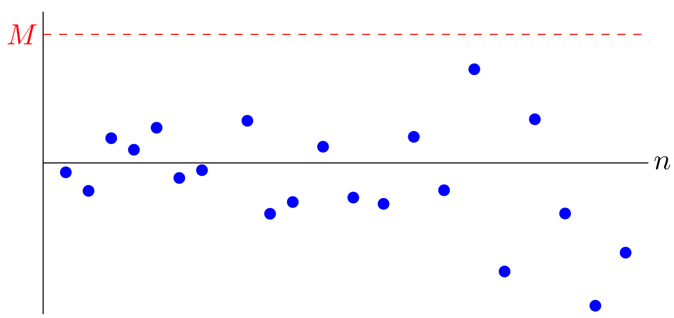
Definisjon 7.4.2: Begrenset følge

En følge $\{a_n\}$ av reelle tall er *begrenset* dersom det finnes en $M \in \mathbb{R}$ slik at $|a_n| \leq M$ for alle $n \in \mathbb{Z}^+$.

En følge er *ubegrenset* dersom den ikke er begrenset.

En følge $\{a_n\}$ av reelle tall er *oppad begrenset* dersom det finnes en $M \in \mathbb{R}$ slik at $a_n \leq M$ for alle $n \in \mathbb{Z}^+$.

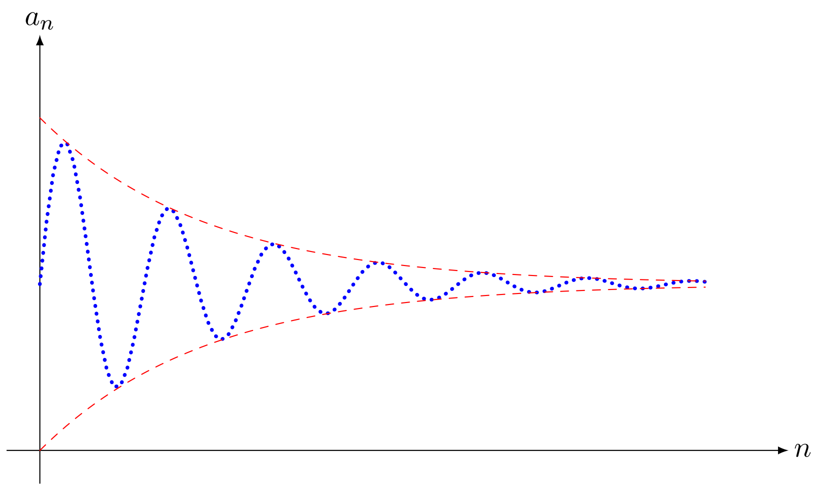
En følge $\{a_n\}$ av reelle tall er *nedad begrenset* dersom det finnes en $M \in \mathbb{R}$ slik at $a_n \geq M$ for alle $n \in \mathbb{Z}^+$.



EN OPPAD BEGRENSET FØLGE

Definisjon 7.4.3: Cauchy-følge

En følge $\{a_n\}$ av reelle tall kalles en *Cauchy-følge* dersom det for ethvert reelt tall $\epsilon > 0$ finnes en $N \in \mathbb{Z}^+$ slik at $|a_n - a_m| < \epsilon$ for alle $m, n \geq N$.



EN CAUCHY-FØLGE

MERKNAD 7.4.4. Igjen vil N i Definisjonene 7.4.1 og 7.4.3 kunne avhenge av ϵ, jf. Merknadene 6.2.8 og 6.3.2.

Eksempel 6.2.9, som viser bruk av definisjonen for å vise en grense av en følge av rasjonale tall, fungerer på samme måte som før for å vise konvergens som følge av reelle tall.

La oss imidlertid se på et eksempel på bruk av definisjonen for å vise at en følge divergerer:

EKSEMPEL 7.4.5. Vis at følgen $\{(-1)^{n+1}\}$ divergerer.

LØSNING. Vi argumenterer ved motsigelse og antar følgen konvergerer, si mot en $L \in \mathbb{R}$. Vi skjønner raskt hva problemet er: Uansett hvor liten $\epsilon > 0$ er, så må tallet L per definisjon ligge i avstand $< \epsilon$ til alle leddene i følgen for stor nok n, det vil si ligge i avstand $< \epsilon$ til begge tallene 1 og -1 samtidig. Siden avstanden mellom disse er 2, er dette umulig for $\epsilon \leq 1$.

Vi lar derfor $\epsilon = 1$. Da vil det finnes en $N \in \mathbb{Z}^+$ slik at

$$\left|(-1)^{n+1} - L\right| < 1 \text{ når } n \geq N. \tag{131}$$

For odde n betyr dette at $|L + 1| < 1$, som er ekvivalent med $L \in (-2, 0)$ og for jevn n betyr (131) at $|L - 1| < \epsilon$, som er ekvivalent med $L \in (0, 2)$. Siden $(-2, 0) \cap (0, 2) = \emptyset$, har vi en motsigelse. $\square$

En del sentrale resultater fra §6.2 om rasjonale tallfølger overføres videre til reelle tallfølger med nøyaktig samme bevis (eneste forskjell er at $\epsilon \in \mathbb{R}^+$ istedenfor $\epsilon \in \mathbb{Q}^+$). Vi repeterer disse resultatene for ordens skyld:

Setning 7.4.6: Entydighet av grenser

Dersom en følge av reelle tall konvergerer mot L og L', da er $L = L'$.

Bevis
Likt beviset for Setning 6.2.6 utledet i Oppgave 6.8. □

Hjelpesetning 7.4.7

Vi har $\lim_{n\to\infty} \frac{1}{n} = 0$ og $\lim_{n\to\infty} c = c$, for enhver $c \in \mathbb{R}$.

Bevis
Likt beviset for Hjelpesetning 6.2.7. □

Setning 7.4.8: Regneregler for grenser av følger

La $\{a_n\}$ og $\{b_n\}$ være konvergente følger av reelle tall. La $a = \lim_{n\to\infty} a_n$ og $b = \lim_{n\to\infty} b_n$. Da gjelder:

(i) $\lim_{n\to\infty} (a_n + b_n) = a + b$;
(ii) $\lim_{n\to\infty} (a_n - b_n) = a - b$;
(iii) $\lim_{n\to\infty} (ca_n) = ca$ for alle $c \in \mathbb{R}$;
(iv) $\lim_{n\to\infty} (a_n b_n) = ab$;
(v) $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}$, dersom $b \neq 0$ og $b_n \neq 0$ for alle $n \in \mathbb{Z}^+$. .

Bevis
Likt beviset for Setning 6.2.14. □

Følgesetning 7.4.9

La P være et polynom av én variabel med reelle koeffisienter (det vil si $P \in \mathbb{R}[x]$ med notasjonen fra §4.1) og $\{a_n\}$ en følge av reelle tall som konvergerer mot a. Da er

$$\lim_{n\to\infty} P(a_n) = P(a).$$

Bevis
Vi har $P(x) = c_0 + c_1 x + \cdots + c_k x^k$ for $k \in \mathbb{N}$ og $c_1, \ldots, c_k \in \mathbb{R}$. Dermed er $P(a_n) = c_0 + c_1 a_n + \cdots + c_k a_n^k$. Ved Setning 7.4.8(iv) har vi at

$$\begin{aligned}
\lim_{n\to\infty} a_n^2 &= a^2, \\
\lim_{n\to\infty} a_n^3 &= \lim_{n\to\infty} \left((a_n^2) \cdot a_n\right) = a^2 \cdot a = a^3, \\
&\cdots \\
\lim_{n\to\infty} a_n^k &= \lim_{n\to\infty} \left((a_n^{k-1}) \cdot a_n\right) = a^{k-1} \cdot a = a^k.
\end{aligned}$$

Ved Setning 7.4.8(iii)-(iv) har vi derfor at $\lim_{n\to\infty} c_i a_n^i = c_i a^i$ for alle $i = 0, \dots, k$, og ved Setning 7.4.8(i) får vi derfor

$$\lim_{n\to\infty} P(a_n) = \lim_{n\to\infty} c_0 + c_1 a_n + \cdots + c_k a_n^k = c_0 + c_1 a + \cdots + c_k a^k = P(a),$$

som fullfører beviset. □

Setning 7.4.10

En konvergent følge av reelle tall er begrenset.

Bevis
Likt beviset for Setning 6.2.11. □

Eksempel 6.2.13 fungerer fremdeles som et eksempel på en ubegrenset følge, som derfor er divergent.

Merknad 7.4.11. Det motsatte av Setning 7.4.10 er ikke sant: for eksempel er følgen i Eksempel 7.4.5 begrenset, men divergent.

Setning 7.4.12

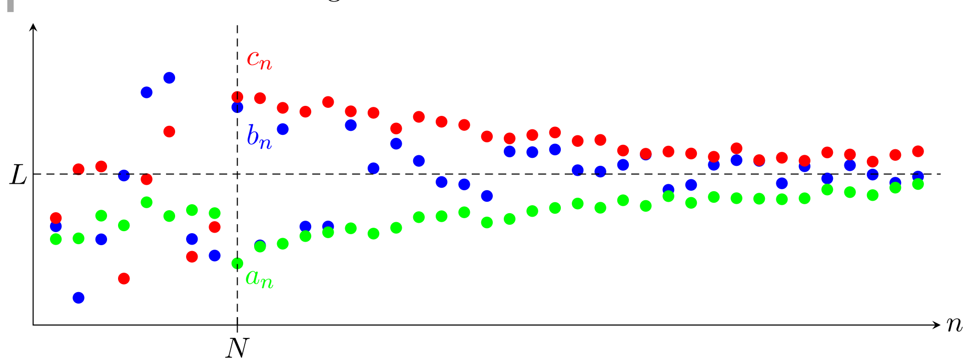
En Cauchy-følge av reelle tall er begrenset.

Bevis
Likt beviset utledet for rasjonale følger i Oppgave 6.13. □

Setning 7.4.13: Skvisesetningen for følger

Anta $\{a_n\}$, $\{b_n\}$ og $\{c_n\}$ er følger av reelle tall slik at $a_n \le b_n \le c_n$ for alle $n \ge N$, hvor $N \in \mathbb{Z}^+$. Hvis $\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = L$, da er $\lim_{n\to\infty} b_n = L$.

Bevis
Likt beviset for Setning 6.2.17 □



Skvisesetning for følger

Vi tar også med en variant:

Setning 7.4.14: Grenser bevarer orden

La $\{a_n\}$ og $\{b_n\}$ være konvergente følger av reelle tall slik at $a_n \leq b_n$ for alle $n \geq N$, hvor $N \in \mathbb{Z}^+$.

Da er $\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n$.

BEVIS

Overlates til Oppgave 7.17(a). □

Det finnes noen vesentlige forskjeller mellom teorien for følger i $\mathbb{Q}$ og i $\mathbb{R}$. Vi vil se på disse i de neste to seksjonene.

Eksempler. La oss se på noen konkrete eksempler.

EKSEMPEL 7.4.15. Finn $\lim_{n\to\infty} \dfrac{2-8n^3}{4n^3+2n^2-n}$.

LØSNING. Begge følgene $2-8n^3$ og $4n^3+2n^2-n$ er ubegrenset og dermed divergente (Setning 7.4.10), og vi kan derfor ikke bruke Setning 7.4.8(v). I eksempler av denne typen lønner det seg ofte å dividere med den høyeste potensen av n i teller og nevner, som i dette tilfellet er n^3:

$$\frac{2-8n^3}{4n^3+2n^2-n} = \frac{\frac{2}{n^3}-8}{4+\frac{2}{n}-\frac{1}{n^2}}.$$

Ved Hjelpesetning 7.4.7 og Setning 7.4.8(iii)-(iv) går alle $\frac{2}{n^3}$, $\frac{2}{n}$ og $\frac{1}{n^2}$ mot 0, slik at teller går mot -8 og nevner går mot 4, igjen ved Setning 7.4.8. Ved Setning 7.4.8(v), har vi derfor

$$\lim_{n\to\infty} \frac{2-8n^3}{4n^3+2n^2-n} = \lim_{n\to\infty} \frac{\frac{2}{n^3}-8}{4+\frac{2}{n}-\frac{1}{n^2}} = \frac{-8}{4} = -2.$$

□

EKSEMPEL 7.4.16. Finn $\lim_{n\to\infty} \dfrac{\cos n}{n}$.

LØSNING. Siden $-1 \leq \cos n \leq 1$ for alle $n \in \mathbb{Z}^+$, har vi $-\frac{1}{n} \leq \frac{\cos n}{n} \leq \frac{1}{n}$ for alle $n \in \mathbb{Z}^+$. Ved Skvisesetningen 7.4.13 og Hjelpesetning 7.4.7, har vi $\lim_{n\to\infty} \dfrac{\cos n}{n} = 0$. □

EKSEMPEL 7.4.17. Finn $\lim_{n\to\infty} \dfrac{1}{\sqrt{n}}$.

Løsning. Siden $\sqrt{n}$ vokser seg så stor vi bare vil, aner vi at grensen er 0. Vi forsøker å vise dette ved definisjonen.

Gitt $\epsilon > 0$. Vi vil finne $N \in \mathbb{Z}^+$ slik at $\frac{1}{\sqrt{n}} = \left|\frac{1}{\sqrt{n}} - 0\right| < \epsilon$ når $n \geq N$. Vi ser at

$$\frac{1}{\sqrt{n}} < \epsilon \Longleftrightarrow \frac{1}{\epsilon} < \sqrt{n} \Longleftrightarrow \frac{1}{\epsilon^2} < n.$$

Derfor holder det å ta $N > \frac{1}{\epsilon^2}$. At det finnes en slik N følger av Arkimedes' prinsipp (7.2.1). □

Eksempel 7.4.18. Finn $\lim_{n\to\infty} \sqrt{n+2} - \sqrt{n}$.

Løsning. Siden begge følgene $\sqrt{n+2}$ og $\sqrt{n}$ er ubegrenset, er de divergente (Setning 7.4.10), og vi kan derfor ikke bruke Setning 7.4.8(ii). Vi gjør derfor følgende "triks":

$$\begin{aligned}\sqrt{n+2} - \sqrt{n} &= \frac{\left(\sqrt{n+2} - \sqrt{n}\right)\left(\sqrt{n+2} + \sqrt{n}\right)}{\sqrt{n+2} + \sqrt{n}} \\ &= \frac{n+2-n}{\sqrt{n+2} + \sqrt{n}} = \frac{2}{\sqrt{n+2} + \sqrt{n}}.\end{aligned}$$

Dette viser at

$$0 \leq \sqrt{n+2} - \sqrt{n} \leq \frac{2}{2\sqrt{n}} = \frac{1}{\sqrt{n}}.$$

Siden $\lim_{n\to\infty} \frac{1}{\sqrt{n}} = 0$ ved Eksempel 7.4.17, gir skvisesetningen (7.4.13) at $\lim_{n\to\infty} \sqrt{n+2} - \sqrt{n} = 0$. □

Eksempel 7.4.19. Vis at $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

Løsning. Denne er litt vrien. Definér $x_n = \sqrt[n]{n} - 1$. Vi vil vise at $\lim_{n\to\infty} x_n = 0$. Da vil $\lim_{n\to\infty} \sqrt[n]{n} = \lim_{n\to\infty} (x_n + 1) = 1$ ved Setning 7.4.8(i).

Siden $1 \leq n$, har vi $1 = \sqrt[n]{1} \leq \sqrt[n]{n}$ (ved Hjelpesetning 7.3.3), slik at $x_n \geq 0$. Ved *Binomialteoremet* 3.8.7 har vi

$$\begin{aligned} n &= (x_n + 1)^n \\ &= x_n^n + n x_n^{n-1} + \frac{n(n-1)}{2} x_n^{n-2} + \cdots + \frac{n(n-1)}{2} x_n^2 + n x_n + 1 \\ &\geq \frac{n(n-1)}{2} x_n^2,\end{aligned}$$

slik at $x_n \leq \sqrt{\frac{2}{n-1}}$. Vi har altså vist at

$$0 \leq x_n \leq \sqrt{\frac{2}{n-1}} = \sqrt{2} \cdot \frac{1}{\sqrt{n-1}}. \tag{132}$$

Siden $\lim_{n\to\infty} \frac{1}{\sqrt{n-1}} = 0$ (ved Eksempel 7.4.17), er $\lim_{n\to\infty} \left(\sqrt{2} \cdot \frac{1}{\sqrt{n-1}}\right) = 0$ ved Setning 7.4.8(iii). Ved Skvisesetningen 7.4.13 på (132) får vi den ønskede grensen $\lim_{n\to\infty} x_n = 0$. □

Uendelige grenser. Per definisjon er en følge $\{a_n\}$ av reelle tall ikke oppad begrenset dersom det til enhver $M \in \mathbb{R}$ finnes en $n \in \mathbb{Z}^+$ slik at $a_n > M$. Dette gjelder for eksempel følgene

$$a_n = n^2 \tag{133}$$

$$a_n = \begin{cases} n^2, & \text{hvis } n \text{ er partall,} \\ 0, & \text{hvis } n \text{ er oddetall.} \end{cases} \tag{134}$$

Gitt en vilkårlig $M \in \mathbb{R}$, vil i tilfellet (133) *enhver* $n \in \mathbb{Z}^+$ med $n > \sqrt{M}$ oppfylle $a_n > M$, mens i tilfellet (134) vil dette være oppfylt kun for *partallene* n med $n > \sqrt{M}$. Oppførselen er altså forskjellig, siden vi i (133) kan få *alle* leddene langt nok ut i følgen så store vi vil, mens vi i (134) kun kan få *noen* av leddene langt nok ut i følgen så store vi vil. Begge følgene er ubegrenset og dermed divergente, men i tilfellet (133) sier vi at *følgen divergerer mot uendelig.* Vi har tilsvarende fenomen og definisjon i det negative tilfellet:

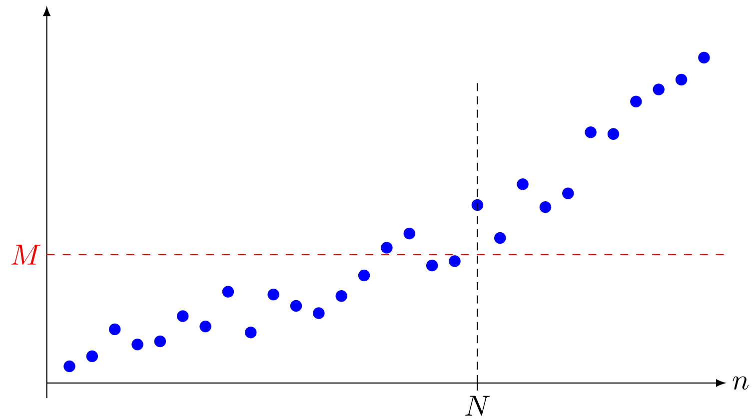
Definisjon 7.4.20

Følgen $\{a_n\}$ av reelle tall *divergerer mot uendelig*, dersom det for ethvert reelt tall M finnes en $N \in \mathbb{Z}^+$ slik at $a_n > M$ for alle $n \geq N$.

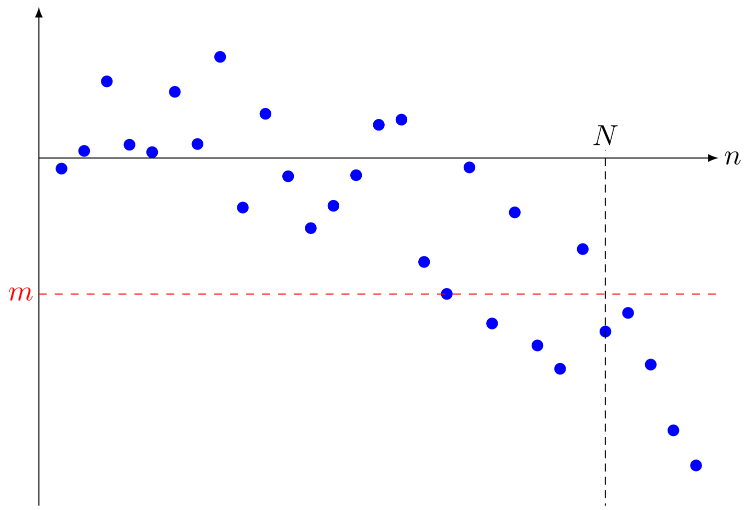
Vi skriver i så fall $\lim_{n\to\infty} a_n = \infty$ eller $a_n \to \infty$.

Følgen $\{a_n\}$ av reelle tall *divergerer mot minus uendelig* , dersom det for ethvert reelt tall m finnes en $N \in \mathbb{Z}^+$ slik at $a_n < m$ for alle $n \geq N$.

Vi skriver i så fall $\lim_{n\to\infty} a_n = -\infty$ eller $a_n \to -\infty$.



DIVERGENS MOT UENDELIG



DIVERGENS MOT MINUS UENDELIG

Vi bemerker at divergens mot $\pm\infty$ er en spesiell type divergens (nærmere bestemt en spesiell type ubegrensethet), selv om vi skriver og sier at "grensen er $\pm\infty$".

EKSEMPEL 7.4.21. Finn $\displaystyle\lim_{n\to\infty} \frac{3n^4 - 6n^2 + 5}{2n^3 + n + 2}$.

LØSNING. Begge følgene $3n^4 - 6n^2 + 5$ og $2n^3 + n + 2$ er ubegrenset og dermed divergente (Setning 7.4.10). Vi merker oss at høyeste potens av n er n^4 i teller og n^3 i nevner, og disse vil dominere de andre leddene når n er stor. For stor n vil altså uttrykket ligne mer og mer på $\frac{3n^4}{2n^3} = \frac{3}{2}n$ som blir så stor vi vil for stor nok n. Det lukter derfor divergens mot uendelig. La oss vise dette skikkelig.

Vi merker at

$$2n^3 + n + 2 \leq 2n^3 + n^3 + 2n^3 = 5n^3 \;\text{ for alle }\; n \in \mathbb{Z}^+$$

og

$$3n^4 - 6n^2 + 5 \geq 3n^4 - n^4 = 2n^4 \text{ for alle } n \geq 3.$$

Dermed har vi at

$$\frac{3n^4 - 2n^2 + 5}{n^3 + n + 2} \geq \frac{2n^4}{5n^3} = \frac{2}{5}n \text{ for alle } n \geq 3.$$

Gitt $M \in \mathbb{R}$, vil derfor

$$\frac{3n^4 - 2n^2 + 5}{n^3 + n + 2} \geq \frac{2n^4}{5n^3} = \frac{2}{5}n \geq M$$

for $n \geq \max\{3, \frac{5}{2}M\}$, som viser at $\lim\limits_{n\to\infty} \dfrac{3n^4 - 2n^2 + 5}{n^3 + n + 2} = \infty$. □

Det finnes ikke helt tilsvarende regneregler for uendelige grenser slik som Setning 7.4.8. Generelt må vi være forsiktige med uendelige grenser og huske at $\pm\infty$ overhodet ikke er tall, som følgende eksempel viser.

EKSEMPEL 7.4.22. La $\{a_n\} = \{n\}$ og $\{b_n\} = \{\frac{1}{n}\}$. Følgene oppfyller $\lim\limits_{n\to\infty} a_n = \infty$ og $\lim\limits_{n\to\infty} b_n = 0$. Siden $a_n b_n = 1$ har vi at

$$\lim_{n\to\infty} a_n b_n = 1.$$

Det ville altså vært helt feil å anvende Setning 7.4.8(iv) og konkludere at

$$\lim_{n\to\infty} a_n b_n = \infty \cdot 0 = 0.$$

(Dette er jo til å få frysninger i ryggen av!)

Rekker. Vi avslutter seksjonen med noen få ord om såkalte *rekker*.

Til en følge av reelle tall $\{a_n\}$, kan vi assosiere en ny følge $\{s_n\}$, hvor

$$s_n = \sum_{k=1}^{n} a_n.$$

Leddet s_n kalles *den nte delsummen*[5] *til følgen* $\{a_n\}$ og følgen $\{s_n\}$ kalles en *uendelig rekke*[6] eller bare *rekke*[7]. For $\{s_n\}$ bruker vi også notasjonen

$$a_1 + a_2 + a_3 + \cdots$$

eller, mer konsist,

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

[5] *Partial sum* på engelsk.
[6] *Infinite series* på engelsk.
[7] *Series* på engelsk.

Hvis følgen $\{s_n\}$ konvergerer mot et tall s, sier vi at *rekken konvergerer (mot s)* og at rekken er *konvergent*. Tallet s kaller vi også for *summen av rekken* og vi skriver

$$\sum_{n=1}^{\infty} a_n = s$$

eller

$$a_1 + a_2 + a_3 + \cdots = s.$$

Det er viktig å være klar over at denne notasjonen betyr at s er grensen til følgen $\{s_n\}$ og ikke "en sum av uendelig mange tall", som ikke er definert på noen som helst måte.

Hvis følgen $\{s_n\}$ ikke konvergerer, sier vi at rekken er *divergent* eller *divergerer*.

Legg merke til at symbolet $\sum_{n=1}^{\infty} a_n$ brukes *både* for selve rekken (som altså er følgen av delsummer $\{s_n\}$) og for dens sum (som altså er *grensen* til følgen av delsummer $\{s_n\}$). Dette er en veldig innarbeidet dobbeltbruk, men kan føre til misforståelser: dersom vi for eksempel skriver $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} b_n$, er det ikke klart om vi mener at rekkene er like (som betyr at $a_n = b_n$ for alle n), eller bare at de har samme sum.

Det er også viktig å være klar over at en rekke er et spesialtilfelle av en følge, slik at all teori om følger overføres til rekker. Det finnes likevel en del flere resultater om rekker. Vi viser til Oppgavene 7.29 og 7.30 for et par av dem, men kommer i denne boken ikke til å betrakte rekker noe særlig mer. Dere kommer til å møte på rekker, inkludert mange resultater som garanterer konvergens av rekker, i kursene *MAT112–Grunnkurs i Matematikk II* og *MAT211–Reell Analyse*. Dessuten er rekker svært nyttige i søken på løsninger til differensialligninger, som dere vil se i kurset *MAT131–Differensialligninger* og andre kurs som omhandler differensialligninger.

7.5. Voksende og avtagende følger i $\mathbb{R}$

Som nevnt ovenfor, finnes det noen vesentlige forskjeller mellom teorien for følger i $\mathbb{Q}$ og i $\mathbb{R}$, og vi skal se på dem i denne og neste seksjon. Vi trenger først en definisjon.

Definisjon 7.5.1: Voksende følger

En følge $\{a_n\}$ av reelle tall er *voksende*[a] (henholdsvis *strengt voksende*[b]) dersom $a_{n+1} \geq a_n$ (hhv. $a_{n+1} > a_n$) for alle $n \in \mathbb{Z}^+$. (Ved Definisjon 7.2.5 er dette det samme som å si at den tilsvarende funksjonen $\mathbb{Z}^+ \to \mathbb{R}$ gitt ved $n \mapsto a_n$ er voksende (hhv. strengt voksende).)

[a] *Increasing* på engelsk.

[b] *Strictly increasing* på engelsk.

Definisjon 7.5.2: Avtagende følger

En følge $\{a_n\}$ av reelle tall er *avtagende*[a] (henholdsvis *strengt avtagende*[b]) dersom $a_{n+1} \leq a_n$ for alle $n \in \mathbb{Z}^+$. (Ved Definisjon 7.2.6 er dette det samme som å si at den tilsvarende funksjonen $\mathbb{Z}^+ \to \mathbb{R}$ gitt ved $n \mapsto a_n$ er avtagende (hhv. strengt avtagende).)

[a] *Decreasing* på engelsk.
[b] *Strictly decreasing* på engelsk.

Definisjon 7.5.3: Monotone følger

En følge $\{a_n\}$ av reelle tall er *monoton*[a] (henholdsvis *strengt monoton*[b]) dersom den er voksende eller avtagende (hhv. strengt voksende eller strengt avtagende). (Ved Definisjon 7.2.7 er dette det samme som å si at den tilsvarende funksjonen $\mathbb{Z}^+ \to \mathbb{R}$ gitt ved $n \mapsto a_n$ er monoton (hhv. strengt monoton).)

[a] *Monotone* på engelsk.
[b] *Strictly monotone* på engelsk.

Det neste resultatet har ikke noe motstykke i $\mathbb{Q}$:

Setning 7.5.4: Monotone begrensede følger konvergerer

(i) La $\{a_n\}$ være en voksende oppad begrenset følge av reelle tall. Da vil følgen konvergere mot $\sup\{a_n \mid n \in \mathbb{Z}^+\}$.

(ii) La $\{a_n\}$ være en avtagende nedad begrenset følge av reelle tall. Da vil følgen konvergere mot $\inf\{a_n \mid n \in \mathbb{Z}^+\}$.

Bevis

(i) Betrakt mengden

$$A = \{a_n \mid n \in \mathbb{Z}^+\} \subset \mathbb{R}$$

av alle leddene i følgen $\{a_n\}$. Per antagelse er A ikke-tom og oppad begrenset. Siden $\mathbb{R}$ oppfyller kompletthetsprinsippet, eksisterer den minste øvre skranken $L = \sup A = \sup\{a_n \mid n \in \mathbb{Z}^+\}$. For å vise at $\lim_{n\to\infty} a_n = L$, må vi finne, til enhver gitt $\epsilon > 0$, en $N \in \mathbb{Z}^+$ slik at $|a_n - L| < \epsilon$ for alle $n \geq N$.

Siden L er den *minste* øvre skranken til A, så må det finnes et ledd a_N slik at $L - \epsilon < a_N \leq L$, for hvis ikke ville $L - \epsilon$ være en øvre skranke for A som er mindre enn den minste øvre skranken. Siden $\{a_n\}$ er voksende, er $a_n \geq a_N$ for alle $n \geq N$. Dermed har vi vist at

$$L - \epsilon < a_N \leq a_n \leq L$$

for alle $n \geq N$, som medfører at

$$|a_n - L| < \epsilon \text{ for alle } n \geq N,$$

som er det vi skulle vise.

Beviset for (ii) er nesten helt identisk og overlates til Oppgave 7.32.
□

Eksempler. Kriteriet gitt i siste setning er et ekstremt nyttig verktøy. La oss se på et par eksempler på bruk:

EKSEMPEL 7.5.5. Avgjør om følgen $\{x_n\}$ definert *rekursivt* ved $x_1 = 3$ og

$$x_{n+1} = \sqrt{3x_n - 2} \tag{135}$$

konvergerer og finn eventuelt grensen.

LØSNING. Legg merke til at denne definisjonen bestemmer hva x_n er for alle n: det første leddet x_1 er oppgitt, og så kan vi bruke formelen (135) til å bestemme x_2, og deretter x_3 osv.

La oss først forsøke å finne hva grensen til følgen må være, *dersom den finnes*. Vi setter $L = \lim_{n\to\infty} x_n$. Vi merker at $x_n \geq 0$ for alle n ved (135). Vi kvadrerer begge sider av (135), og finner

$$x_{n+1}^2 = 3x_n - 2. \tag{136}$$

Vi merker at $\lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1}$, siden følgene $\{x_n\}$ og $\{x_{n+1}\}$ har de samme leddene, forskjøvet med 1, slik at definisjonen av grense viser at de må ha samme grenseverdi. Tar vi grenser på begge sider av (136) og bruker reglene i Setning 7.4.8, finner vi

$$L^2 = 3L - 2.$$

Løser vi denne annengradsligningen, finner vi $L = 1$ eller 2. Dette er altså de eneste kandidatene for grensen. Så langt kan vi imidlertid ikke si noe som helst om hvorvidt følgen faktisk konvergerer.

Nå kan det være lurt å regne ut noen ledd i følgen, for å få en følelse av hvordan følgen oppfører seg:

$$\begin{aligned} x_1 &= 3 \\ x_2 &= \sqrt{3 \cdot 3 - 2} = \sqrt{7} < 3 = x_1 \\ x_3 &= \sqrt{3\sqrt{7} - 2} < \sqrt{7} = x_2, \end{aligned}$$

som tyder på at følgen avtar. (Dette er kompatibelt med begge kandidatene 1 og 2 for grensen.) Hvis vi klarer å vise at følgen virkelig er avtagende, og dessuten nedad begrenset, da vil vi ha vist at den må konvergere, ved Setning 7.5.4(ii). Dersom vi viser at 2 er en nedre skranke, det vil si at $a_n \geq 2$ for alle n, da må grensen være 2 (enten fordi Setning 7.5.4(ii) sier at grensen må være den største nedre skranken til følgen, som da ikke kan være 1, eller siden grenser bevarer orden (Setning 7.4.14). Dersom a_n "kommer seg under 2", det vil si $a_n < 2$ for en n, da må av samme grunn grensen være 1.

La oss tippe at vi er i det første tilfellet, slik at vi vil vise at $\{x_n\}$ er avtagende og nedad begrenset av 2, det vil si:

$$2 \leq x_{n+1} \leq x_n \quad \text{for alle} \quad n \geq 1 \tag{137}$$

Dette er jo et typisk utsagn man vil forsøke å vise ved induksjon.

Vi har allerede vist at tilfellet $n = 1$ er sant:

$$2 \leq \sqrt{7} = x_2 \leq 3 = x_1.$$

Anta nå at (137) er sant for $n = k$, der $k \geq 1$ er et heltall, det vil si

$$2 \leq x_{k+1} \leq x_k \quad \text{for en} \quad k \geq 1. \tag{138}$$

Det følger at

$$4 = 3 \cdot 2 - 2 \leq 3x_{k+1} - 2 \leq 3x_k - 2$$

og dermed at

$$2 = \sqrt{4} \leq \sqrt{3x_{k+1} - 2} \leq \sqrt{3x_k - 2}$$

ved Hjelpesetning 7.3.3. Dette er ekvivalent med

$$2 \leq x_{k+2} \leq x_{k+1},$$

som er (137) med $n = k + 1$. Vi har dermed vist at (137) er sant ved induksjon.

Dermed er $\{x_n\}$ avtagende og nedad begrenset, og dermed konvergent ved Setning 7.5.4(ii). Siden grenser bevarer orden (Setning 7.4.14), gir (137) at $2 \leq L = \lim_{n\to\infty} x_n$. Siden vi har vist at $L = 1$ eller 2, må vi ha at $L = 2$. □

La oss også gi et bevis på at følgen fra Setning 6.3.4, som ikke konvergerer i $\mathbb{Q}$, faktisk konvergerer mot $\sqrt{2}$ i $\mathbb{R}$:

Eksempel 7.5.6. Cauchy-følgen fra Setning 6.3.4, definert rekursivt ved

$$p_{n+1} = p_n - \frac{p_n^2 - 2}{p_n + 2} = 2 - \frac{2}{p_n + 2}, \tag{139}$$

med en vilkårlig valgt $p_1 \in \mathbb{R}^+$, konvergerer mot $\sqrt{2}$.

Bevis. Fra høyre likhet i (139) og induksjon ser vi at $p_n > 0$ for alle $n \in \mathbb{Z}^+$ (som i Oppgave 6.11(a)). Høyre likhet i (139) gir derfor også at $p_n < 2$ for alle $n \in \mathbb{Z}^+$. Altså er følgen $\{p_n\}$ både oppad og nedad begrenset.

Siden alle $p_n > 0$, gir høyre likhet i (139) og litt regning derfor at

$$p_{n+1} = 2 - \frac{2}{p_n + 2} \geq 2 - \frac{2}{p_{n-1} + 2} = p_n \iff p_n \geq p_{n-1}, \tag{140}$$

og likeledes

$$p_{n+1} = 2 - \frac{2}{p_n + 2} \leq 2 - \frac{2}{p_{n-1} + 2} = p_n \iff p_n \leq p_{n-1}. \tag{141}$$

Fra venstre likhet i (139) med $n = 1$ ser vi at

$$p_2 \geq p_1 \iff p_1 \leq \sqrt{2} \tag{142}$$

og likeledes

$$p_2 \leq p_1 \iff p_1 \geq \sqrt{2}. \tag{143}$$

Altså sier (140)-(143) oss, sammen med induksjon, at

$$\{p_n\} \text{ er voksende (henholdsvis avtagende)} \iff p_1 \leq \sqrt{2} \text{ (hhv. } p_1 \geq \sqrt{2}).$$

Følgen $\{p_n\}$ er derfor voksende eller avtagende, det vil si monoton, og både oppad og nedad begrenset. Den konvergerer derfor mot en grense $L \in \mathbb{R}$ ved Setning 7.5.4. Siden alle $p_n > 0$ og grenser bevarer orden (Setning 7.4.14), må også $L \geq 0$. Ved å ta grenser på begge sider av (139) som i beviset for Setning 6.3.4, finner vi som før at $L^2 = 2$. Dermed er $L = \sqrt{2}$ ved Definisjon 7.3.2. □

Eulertallet e**.** Som en ytterligere anvendelse av Setning 7.5.4 vil vi utlede eksistensen og definisjonen av eulertallet e, en av de aller viktigste konstantene i matematikken.

Betrakt følgen $\left\{\left(1+\frac{1}{n}\right)^n\right\}$. Vi overlater det til Oppgave 7.41 å vise at den er voksende og oppad begrenset av 3. Følgen er derfor konvergent ved Setning 7.5.4.

Definisjon 7.5.7

Vi definerer

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

og kaller e for *eulertallet*.

Merknad 7.5.8. Det går også an å vise, jf. Oppgave 7.42, at

$$e = \lim_{n\to\infty}\left(1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{n!}\right) = \lim_{n\to\infty}\sum_{k=0}^{n}\frac{1}{k!},$$

og dette kan betraktes som en ekvivalent definisjon på e. Med notasjonen for rekker fra slutten av §7.4 kan vi skrive

$$e = \sum_{n=0}^{\infty}\frac{1}{n!} = 1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\cdots.$$

Siden

$$2 \leq \left(1+\frac{1}{n}\right)^n \leq 3$$

for alle $n \in \mathbb{Z}^+$, vet vi at $2 \leq e \leq 3$ ved Setning 7.4.14. Tallet e ble "oppdaget" av den sveitsiske matematikeren Jacob Bernoulli (1654–1705) i 1683 i forbindelse med hans studium av *rentes rente*[8], som vi skal forklare

[8] *Compound interest* på engelsk.

nedenunder. Betegnelsen e ble introdusert av den sveitsiske matematikeren Leonhard Euler (1707–1783) i to-bindsverket *Mechanica* i 1736. Han viste at e er irrasjonalt i 1737, noe vi skal vise i Oppgave 7.42, og beregnet de første 23 desimalene i sin berømte *Introductio in analysin infinitorum*[9] fra 1748, nemlig

$$e = 2,71828182845904523536028\ldots$$

(Vi skal komme til desimaltallsrepresentasjoner av reelle tall i §7.8.) Av disse grunnene bærer konstanten navnet *eulertallet*.

AC LOGARITHM. PER SERIES EXPLICAT. 93

tero Logarithmorum hyperbolicorum usus in calculo integrali Cap.VII. fusius demonstrabitur.

CAPUT VIII.

De quantitatibus transcendentibus ex Circulo ortis.

126. Post Logarithmos & quantitates exponentiales considerari debent Arcus circulares eorumque Sinus & Cosinus, quia non solum aliud quantitatum transcendentium genus constituunt, sed etiam ex ipsis Logarithmis & exponentialibus, quando imaginariis quantitatibus involvuntur, proveniunt, id quod infra clarius patebit.

Ponamus ergo Radium Circuli seu Sinum totum esse = 1, atque satis liquet Peripheriam hujus Circuli in numeris rationalibus exacte exprimi non posse, per approximationes autem inventa est Semicircumferentia hujus Circuli esse = 3,[illegible] +, pro quo numero, brevitatis ergo, scribam π, ita ut sit π = Semicircumferentiæ Circuli, cujus Radius = 1, seu π erit longitudo Arcus 180 graduum.

127. Denotante z Arcum hujus Circuli quemcunque, cujus Radium perpetuo assumo = 1, hujus Arcus z considerari potissimum solent Sinus & Cosinus. Sinum autem Arcus z in posterum hoc modo indicabo *sin. A.* z, seu tantum *sin.* z. Cosinum vero hoc modo *cos. A.* z, seu tantum *cos.* z. Ita, cum π sit Arcus 180°, erit *sin.* $0\pi = 0$; *cos.* $0\pi = 1$; & *sin.* $\frac{1}{2}\pi = 1$, *cos.* $\frac{1}{2}\pi = 0$; *sin.* $\pi = 0$; *cos.* $\pi = -1$; *sin.* $\frac{3}{2}\pi = -1$; *cos.* $\frac{3}{2}\pi = 0$; *sin.* $2\pi = 0$; & *cos.* $2\pi = 1$.

Omnes ergo Sinus & Cosinus intra limites +1 & −1 continen-

M 3

SIDE FRA *Introductio in analysin infinitorum* (KILDE: WIKIPEDIA, PULIC DOMAIN)

MERKNAD 7.5.9. En klassisk huskeregel for de første desimalene til e er "2, 7IbsenIbsen", siden Henrik Ibsen ble født i 1828. De neste 6 desimalene 459045 er lette å huske (som "en halv rett vinkel, en rett vinkel, og igjen en halv rett vinkel").

Mer generelt går det an å vise (jf. Oppgave 9.19) at

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x \quad \text{for enhver } x \in \mathbb{R} \tag{144}$$

og denne grensen har stor betydning i matematikk og naturvitenskap, som vi nå kort forklarer.

Mange prosesser i naturen og i samfunnet følger "naturloven" at en størrelse endrer seg proporsjonalt med størrelsen selv. Et eksempel er kapital i banken. Anta at man har et innskudd U i en bankkonto ved begynnelsen av året, at årsrenten er r (i prosent) og at banken legger rentebeløpet til

[9] Latin for "Introduksjon til infinitesimal analyse".

saldoen mot slutten av året. Da vil saldoen i begynnelsen av neste kalenderår være $\left(1+\frac{r}{100}\right)U$. Anta istedenfor at banken "fordeler renten" på hver måned med rente $\frac{r}{12}$ og legger rentebeløpet til saldoen hver måned, slik at den tillagte renten også vil forrentes. Dette kalles *rentes rente*. Hva vil saldoen da være mot slutten av året? Det var dette problemet Bernoulli studerte og som ledet frem til "oppdagelsen" av e.

For å besvare spørsmålet, la oss anta at året deles inn i n like store deler (måneder, uker, dager,...), som kalles *terminer* og at $\frac{r}{n}$ er renten per termin i prosent. La u_j være saldoen etter j terminer. Da har vi

$$
\begin{aligned}
u_1 &= \left(1+\frac{r}{n}\right)U \\
u_2 &= \left(1+\frac{r}{n}\right)u_1 = \left(1+\frac{r}{n}\right)^2 U \\
u_3 &= \left(1+\frac{r}{n}\right)u_2 = \left(1+\frac{r}{n}\right)^3 U \\
\dots &\quad \dots \\
u_n &= \left(1+\frac{r}{n}\right)u_{n-1} = \left(1+\frac{r}{n}\right)^n U,
\end{aligned}
$$

hvor u_n er saldoen etter ett år. Her er $n = 12$ hvis renten legges til hver måned, $n = 52$ hvis den legges til hver uke, $n = 365$ hvis den legges til hver dag, osv. Vi ser at hvis terminene blir kortere og kortere (som vil tilsvare at renten legges til "hvert øyeblikk"), så vil ved (144) saldoen nærme seg

$$\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n U = e^r U.$$

JACOB BERNOULLI (1654–1705) (KILDE: WIKIPEDIA, PULIC DOMAIN)

7.6. Cauchy-følger i $\mathbb{R}$

En annen vesentlig forskjell mellom $\mathbb{Q}$ og $\mathbb{R}$ er oppførselen til Cauchy-følger: I $\mathbb{Q}$ så er det å være en Cauchy-følge en nødvendig, men ikke tilstrekkelig, betingelse for å være en konvergent følge, ved Setningene 6.3.3 og 6.3.4; i $\mathbb{R}$ er betingelsen imidlertid også tilstrekkelig:

Setning 7.6.1: Cauchy-kriteriet for konvergens

En følge av reelle tall konvergerer hvis og bare hvis den er en Cauchy-følge.

Bevis

Beviset for at en konvergent følge av reelle tall er Cauchy er likt beviset for Setning 6.3.3.

La oss nå vise at en Cauchy-følge av reelle tall konvergerer. Dette beviset er av det mer krevende slaget.

La $\{a_n\}$ være en Cauchy-følge. Ved Setning 7.4.12 er $\{a_n\}$ begrenset, det vil si at det finnes en $M \in \mathbb{R}$ slik at $a_n \in [-M, M]$ for alle $n \in \mathbb{Z}^+$. La $d = 2M$, som altså er lengden på intervallet $[-M, M]$. Siden $\{a_n\}$ er Cauchy, finnes en $N_1 \in \mathbb{Z}^+$ slik at $|a_n - a_m| < \frac{d}{2}$ for alle $m, n \geq N_1$. Spesielt finnes det et delintervall $[x_1, y_1] \subset [-M, M]$ av lengde $y_1 - x_1 = \frac{d}{2}$ slik at $a_n \in [x_1, y_1]$ for alle $n \geq N_1$. Vi har altså halvert det opprinnelige intervallet $[-M, M]$. Vi fortsetter prosedyren: det finnes en $N_2 \in \mathbb{Z}^+$ slik at $|a_n - a_m| < \frac{d}{4}$ for alle $m, n \geq N_2$. Spesielt finnes det et delintervall $[x_2, y_2] \subset [x_1, y_1]$ av lengde $y_2 - x_2 = \frac{d}{4}$ slik at $a_n \in [x_2, y_2]$ for alle $n \geq N_2$. Vi gjentar prosedyren så ofte vi vil: når vi har funnet et heltall N_k og et delintervall $[x_k, y_k]$ av lengde $y_k - x_k = \frac{d}{2^k}$ slik at $a_n \in [x_k, y_k]$ for alle $n \geq N_k$, da kan vi finne en ny $N_{k+1} \in \mathbb{Z}^+$ og et nytt delintervall $[x_{k+1}, y_{k+1}] \subset [x_k, y_k]$ av lengde $y_{k+1} - x_{k+1} = \frac{d}{2^{k+1}}$ slik at $a_n \in [x_{k+1}, y_{k+1}]$ for alle $n \geq N_{k+1}$. Vi merker at følgen $\{x_n\}$ er voksende og oppad begrenset av M. Ved Setning 7.5.4 vil $\{x_n\}$ konvergere, si mot $L \in \mathbb{R}$. Vi vil vise at a_n også konvergerer mot L. (Det går an å vise at y_n også konvergerer mot L.)

Idéen er som følger: uansett k er $a_n \in [x_k, y_k]$ for stor nok n. Da er $|a_n - x_k| \leq y_k - x_k = \frac{d}{2^k}$, som vi kan gjøre så liten vi vil for stor k. Likeledes kan vi gjøre $|x_k - L|$ så liten vi vil for stor k, siden $x_k \to L$. Ved trekantulikheten har vi $|a_n - L| = |a_n - x_k + x_k - L| \leq |a_n - x_k| + |x_k - L|$, som vi derfor kan gjøre så liten vi vil for stor n. Mer presist kan vi gjøre begge ledd mindre enn $\frac{\epsilon}{2}$: leddet $|a_n - x_k|$ er mindre enn $\frac{d}{2^k}$, som er mindre enn $\frac{\epsilon}{2}$ hvis $\frac{d}{2^{k-1}} < \epsilon$; leddet $|x_k - L|$ er mindre enn $\frac{\epsilon}{2}$ for passe k siden $x_k \to L$. La oss skrive dette skikkelig.

Gitt $\epsilon > 0$. Vi vil vise at det finnes en $N \in \mathbb{Z}^+$ slik at

$$|a_n - L| < \epsilon \quad \text{for alle} \quad n \geq N. \tag{145}$$

Siden $x_k \to L$, vil det finnes en $K_1 \in \mathbb{Z}^+$ slik at $|x_k - L| < \frac{\epsilon}{2}$ for alle $k \geq K_1$. Samtidig vil det finnes en $K_2 \in \mathbb{Z}^+$ slik at $\frac{d}{2^{K_2-1}} < \epsilon$. Velg derfor en $k \geq \max\{K_1, K_2\}$. Da vil

$$|x_k - L| < \frac{\epsilon}{2} \quad \text{og} \quad \frac{d}{2^k} \leq \frac{d}{2^{K_2}} < \frac{\epsilon}{2}. \tag{146}$$

La nå $N = N_k$. For alle $n \geq N$, er derfor $a_n \in [x_k, y_k]$ ved det vi viste ovenfor. Dermed er $|a_n - x_k| \leq y_k - x_k = \frac{d}{2^k}$ og vi har

$$\begin{aligned} |a_n - L| &= |a_n - x_k + x_k - L| \leq |a_n - x_k| + |x_k - L| \\ &\leq \frac{d}{2^k} + |x_k - L| \overset{(146)}{<} \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

□

Cauchy-kriteriet for konvergens er ekstremt hendig å ha, fordi vi med dette kan avgjøre om en følge er konvergent eller ikke uten å kjenne til en kandidat for grensen. La oss bruke kriteriet til å gi et annet bevis på at Cauchy-følgen fra Setning 6.3.4 konvergerer mot $\sqrt{2}$ i $\mathbb{R}$:

EKSEMPEL 7.6.2. Vis at Cauchy-følgen fra Setning 6.3.4, definert ved

$$p_{n+1} = p_n - \frac{p_n^2 - 2}{p_n + 2} = 2 - \frac{2}{p_n + 2}, \tag{147}$$

med en vilkårlig valgt $p_1 \in \mathbb{R}^+$, konvergerer mot $\sqrt{2}$.

BEVIS. Ved Oppgave 6.11 er $\{p_n\}$ er en Cauchy-følge. Ved Setning 7.6.1 vil den derfor konvergere mot en grense $L \in \mathbb{R}$. Ved å ta grenser på begge sider av (147) som i beviset for Setning 6.3.4, finner som før at $L^2 = 2$. Vi har altså de to mulighetene $L = \pm\sqrt{2}$ ved Definisjon 7.3.2 og Merknad 7.3.4.

Ved Oppgave 6.11(a) er $p_n > 0$ for alle n. Siden grenser bevarer orden (Setning 7.4.14), er $L = \lim_{n\to\infty} p_n \geq 0$. Da må $L = \sqrt{2}$. □

7.7. Delfølger og Bolzano-Weierstrass Teoremet*

Vi tar med en liten seksjon med en definisjon og et berømt resultat som vi kommer til å få bruk for i beviset for *algebraens fundamentalteorem* i §8.6 og danner grunnlaget for generalisering av begrepet lukket, begrenset intervall i §10.6. Begge disse seksjonene er merket *, så også denne seksjonen kan hoppes over ved første lesing.

Vi starter med å definere hva vi mener med en *delfølge* av en følge:

EKSEMPEL 7.7.1. Betrakt en følge:

$$x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6,\ x_7,\ x_8,\ x_9,\ x_{10},\ x_{11},\ x_{12},\ x_{13},\ x_{14},\ x_{15},\ x_{16},\ \ldots$$

En delfølge av denne får vi ved å plukke ut uendelig mange av leddene, som for eksempel de som er markert i rødt:

$$x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6,\ x_7,\ x_8,\ x_9,\ x_{10},\ x_{11},\ x_{12},\ x_{13},\ x_{14},\ x_{15},\ x_{16},\ \ldots$$

Da får vi en ny følge

$$x_2,\ x_5,\ x_6,\ x_8,\ x_{12},\ \ x_{14},\ \ x_{15},\ldots. \tag{148}$$

Dette tilsvarer å spesifisere indeksene på de utvalgte leddene, nemlig

$$2,\ \ 5,\ \ 6,\ \ 8,\ \ 12,\ \ 14,\ \ 15,\ldots.$$

La oss si at det $k$te leddet vi plukket ut hadde index n_k. I vårt tilfelle betyr dette

$$n_1 = 2,\ \ n_2 = 5,\ \ n_3 = 6,\ \ n_4 = 8,\ \ n_5 = 12,\ \ n_6 = 14,\ \ n_7 = 15,\ldots.$$

Da er den nye følgen i (148) gitt ved

$$\{x_{n_k}\}_{k=1}^{\infty},$$

det vil si at det k-te leddet i denne følgen er x_{n_k}.

Vi tar dette som utgangspunkt for vår formelle definisjon av delfølge:

Definisjon 7.7.2: Delfølge

La $\{a_n\}$ være en følge i en mengde og $n_1 < n_2 < n_3 < \cdots$ en strengt voksende følge i $\mathbb{Z}^+$. Da kaller vi følgen $\{a_{n_k}\}$ en *delfølge*[a] av $\{a_n\}$. (Merk at $\{a_n\}$ er en delfølge av seg selv, ved å la $n_k = k$ for alle k.)

[a] *Subsequence* på engelsk.

Følgende er nyttig:

Observasjon 7.7.3

En følge $\{a_n\}$ av reelle tall konvergerer mot a hvis og bare hvis enhver delfølge av $\{a_n\}$ konvergerer mot a.

BEVIS
Siden $\{a_n\}$ er en delfølge av seg selv, er "hvis"-retningen ("$\Leftarrow$") umiddelbar. Den motsatte retningen overlates til Oppgave 7.27. □

Det neste resultatet er kanskje mer overraskende:

Setning 7.7.4

Enhver følge av reelle tall har en monoton delfølge.

Bevis
La $\{a_n\}$ være en følge av reelle tall. Vi kaller et ledd a_n for en *topp* dersom alle påfølgende ledd er mindre eller lik a_n, det vil si:

$$a_n \text{ er en topp} \iff a_n \geq a_k \text{ for alle } k > n.$$

Vi deler resten av beviset i to tilfeller:

(i) $\{a_n\}$ har uendelig mange topper;
(ii) $\{a_n\}$ har endelig mange topper.

I tilfellet (i) lar vi

$$a_{n_1}, a_{n_2}, a_{n_3}, \ldots \tag{149}$$

være delfølgen av alle toppene. Per definisjon av topp er

$$a_{n_1} \geq a_{n_2} \geq a_{n_3} \geq \cdots,$$

slik at (149) er en avtagende delfølge av $\{a_n\}$, og setningen er bevist.

I tilfellet (ii) lar vi a_N være den siste toppen, det vil si at $n = N$ er den største indeksen slik at a_n er en topp. La $n_1 = N + 1$. Siden a_{N+1} ikke er en topp, finnes en indeks $n_2 > n_1 = N + 1$ slik at $a_{N+1} = a_{n_1} \leq a_{n_2}$. Siden $n_2 > N$, er a_{n_2} ikke en topp, og dermed finnes det en indeks $n_3 > n_2$ slik at $a_{n_2} \leq a_{n_3}$. Ved å fortsette på denne måten, finner vi en voksende delfølge $\{a_{n_k}\}$ av $\{a_n\}$, og setningen er igjen bevist. □

Vi kan nå presentere hovedresultatet i seksjonen, som er ett av de viktigste resultatene i matematisk analyse. Det bærer navnet etter den bohemske matematikeren, logikeren, filosofen, teologen og katolske presten Bernard Bolzano (1781–1848), som beviste resultatet i 1817, og den tyske matematikeren Karl Weierstrass (1815–1897), som gjenoppdaget resultatet 50 år senere og innså dets fulle dybde.

Teorem 7.7.5: Bolzano-Weierstrass

Enhver begrenset følge av reelle tall har en konvergent delfølge.

Bevis
Ifølge Setning 7.7.4 har enhver slik følge en monoton delfølge. Siden denne er begrenset, er den konvergent ved Setning 7.5.4. □

Eksempel 7.7.6. La $\{x_n\}$ være følgen av reelle tall gitt ved $x_n = (-1)^{n+1}$. Da er $\{x_n\}$ divergent ved Eksempel 7.4.5, og begrenset. Lar vi $n_k = 2k - 1$, får vi delfølgen $\{x_{n_k}\} = \{(-1)^{n_k+1}\} = \{(-1)^{2k}\} = \{1\}$, som konvergerer mot 1. Lar vi $n_k = 2k$, får vi $\{x_{n_k}\} = \{(-1)^{n_k+1}\} = \{(-1)^{2k+1}\} = \{-1\}$, som konvergerer mot -1.

En ekvivalent formulering er:

Teorem 7.7.7: Bolzano-Weierstrass (versjon II)

Enhver følge i et lukket, begrenset intervall $I \subset \mathbb{R}$ har en delfølge som konvergerer i I.

Bevis

La $I = [a, b]$ og $\{x_n\}$ en følge i I. Da er $\{x_n\}$ begrenset og har dermed en konvergent delfølge $\{x_{n_k}\}$ i $\mathbb{R}$ ved Teorem 7.7.5. Siden $a \leq x_{n_k} \leq b$ for alle $k \in \mathbb{Z}$ og grenser bevarer orden (Setning 7.4.14), er $a \leq \lim_{k\to\infty} x_{n_k} \leq b$, som viser at delfølgen konvergererer i I. □

Spesielt interessant er det at dette gir oss en karakterisering av lukkede, begrensede intervaller, noe vi vil bygge videre på i §10.6.

Følgesetning 7.7.8: Bolzano-Weierstrass (versjon III)

Et intervall $I \subset \mathbb{R}$ er lukket og begrenset hvis og bare hvis enhver følge i I har en delfølge som konvergerer i I.

Bevis

"Bare hvis"-retningen ("$\Rightarrow$") følger fra Teorem 7.7.7. "Hvis"-retningen ("$\Leftarrow$") overlates til Oppgave 7.47. □

7.8. Desimaltallsrepresentasjonen av de reelle tallene

Fra skolen er vi vant til å uttrykke reelle tall som desimaltall. I matematikk pleier man ikke å uttrykke reelle tall som desimaltall, med mindre man jobber med tilnærminger, men beholder for eksempel uttrykkene $\sqrt{2}$ og π, istedenfor $1,4142$ og $3,14$. For det første er disse desimaltalluttrykkene ikke eksakt riktige, og det kan føre til store avvik i sluttsvarene. For det andre fører desimaltall gjerne til at uttrykk blir lange og uoversiktlige.

I denne seksjonen skal vi likevel se på hvordan desimaltall passer inn med det vi har lært så langt. Stoffet i denne seksjonen trengs ikke i den videre lesingen, bortsett fra at vi kommer til å ta i bruk desimaltall i beviset for Setning 7.9.10 nedenunder og i §9.8.

Hva mener vi så når vi for eksempel skriver

$$\pi = 3,14159265358979\ldots?$$

Jo, med det mener vi at tallet π er grensen til følgen med ledd

$$\begin{aligned}
&3 \\
3,1 &= 3 + \frac{1}{10} \\
3,14 &= 3 + \frac{1}{10} + \frac{4}{10^2} \\
3,141 &= 3 + \frac{1}{10} + \frac{4}{10^2} + \frac{1}{10^3}
\end{aligned}$$

$$
\begin{array}{rcl}
3,1415 & = & 3+\frac{1}{10}+\frac{4}{10^2}+\frac{1}{10^3}+\frac{5}{10^4} \\
3,14159 & = & 3+\frac{1}{10}+\frac{4}{10^2}+\frac{1}{10^3}+\frac{5}{10^4}+\frac{9}{10^5} \\
3,141592 & = & 3+\frac{1}{10}+\frac{4}{10^2}+\frac{1}{10^3}+\frac{5}{10^4}+\frac{9}{10^5}+\frac{2}{10^6} \\
3,1415926 & = & 3+\frac{1}{10}+\frac{4}{10^2}+\frac{1}{10^3}+\frac{5}{10^4}+\frac{9}{10^5}+\frac{2}{10^6}+\frac{6}{10^7} \\
3,14159265 & = & 3+\frac{1}{10}+\frac{4}{10^2}+\frac{1}{10^3}+\frac{5}{10^4}+\frac{9}{10^5}+\frac{2}{10^6}+\frac{6}{10^7}+\frac{5}{10^8} \\
\cdots & \cdots & \cdots
\end{array}
$$

Prosedyren fortsetter uten stans og uten at vi får noe gjentagende mønster i desimalene, det vil si i siffrene bak komma.

Vi skal videre i denne seksjonen se hvordan vi kan finne slike desimaltallsrepresentasjoner av alle ikke-negative reelle tall og starter med rasjonale tall i første delseksjon. (For negative tall, legger vi bare til et minustegn foran representasjonen av absoluttverdien til tallet.)

MERKNAD 7.8.1. Noen synes det er morsomt å kunne mange desimaler av π. Det finnes mange huskeregler for å memorisere disse. En av de enkleste er å memorisere setningen:

> "*How I want a drink, alcoholic of course, after the heavy chapters involving quantum mechanics!*"

Antallet bokstaver i hvert ord gir de første 15 siffrene i π. Setningen er laget av den engelske fysikeren, astronomen og matematikeren James Hopwood Jeans (1877–1946). Det finnes lengre eksempler: Antallet bokstaver i ordene i rekkefølge i novellen *Cadaeic Cadenza* av Mike Keith gir de første 3834 siffrene i π, mens boken *Not A Wake: A Dream Embodying π's Digits Fully For 10000 Decimals* av Michael Keith gir de første 10000 siffrene.

Forresten er verdensrekorden i memorisering av antall siffre i π for tiden på 70030 ifølge nettsiden

https://www.pi-world-ranking-list.com/?page=lists&category=pi.

Desimaltallsrepresentasjonen av rasjonale tall. La oss først se på hvordan vi finner det vi er vant til å kalle for *desimaltallsrepresentasjonen* eller *desimalutviklingen* av et ikke-negativt rasjonalt tall. Det vi gjør er ikke annet enn å bruke *Divisjonsalgoritmen* (Setning 3.7.3) mange ganger (og stanser hvis vi får null i rest).

La da $\frac{l}{m} \in \mathbb{Q}^{\geq 0}$, med $l \in \mathbb{N}$ og $m \in \mathbb{Z}^+$. Vi bruker *Divisjonsalgoritmen* (Setning 3.7.3) og skriver

$$l = md_0 + r_0, \quad \text{for entydige } d_0, r_0 \in \mathbb{N} \text{ slik at } 0 \leq r_0 < m,$$

eller, som vi kanskje er mest vant til,

$$\frac{l}{m} = d_0 + \frac{r_0}{m}, \quad \text{med} \quad r_0 \in \{0, \ldots, m-1\}. \tag{150}$$

Merk at siden $0 \leq \frac{r_0}{m} < 1$, så er

$$d_0 \leq \frac{l}{m} < d_0 + 1, \tag{151}$$

slik at vi rent "geometrisk" har lokalisert "heltallsintervallet" der $\frac{l}{m}$ ligger, nemlig $[d_0, d_0 + 1)$. Dessuten er $0 \leq \frac{10r_0}{m} < 10$, slik at vi kan gjenta divisjonsalgoritmen på tallet $10r_0$ og få:

$$\frac{10r_0}{m} = d_1 + \frac{r_1}{m}, \quad \text{for entydige } d_1, r_1 \in \mathbb{N} \text{ slik at } 0 \leq r_1 < m,$$

med $d_1 \in \{0, \ldots, 9\}$. (Merk at dersom $r_0 = 0$, da får vi $d_1 = r_1 = 0$.) Ved å dele på 10, får vi $\frac{r_0}{m} = \frac{d_1}{10} + \frac{r_1}{10m}$. Ved å sette dette inn i (150), får vi:

$$\frac{l}{m} = d_0 + \frac{d_1}{10} + \frac{r_1}{10m}, \quad \text{med} \quad r_1 \in \{0, \ldots, m-1\}. \tag{152}$$

Igjen, siden $\frac{r_1}{10m} < \frac{1}{10}$, får vi:

$$d_0 + \frac{d_1}{10} \leq \frac{l}{m} < d_0 + \frac{d_1}{10} + \frac{1}{10}. \tag{153}$$

"Geometrisk" betyr det at vi har delt intervallet $[d_0, d_0 + 1)$ i 10 like store deler og lokalisert hvilket av de ti delintervallene som inneholder $\frac{l}{m}$, nemlig $[d_0 + \frac{d_1}{10}, d_0 + \frac{d_1}{10} + \frac{1}{10})$.

Siden $\frac{10r_1}{m} < 10$, fortsetter vi prosedyren og bruker divisjonsalgoritmen til å skrive

$$\frac{10r_1}{m} = d_2 + \frac{r_2}{m}, \quad \text{for entydige } d_2 \in \{0, \ldots, 9\} \text{ og } r_2 \in \{0, \ldots, m-1\}.$$

(Igjen, dersom $r_1 = 0$, da får vi $d_2 = r_2 = 0$.) Ved å sette inn i (152), får vi:

$$\frac{l}{m} = d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \frac{r_2}{10^2 m}. \tag{154}$$

Igjen, siden $\frac{r_2}{10^2 m} < \frac{1}{10^2}$, får vi:

$$d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} \leq \frac{l}{m} < d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \frac{1}{10^2}. \tag{155}$$

Det er nå klart hvordan prosedyren fortsetter induktivt: når vi har bestemt $d_1, \ldots, d_n \in \{0, \ldots, 9\}$ og $r_1, \ldots, r_n \in \{0, \ldots, m-1\}$ som oppfyller

$$\frac{l}{m} = d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_n}{10^n} + \frac{r_n}{10^n m}, \tag{156}$$

bruker vi divisjonsalgoritmen til å finne r_{n+1} og d_{n+1}:

$$10r_n = md_{n+1} + r_{n+1}, \quad \text{for entydige } d_{n+1} \in \{0, \ldots, 9\}, \ r_{n+1} \in \{0, \ldots, m-1\}.$$

Ved å sette inn i (156), får vi:

$$\frac{l}{m} = d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_{n+1}}{10^{n+1}} + \frac{r_{n+1}}{10^{n+1}m}. \tag{157}$$

Igjen, siden $\frac{r_2}{10^{n+1}m} < \frac{1}{10^{n+1}}$, får vi:

$$d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_{n+1}}{10^{n+1}} \leq \frac{l}{m} < d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \frac{d_{n+1}}{10^{n+1}} + \frac{1}{10^{n+1}}. \tag{158}$$

I prinsippet fortsetter prosedyren i det uendelige, men dersom $r_N = 0$ i ett trinn, da er det lett å se at $r_n = 0$ for alle $n \geq N$ og $d_n = 0$ for alle $n > N$; dessuten har vi da fra (156) at

$$\frac{l}{m} = d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_N}{10^N}.$$

Hvis $r_n > 0$ for alle n, da stanser aldri prosedyren.

Ser vi på dette med språket av følger som vi har utviklet så langt, så ser vi at vi har funnet en følge $\{a_n\}$ av rasjonale tall, der

$$a_n = d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_n}{10^n}, \tag{159}$$

slik at

$$a_n \leq \frac{l}{m} < a_n + \frac{1}{10^n}. \tag{160}$$

Siden $|a_{n'} - a_n| < \left(\frac{1}{10}\right)^n$ for $n' \geq n$, er $\{a_n\}$ en Cauchy-følge, og er dermed konvergent (i $\mathbb{R}$) ved Setning 7.6.1. Bruker vi *Skvisesetningen* (Setning 7.4.13) på (160), ser vi at $\lim_{n\to\infty} a_n = \frac{l}{m}$. Prosedyren gir oss altså en følge $\{a_n\}$ definert ved (159) som konvergerer mot $\frac{l}{m}$.

Det vi kaller *desimaltallsrepresentasjonen*[10] eller *desimalutviklingen*[11] til $\frac{l}{m}$ er uttrykket

$$\frac{l}{m} = d_0, d_1 d_2 d_3 d_4 d_5 \ldots, \tag{161}$$

hvor $d_0 \in \mathbb{N}$ og $d_i \in \{0, \ldots, 9\}$ for $i > 0$. Med denne likheten mener vi at følgen (159) konvergerer mot $\frac{l}{m}$. Sifrene $d_1, d_2, d_3, \ldots$ bak kommaet kaller vi *desimaler*[12]. Dersom $d_n = 0$ for alle $n > N$, skriver vi

$$\frac{l}{m} = d_0, d_1 d_2 d_3 \ldots d_N \tag{162}$$

istedenfor

$$\frac{l}{m} = d_0, d_1 d_2 d_3 \ldots d_N 00000 \ldots$$

[10] *Decimal representation* på engelsk.
[11] *Decimal expansion* på engelsk.
[12] *Digits* på engelsk.

Ved (159) betyr (162) at

$$d_0, d_1 d_2 d_3 \dots d_N = d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \dots + \frac{d_N}{10^N} = a_N. \tag{163}$$

Vi har sett at restene $r_n \in \{0, \dots, m-1\}$. Derfor må vi etter maksimalt m trinn få en gjentagende rest, det vil si at $r_N = r_k$ for et par av positive heltall $N < k$. Da er det lett å se at prosedyren gir ut de samme påfølgende d_n og r_n, det vil si at $d_{N+i} = d_{k+i}$ og $r_{N+i} = r_{k+i}$ for alle $i > 0$. Vi får altså en representasjon på formen

$$d_0, d_1 \dots d_N d_{N+1} \dots d_{k-1} d_N d_{N+1} \dots d_{k-1} d_N d_{N+1} \dots d_{k-1} d_N d_{N+1} \dots d_{k-1} \dots$$

der "mønsteret" $d_N d_{N+1} \dots d_{k-1}$ gjentar seg, som vi gjerne uttrykker som

$$d_0, d_1, \dots d_{N-1} \overline{d_N d_{N+1} \dots d_{k-1}}$$

Vi sier da at representasjonen er *periodisk* med *periode* $k - N$ (antallet sifre i det gjentagende mønsteret).

Eksempel 7.8.2. Et ikke-negativt heltall l får ved denne prosedyren desimaltallsrepresentasjonen $l, 00000 \dots$, det vil si l i den forenklede notasjonen.

Eksempel 7.8.3. La oss se på hva prosedyren ovenfor gir for det rasjonale tallet $\frac{1}{35}$. De forskjellige trinnene gir:

$$\begin{array}{lrcllll}
\text{Trinn } 0: & 1 & = & 35 \cdot 0 + 1 & \Rightarrow & d_0 = 0 & \text{og} & r_0 = 1 \\
\text{Trinn } 1: & 10 \cdot 1 & = & 35 \cdot 0 + 10 & \Rightarrow & d_1 = 0 & \text{og} & r_1 = 10 \\
\text{Trinn } 2: & 10 \cdot 10 & = & 35 \cdot 2 + 30 & \Rightarrow & d_2 = 2 & \text{og} & r_2 = 30 \\
\text{Trinn } 3: & 10 \cdot 30 & = & 35 \cdot 8 + 20 & \Rightarrow & d_3 = 8 & \text{og} & r_3 = 20 \\
\text{Trinn } 4: & 10 \cdot 20 & = & 35 \cdot 5 + 25 & \Rightarrow & d_4 = 5 & \text{og} & r_4 = 25 \\
\text{Trinn } 5: & 10 \cdot 25 & = & 35 \cdot 7 + 5 & \Rightarrow & d_5 = 7 & \text{og} & r_5 = 5 \\
\text{Trinn } 6: & 10 \cdot 5 & = & 35 \cdot 1 + 15 & \Rightarrow & d_6 = 1 & \text{og} & r_6 = 15 \\
\text{Trinn } 7: & 10 \cdot 15 & = & 35 \cdot 4 + 10 & \Rightarrow & d_7 = 4 & \text{og} & r_7 = 10 = r_1 \\
\text{Trinn } 8: & 10 \cdot 10 & = & 35 \cdot 2 + 30 & \Rightarrow & d_8 = 2 = d_2 & \text{og} & r_8 = 30 = r_2 \\
\text{Trinn } 9: & 10 \cdot 30 & = & 35 \cdot 8 + 20 & \Rightarrow & d_9 = 8 = d_3 & \text{og} & r_9 = 20 = r_3 \\
\dots
\end{array}$$

Vi ser at vi får desimaltallsrepresentasjonen

$$\frac{1}{35} = 0,028571428571428571 4 \dots = 0,0\overline{285714},$$

som er periodisk med periode 6. Periodisiteten viser seg først i andre desimal.

Desimaltall. Motivert av prosedyren ovenfor, lager vi følgende definisjon:

Definisjon 7.8.4: Desimaltall

Et *desimaltall*[a] er et uttrykk på formen

$$d_0, d_1 d_2 d_3 \ldots$$

der $d_0 \in \mathbb{N}$ og $\{d_i\}$ er en følge av heltall med $d_i \in \{0, \ldots, 9\}$. Leddet d_i i følgen kalles den *nte desimalen*[b] til desimaltallet.

Desimaltallet kalles *endelig*[c] dersom det finnes en $N \in \mathbb{Z}^+$ slik at $d_n = 0$ for alle $n \geq N$. Vi skriver et slikt tall som

$$d_0, d_1 \ldots d_N.$$

Desimaltallet kalles *periodisk*[d] dersom det er på formen

$$d_0, d_1 \ldots d_N d_{N+1} \ldots d_k d_N d_{N+1} \ldots d_k d_N d_{N+1} \ldots d_k d_N d_{N+1} \ldots d_k \ldots$$

for et eller annet par av positive heltall $N < k$; vi uttrykker dette som

$$d_0, d_1, \ldots d_{N-1} \overline{d_N d_{N+1} \ldots d_{k-1}}.$$

[a] *Decimal* på engelsk.
[b] *nth digit* på engelsk. Etymologien til det norske utrykket "desimal" kommer fra latin "tiende"
[c] *Finite* eller *terminating* på engelsk.
[d] *Periodic* eller *repeating* på engelsk.

Merk at et endelig desimaltall er et spesialtilfelle av et periodisk desimaltall.

Hjelpesetning 7.8.5

Til ethvert desimaltall $d_0, d_1 d_2 d_3 \ldots$ kan vi assosiere følgen $\{a_n\}$ gitt ved

$$a_n = d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_n}{10^n}, \tag{164}$$

som er konvergent.

Definisjon 7.8.6

La $r = \lim_{n \to \infty} a_n$. Da sier vi at *desimaltallet* $d_0, d_1 d_2 d_3 \ldots$ *representerer* r eller er *desimaltallsrepresentasjonen* eller *desimalutviklingen til* r, og skriver også

$$r = d_0, d_1 d_2 d_3 \ldots$$

Bevis for Hjelpesetning 7.8.5
For $m \geq n$ er

$$a_m - a_n = \frac{d_m}{10^m} + \cdots + \frac{d_{n+1}}{10^{n+1}} < \left(\frac{1}{10}\right)^n,$$

slik at $\{a_n\}$ er en Cauchy-følge, som derfor er konvergent (ved Setning 7.6.1). □

Hjelpesetning 7.8.5 sier spesielt at vi til ethvert desimaltall $d_0, d_1 d_2 d_3 \dots$ kan assosiere det reelle tallet $r = \lim_{n\to\infty} a_n$. Siden $a_n \geq 0$ for alle n og grenser bevarer orden (Setning 7.4.14), er $r \geq 0$. Med andre ord har vi en naturlig avbildning

$$\begin{aligned} s : \{\text{desimaltall}\} &\longrightarrow \mathbb{R}^{\geq 0} \\ d_0, d_1 d_2 d_3 \dots &\mapsto \lim_{n\to\infty} a_n = \lim_{n\to\infty} \left(d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_n}{10^n} \right). \end{aligned}$$

Det er naturlig å spørre seg:

- Kan alle ikke-negative reelle tall representeres av desimaltall? Med andre ord, er avbildningen s *surjektiv*?
- Vil forskjellige desimaltall alltid representere forskjellige reelle tall? Med andre ord, er avbildningen s *injektiv*?
- Vi har sett, gjennom prosedyren i forrige delseksjon, at alle ikke-negative rasjonale tall kan representeres ved desimaltall, med andre ord at verdimengden til s inneholder $\mathbb{Q}^{\geq 0}$. Kan vi beskrive hvilke desimaltall som representerer rasjonale tall? Med andre ord, kan vi beskrive $s^{-1}(\mathbb{Q}^{\geq 0})$?

Før vi ser på neste resultat, som gir svar på alle spørsmålene, ser vi på et velkjent eksempel som viser at svaret på det andre spørsmålet er negativt:

EKSEMPEL 7.8.7. Vis at $0, \overline{9} = 1$ som reelle tall, med andre ord: desimaltallene $0, \overline{9}$ og 1 representerer samme reelle tall.

LØSNING. Desimaltallet $0, \overline{9}$ representerer det reelle tallet $\lim_{n\to\infty} a_n$ der

$$a_n = \frac{9}{10} + \frac{9}{10^2} + \frac{9}{10^3} + \cdots + \frac{9}{10^n} = \frac{9}{10} \left(1 + \frac{1}{10^2} + \frac{1}{10^3} + \cdots + \frac{1}{10^n} \right).$$

Grensen til uttrykket i parentesen når $n \to \infty$ er velkjent, jf. Oppgave 7.21(c), nemlig $\frac{1}{1-\frac{1}{10}} = \frac{10}{9}$. Dermed er $\lim_{n\to\infty} a_n = \frac{9}{10} \cdot \frac{10}{9} = 1$. □

Setning 7.8.8

(i) Alle ikke-negative reelle tall har en desimaltallsrepresentasjon.
(ii) To forskjellige desimaltall representerer samme ikke-negative reelle tall hvis og bare hvis de er på formen

$$d_0, d_1 \dots d_N \overline{9} \ \text{ og } \ d_0, d_1 \dots d_{N-1}(d_N + 1).$$

(Hvis $N = 0$, betyr dette $d_0, \overline{9}$ og $d_0 + 1$.)
(iii) En desimaltallsrepresentasjon til et ikke-negativt reelt tall er periodisk (inkludert endelig) hvis og bare hvis tallet er rasjonalt.

MERKNAD 7.8.9. Vi kan også snakke om desimaltallrepresentasjonen til et negativt reelt tall. Hvis $r \in \mathbb{R}^+$ representeres ved $d_0, d_1 d_2 d_3 \ldots$, da betyr dette at $r = \lim_{n\to\infty} a_n$, med $a_n = d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_n}{10^n}$. Da vil $-r = \lim_{n\to\infty} (-a_n)$ og vi sier at $-r$ representeres ved $-d_0, d_1 d_2 d_3 \ldots$.

Før vi går løs på beviset for Setning 7.8.8 merker vi at (ii) sier for eksempel at

$$0,\overline{9} = 0,999999999\ldots = 1, \tag{165}$$

som vi allerede så i Eksempel 7.8.7, og

$$4,2975569\overline{9} = 4,297557, \tag{166}$$

som reelle tall og at disse og lignende tilfeller er de eneste der to ulike desimaltall representerer det samme reelle tallet. Utsagnene (i) og (ii) til sammen gir altså at alle ikke-negative reelle tall har en desimaltallsrepresentasjon som er entydig bortsett fra at $d_0, d_1 \ldots d_{N-1}(d_N+1) = d_0, d_1 \ldots d_N\overline{9}$. Definerer vi ekvivalensrelasjonen $\sim$ på mengden av alle desimalltall ved

$$d_0, d_1 \ldots d_{N-1}(d_N + 1) \sim d_0, d_1 \ldots d_N\overline{9},$$

da får vi altså at mengden $\mathbb{R}^{\geq 0}$ er i én-til-én korrespondanse med mengden av ekvivalensklasser $\{\text{desimaltall}\}/\sim$. Dette gir oss faktisk en alternativ måte å definere reelle tall på som i praksis er lik definisjonen til Weierstrass vi nevnte i introduksjonen til kapitlet. Vi kan addere og multiplisere desimaltall ved å utnytte at de representerer følger og benytte reglene i Setning 7.4.8. Dette gir oss de reglene vi har lært på skolen, men vi skal ikke gå nærmere inn på det her (se imidlertid Hjelpesetning 7.8.11 nedenunder, som vi vil trenge i beviset på del (iii) av Setning 7.8.8).

Vi vil nå gå i gang med beviset for Setning 7.8.8, som vi beviser del for del. Den viktigste og mest intuitive delen er (i): Vi kan rett og slett utvide prosedyren fra forrige delseksjon der vi fant desimalutviklinger til ikke-negative rasjonale tall til å finne desimalutviklinger til alle ikke-negative reelle tall. Vi kan da selvsagt ikke bruke divisjonsalgoritmen, men deler intervaller suksessivt i 10 like store deler for å lokalisere hvor tallet ligger, som er det vi i praksis gjorde for rasjonale tall.

Bevisene for del (ii) og (iii) er nokså tekniske og for spesielt interesserte. Man taper ikke noe for den videre lesing om man hopper over dem.

BEVIS FOR SETNING 7.8.8(I)

La $r \in \mathbb{R}^{\geq 0}$. Vi lar d_0 være det entydige naturlige tallet slik at

$$d_0 \leq r \in [d_0, d_0 + 1) < d_0 + 1, \tag{167}$$

som finnes ved Arkimedes' prinsipp og velordningsprinsippet, jf. Oppgave 7.11). Vi deler så dette intervallet $[d_0, d_0 + 1)$ i 10 like store deler og identifiserer delintervallet der r ligger, det vil si vi finner det entydige naturlige tallet d_1 slik at

$$d_0 + \frac{d_1}{10} \leq r < d_0 + \frac{d_1}{10} + \frac{1}{10}.$$

Vi har at $d_1 \in \{0, \dots, 9\}$ pga. (167). Vi fortsetter prosedyren induktivt: etter at vi har funnet $d_1, \dots, d_n \in \{0, \dots, 9\}$ slik at

$$d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_n}{10^n} \leq r < d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \frac{d_n}{10^n} + \frac{1}{10^n}, \tag{168}$$

finner vi det entydige naturlige tallet d_{n+1} slik at

$$d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_{n+1}}{10^{n+1}} \leq r < d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \frac{d_{n+1}}{10^{n+1}} + \frac{1}{10^{n+1}}.$$

Igjen gir (168) at $d_{n+1} \in \{0, \dots, 9\}$.

Ved *Skvisesetningen* (7.4.13) får vi at

$$\lim_{n \to \infty} \left(d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_n}{10^n} \right) = r,$$

som viser ved Definisjon 7.8.6 at $r = d_0, d_1 d_2 d_3 \dots$. Dette fullfører beviset. □

Eksempel 7.8.10. Vi kan finne desimaltallsrepresentasjonen til $\sqrt{2}$ så langt vi vil ved kvadrering:

$$\begin{array}{lll}
1 < 2 < 4 = 2^2 \Rightarrow 1 < \sqrt{2} < 2 & \Rightarrow & d_0 = 1 \\
(1,4)^2 < 1,96 < 2 < 2,25 = (1,5)^2 & \Rightarrow & d_1 = 4 \\
(1,41)^2 = 1,9881 < 2 < 2,0164 = (1,42)^2 & \Rightarrow & d_2 = 1 \\
(1,414)^2 = 1,999396 < 2 < 2,0002225 = (1,415)^2 & \Rightarrow & d_3 = 4
\end{array}$$

Altså har vi $\sqrt{2} = 1,414\dots$.

Vi trenger følgende resultat i beviset for del (iii) av Setning 7.8.8:

Hjelpesetning 7.8.11

Vi har

(i) $d_0, d_1 d_2 d_3 \dots = \Big(d_0, d_1 \dots d_n\Big) + \left(0, \underbrace{0 \dots 0}_{n \text{ ganger}} d_{n+1} d_{n+2} \dots \right),$

(ii) $10^n \cdot (d_0, d_1 d_2 d_3 \dots) = d^*, d_{n+1} d_{n+2} d_{n+3} \dots$, med
$$d^* = 10^n d_0 + 10^{n-1} d_1 + \cdots + 10 d_{n-1} + d_n.$$

Bevis

Overlates til Oppgave 7.50. □

Bevis for Setning 7.8.8(iii)

Vi har allerede sett at et rasjonalt tall har en desimaltallsrepresentasjon som er periodisk. (Dette viser "hvis"-delen av "hvis og bare hvis"-utsagnet.)

Motsatt, anta at $r \in \mathbb{R}^{\geq 0}$ har en periodisk desimaltallsrepresentasjon, det vil si

$$r = d_0, d_1 d_2 \dots d_{N-1}\overline{d_N \dots d_p}, \quad \text{for heltall} \ \ N \leq p.$$

Da har vi, ved Hjelpesetning 7.8.11(ii), at

$$10^{N-1} r = (10^{N-1} d_0 + 10^{N-2} d_1 + \dots + 10 d_{N-2} + d_{N-1}), \overline{d_N \dots d_p}$$

og

$$10^p r = (10^p d_0 + 10^{p-1} d_1 + \dots + 10 d_{p-1} + d_p), \overline{d_N \dots d_p}.$$

Ved å subtrahere dem (og bruke Hjelpesetning 7.8.11(i)), får vi:

$$10^p r - 10^{N-1} r = (10^p d_0 + \dots + d_p) - (10^{N-1} d_0 + \dots + d_{N-1}),$$

og dermed

$$r = \frac{(10^p d_0 + \dots + d_p) - (10^{N-1} d_0 + \dots + d_{N-1})}{10^p - 10^{N-1}},$$

som er et rasjonalt tall. □

Eksempel 7.8.12. For å finne det rasjonale tallet som har desimaltallsrepresentasjon $1, 11\overline{12}$ regner vi ut

$$r = \frac{11112 - 111}{10^4 - 10^2} = \frac{11001}{9900} = \frac{3667}{3300}.$$

Vi tar også med følgende:

Setning 7.8.13

Et ikke-negativt rasjonalt tall har en endelig desimaltallsrepresentasjon hvis og bare hvis primfaktorene (jf. Definisjon 3.7.8) i dets nevner i full forkortet form kun er 2 og 5.

Bevis

Overlates til Oppgave 7.51. □

Vi vil nå avslutte beviset for Setning 7.8.8:

Bevis for Setning 7.8.8(ii)

Anta at desimaltallene $d_0, d_1 d_2 d_3 \dots$ og $d'_0, d'_1 d'_2 \dots$ representerer samme reelle tall r. Med dette menes at følgene

$$a_n = d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \dots + \frac{d_n}{10^n}$$

og

$$a'_n = d'_0 + \frac{d'_1}{10} + \frac{d'_2}{10^2} + \dots + \frac{d'_n}{10^n}$$

begge konvergerer mot r. La N være den minste indeksen slik at $d_N \neq d'_N$. Ved symmetri kan vi anta at $d_N < d'_N$, og dermed at $d_N + 1 \leq d'_N$. La oss derfor skrive $d'_N = d_N + 1 + s$ for en $s \in \{0, \dots, 8\}$. Vi har, for $n > N$:

$$\begin{aligned}
a_n &= d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_{N-1}}{10^{N-1}} + \frac{d_N}{10^N} + \frac{d_{N+1}}{10^{N+1}} + \cdots + \frac{d_n}{10^n} \\
&= d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_{N-1}}{10^{N-1}} + \frac{d_N}{10^N} \\
&\qquad + \left(\frac{1}{10^N} - \frac{1}{10^N}\right) + \frac{d_{N+1}}{10^{N+1}} + \cdots + \frac{d_n}{10^n} \\
&= \left(d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_{N-1}}{10^{N-1}}\right) + \frac{d_N + 1}{10^N} \\
&\qquad - \left(\frac{1}{10^N} - \frac{d_{N+1}}{10^{N+1}} - \cdots - \frac{d_n}{10^n}\right) \\
&= \left(d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_{N-1}}{10^{N-1}}\right) + \frac{d'_N - s}{10^N} \\
&\qquad - \frac{1}{10^N}\left(1 - \frac{d_{N+1}}{10} - \cdots - \frac{d_n}{10^{n-N}}\right) \\
&\le \left(d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_{N-1}}{10^{N-1}}\right) + \frac{d'_N - s}{10^N},
\end{aligned}$$

siden $\frac{d_{N+1}}{10} + \cdots + \frac{d_n}{10^{n-N}} \le 1$. Samtidig har vi

$$\begin{aligned}
a'_n &= d'_0 + \frac{d'_1}{10} + \frac{d'_2}{10^2} + \cdots + \frac{d'_N}{10^N} + \frac{d'_{N+1}}{10^{N+1}} \cdots + \frac{d'_n}{10^n} \\
&= \left(d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_{N-1}}{10^N}\right) + \frac{d'_N}{10^N} + \left(\frac{d'_{N+1}}{10^{N+1}} \cdots + \frac{d'_n}{10^n}\right) \\
&\ge \left(d_0 + \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_{N-1}}{10^N}\right) + \frac{d'_N}{10^N}.
\end{aligned}$$

Sammenligner vi nå ulikhetene for a_n og a'_n, ser vi at for at $\lim_{n\to\infty} a_n = \lim_{n\to\infty} a'_n$, må vi (ved Skvisesetningen 7.4.13) ha $s = 0$, det vil si

$$d'_N = d_N + 1, \tag{169}$$

samtidig som at både

$$\lim_{n\to\infty}\left(\frac{d'_{N+1}}{10^{N+1}} + \cdots + \frac{d'_n}{10^n}\right) = 0 \tag{170}$$

og

$$\lim_{n\to\infty}\left(\frac{d_{N+1}}{10} + \cdots + \frac{d_n}{10^{n-N}}\right) = 1. \tag{171}$$

Følgen i (170) er for $n > N$ en voksende følge av ikke-negative reelle tall, slik at grensen er 0 hvis og bare hvis alle $d'_n = 0$ for $n > N$. På grunn av (169) mangler det derfor bare å vise at $d_n = 9$ for $n > N$ for å vise setningen.

Vi studerer dermed følgen i (171) for $n > N$. Anta ved selvmotsigelse at $d_k < 9$ for en $k > N$. Da er, for $n \geq k$:

$$\begin{aligned}
\left(\frac{d_{N+1}}{10} + \cdots + \frac{d_n}{10^{n-N}}\right) &\leq \left(\frac{9}{10} + \cdots + \frac{8}{10^k} + \cdots + \frac{9}{10^{n-N}}\right) \\
&= \left(\frac{9}{10} + \cdots + \frac{9}{10^k} + \cdots + \frac{9}{10^{n-N}}\right) - \frac{1}{10^k} \\
&= \frac{9}{10}\left(1 + \frac{1}{10} + \cdots + \frac{1}{10^{n-N-1}}\right) - \frac{1}{10^k} \\
&\overset{(*)}{=} \frac{9}{10}\left(\frac{\frac{1}{10^{n-N}} - 1}{\frac{1}{10} - 1}\right) - \frac{1}{10^k} \\
&= \left(1 - \frac{1}{10^{n-N}}\right) - \frac{1}{10^k},
\end{aligned}$$

hvor vi i overgangen $(*)$ brukte resultatet fra Oppgave 5.10. Uttrykket i siste parentes går mot 1 når $n \to \infty$; siden grenser bevarer orden ved Setning 7.4.14 er derfor $\lim_{n\to\infty}\left(\frac{d_{N+1}}{10} + \cdots + \frac{d_n}{10^{n-N}}\right) \leq 1 - \frac{1}{10^k}$, som motsier (171).

Vi har derfor oppnådd en selvmotsigelse og dermed bevist setningen.
□

7.9. Tellbarhet og ikke-tellbarhet

Vi skal i denne seksjonen bli kjent med nok en vesentlig forskjell mellom $\mathbb{Q}$ og $\mathbb{R}$. Til tross for at begge mengdene er uendelige, og $\mathbb{Q}$ er *tett* i $\mathbb{R}$, så er $\mathbb{Q}$ og $\mathbb{R}$ uendelige på vidt forskjellige måter. Det er den tyske matematikeren Georg Cantor (1845–1918) som står bak konseptene og resultatene i denne seksjonen. Faktisk er det han som i en artikkel i 1874 introduserte konseptet *bijeksjon* (jf. Definisjon 2.3.4). Mer om stoffet i denne seksjonen vil dere møte i kurset *MAT211–Reell Analyse*.

Vi starter med et berømt paradoks for å illustrere temaet:

Eksempel 7.9.1 (Hilberts hotell). Dette er et paradoks om et hypotetisk hotell som den tyske matematikeren David Hilbert (1862–1943) brukte under en forelesning i 1924 for å illustrere uendelighetsbegrepet. Hotellet har uendelig mange rom og alle er opptatt.

En dag kommer en ny gjest og ber om et rom. Resepsjonisten svarer at dessverre er alle rommene opptatt. Hotelldirektøren er imidlertid en smart fyr: han ber gjesten i rom 1 flytte til rom 2, gjesten i rom 2 flytte til rom 3, osv., slik at hver gjest i rom n flytter til rom $n+1$. Slik blir rom 1 ledig, og den nye gjesten kan innlosjeres.

Dette fortsetter å fungere utmerket i en periode, hvor hotellet får 1, 2, 50, 100 nye gjester hver dag: kommer det k nye gjester, ber man nemlig hver gjest i rom n flytte til rom $n+k$, slik at rommene $1, \ldots, k$ blir ledige og kan huse de k nye gjestene.

Men en dag kommer en turistbuss fra kjeden “Infini-tours” til hotellet med uendelig mange passasjerer. Kan hotellet huse alle disse? Ja, sier hotelldirektøren: han ber gjesten i rom 1 flytte til rom 2, gjesten i rom 2 flytte til rom 4, gjesten i rom 3 flytte til rom 6, osv., det vil si at hver gjest i rom n flytter til rom $2n$. Slik blir oddetallsrommene 1, 3, 5, osv. ledige, og passasjer nr. 1 kan innlosjeres i rom 1, passasjer nr. 2 kan innlosjeres i rom 3, passasjer nr. 3 kan innlosjeres i rom 5, osv., det vil si at passasjer nr. n kan innlosjeres i rom $2n - 1$.

Hotellet blir mer og mer populært og en sommerdag kommer det uendelig mange busser fra “Infini-tours” med uendelig mange passasjerer. Dette er den største utfordringen hotelldirektøren har stått ovenfor så langt. Han kommer på at det finnes uendelig mange primtall og bestemmer seg for å bruke dette til å flytte på gjestene. Han ber gjestene som allerede er innlosjert å flytte til alle rommene som er potenser av 2: gjest i rom n flyttes dermed til rom 2^n. Deretter ber han passasjerene i buss nr. 1 om å flytte inn til rommene som er potenser av neste primtall, som er 3: passasjer nr. n i buss nr. 1 flytter altså inn til rom 3^n. Deretter ber han passasjerene i buss nr. 2 om å flytte inn til rommene som er potenser av neste primtall, som er 5: passasjer nr. n i buss nr. 2 flytter altså inn til rom 5^n. Slik fortsetter fordelingen. I motsetning til de andre tilfellene, har denne prosedyren en ytterligere fordel: hotellet vil fremdeles ha uendelig mange ledige rom etter fordelingen, nemlig alle rommene som ikke er potenser av et primtall, slik som $6, 10, 12, 14, 15, \ldots$.

Vi ser dermed at et hotell med uendelig mange rom har helt andre egenskaper enn hoteller med endelig mange rom!

“Paradokset” kan forklares matematisk ved at det finnes bijeksjoner mellom mengden $\mathbb{Z}^+$, som er mengden av hotellrom, og *ekte* delmengder av denne, selv om de er “mindre”. Forskyvningen av gjest i rom n til rom $n+1$ gir en bijeksjon

$$\begin{array}{ccc} \mathbb{Z}^+ & \longrightarrow & \mathbb{Z}^+ \setminus \{1\} \\ n & \mapsto & n+1, \end{array}$$

som gjør rom 1, som ikke ligger i verdimengden, ledig. Forskyvningen av gjest i rom n til rom $2n$ gir en bijeksjon

$$\begin{array}{ccc} \mathbb{Z}^+ & \longrightarrow & \{\text{partall}\} \\ n & \mapsto & 2n, \end{array}$$

som gjør alle oddetalsrommene, som er komplementet til verdimengden, ledige. Innlosjeringen av alle passasjerene i bussen til oddeltalsrommene beskrives da ved bijeksjonen

$$\begin{array}{ccc} \mathbb{Z}^+ & \longrightarrow & \{\text{oddetall}\} \\ n & \mapsto & 2n-1. \end{array}$$

MERKNAD 7.9.2. David Hilbert er en av de mest innflytelsesrike matematikerne gjennom tidene, med viktige bidrag innenfor algebra, geometri, analyse, tallteori og fysikk. Såkalte *hilbertrom*[13], som er generaliseringer av *euklidske rom* vi vil møte på i §10.1, er oppkalt etter ham og spiller en viktig rolle i både matematisk analyse og kvantemekanikk. Spesielt berømt er *Hilberts 23 problemer* publisert i 1900, en liste over viktige problemer som da var uløst. Mange av disse problemene ble inspirerende for 1900-tallets matematikk og flere er fremdelest uløst.

DAVID HILBERT (KILDE: WIKIPEDIA, PUBLIC DOMAIN)

Vi minner om definisjonen på en *endelig* mengde i Definisjon 2.3.15. Neste trinn er å definere:

Definisjon 7.9.3

En mengde A er *tellbar*[a] dersom det finnes en bijeksjon $\mathbb{Z}^+ \to A$. I dette tilfellet sies A å ha *kardinalitet* $\aleph_0$ (leses: "alef-null").

En mengde A er *ikke-tellbar* eller *overtellbar*[b] dersom den verken er endelig eller tellbar.

[a] *Countable* på engelsk.
[b] *Uncountable* på engelsk.

Tegnet $\aleph$ er bokstaven "alef" i det hebraiske alfabetet.

MERKNAD 7.9.4. Noe litteratur definerer en mengde A til å være *tellbar* dersom det finnes en bijeksjon $\mathbb{Z}^+ \to A$ *eller* A er endelig. Begrepet *tellbar* i vår forstand kalles da gjerne *uendelig tellbar*.

[13] *Hilbert spaces* på engelsk.

MERKNAD 7.9.5. Alle de uendelige mengdene av hotellrom, gjester og busser i Eksempel 7.9.1 (Hilberts hotell) er tellbare, noe vi strengt tatt burde ha spesifisert. Hvis ikke, fungerer ikke eksemplet.

Sammenligner vi definisjonen av tellbar med definisjonen av en følge, ser vi at en mengde A er tellbar hvis og bare hvis den er verdimengden til en følge med forskjellige ledd. Derfor har vi:

Observasjon 7.9.6

En mengde er tellbar hvis og bare hvis dens elementer kan listes opp i en følge hvor leddene er forskjellige.

EKSEMPEL 7.9.7. Mengden $\mathbb{Z}$ er tellbar, siden det finnes en bijeksjon mellom $\mathbb{Z}$ og $\mathbb{Z}^+$ som kan visualiseres på følgende måte:

$$\begin{array}{rccccccccccc} \mathbb{Z}: & 0 & 1 & -1 & 2 & -2 & 3 & -3 & 4 & -4 & \dots \\ \mathbb{Z}^+: & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & \dots \end{array}$$

Det overlates til Oppgave 7.56 å finne uttrykket for bijeksjonen.

Det neste resultatet er kanskje litt overraskende:

Setning 7.9.8: Tellbarhet av rasjonale tall

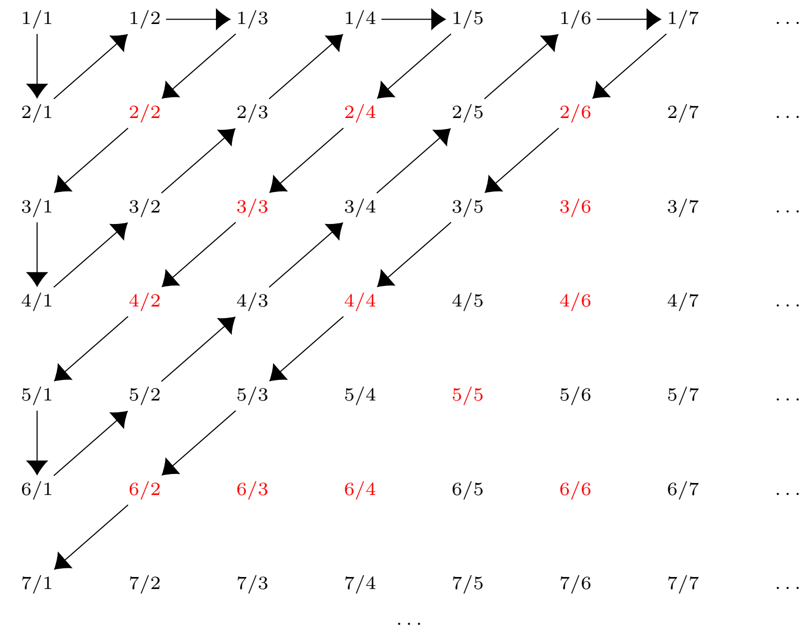
Mengden $\mathbb{Q}$ er tellbar.

BEVIS

Vi viser først at mengden $\mathbb{Q}^+$ av alle positive rasjonale tall er tellbar. Vi lister alle positive rasjonale tall i rader med samme teller:

$$\begin{array}{cccccccccc} 1/1 & 1/2 & 1/3 & 1/4 & 1/5 & 1/6 & 1/7 & 1/8 & 1/9 & \dots \\ 2/1 & 2/2 & 2/3 & 2/4 & 2/5 & 2/6 & 2/7 & 2/8 & 2/9 & \dots \\ 3/1 & 3/2 & 3/3 & 3/4 & 3/5 & 3/6 & 3/7 & 3/8 & 3/9 & \dots \\ 4/1 & 4/2 & 4/3 & 4/4 & 4/5 & 4/6 & 4/7 & 4/8 & 4/9 & \dots \\ 5/1 & 5/2 & 5/3 & 5/4 & 5/5 & 5/6 & 5/7 & 5/8 & 5/9 & \dots \\ 6/1 & 6/2 & 6/3 & 6/4 & 6/5 & 6/6 & 6/7 & 6/8 & 6/9 & \dots \\ 7/1 & 7/2 & 7/3 & 7/4 & 7/5 & 7/6 & 7/7 & 7/8 & 7/9 & \dots \\ \dots \end{array}$$

Merk at det her finnes gjentagelser: for eksempel er $1/1 = 2/2 = 3/3 = \cdots$. Nå kan vi ordne disse tallene i en følge ved å gå gjennom tabellen som vist i figuren, hvor tallene merket med rødt er gjentagelser, som vi derfor hopper over.



Dette viser at $\mathbb{Q}^+$ er tellbar. Samme argument som i Eksempel 7.9.7 viser at det finnes en bijeksjon mellom $\mathbb{Q}$ og $\mathbb{Q}^+$ (jf. Oppgave 7.56), og dermed også en bijeksjon mellom $\mathbb{Q}$ og $\mathbb{Z}^+$, som viser at $\mathbb{Q}$ er tellbar. □

Teknikken fra det siste beviset kan generaliseres: Hvis $A_1, A_2, A_3, \dots$ er mengder (indeksert med $\mathbb{Z}^+$) definerer vi *unionen* til å være alle x som oppfyller $x \in A_n$ for en $n \in \mathbb{Z}^+$. Denne betegnes med $A_1 \cup A_2 \cup A_3 \cup \cdots$ eller $\cup_{n=1}^{\infty} A_n$. Altså:

$$\cup_{n=1}^{\infty} A_n = A_1 \cup A_2 \cup A_3 \cup \cdots = \{x \mid x \in A_n \text{ for en } n \in \mathbb{Z}^+\}.$$

Dette kalles en *tellbar union*. Hvis hver A_n er tellbar, kalles dette en *tellbar union av tellbare delmengder*. Argumentet ovenfor kan brukes til å vise:

Setning 7.9.9

En tellbar union av tellbare delmengder er tellbar.

Bevis
Detaljene overlates til Oppgave 7.58. □

Imidlertid er $\mathbb{R}$ en ikke-tellbar mengde:

Setning 7.9.10: Overtellbarhet av reelle tall

Mengden $\mathbb{R}$ er ikke-tellbar.

Bevis
Mengden $\mathbb{R}$ er uendelig. Vi argumenterer ved selvmotsigelse og antar at $\mathbb{R}$ er tellbar, det vil si at vi ved Observasjon 7.9.6 kan liste alle reelle tall i en følge $\{x_n\}$. La $I \subset \mathbb{R}$ bestå av alle tall i intervallet $(0, 1)$ hvor alle

desimaler i desimaltallrepresentasjonen er 0 eller 1. Ved å plukke ut de tallene i følgen $\{x_n\}$ som ligger i I, *i rekkefølge*, får vi også listet opp alle tallene i I i en følge:

$$\begin{aligned}
a_1 &= 0, a_{11}a_{12}a_{13}a_{14}a_{15}a_{16}a_{17}a_{18}a_{19}\ldots a_{1n}\ldots \\
a_2 &= 0, a_{21}a_{22}a_{23}a_{24}a_{25}a_{26}a_{27}a_{28}a_{29}\ldots a_{2n}\ldots \\
a_3 &= 0, a_{31}a_{32}a_{33}a_{34}a_{35}a_{36}a_{37}a_{38}a_{39}\ldots a_{3n}\ldots \\
a_4 &= 0, a_{41}a_{42}a_{43}a_{44}a_{45}a_{46}a_{47}a_{48}a_{49}\ldots a_{4n}\ldots \\
&\ldots \\
a_n &= 0, a_{n1}a_{n2}a_{n3}a_{n4}a_{n5}a_{n6}a_{n7}a_{n8}a_{n9}\ldots a_{nn}\ldots
\end{aligned}$$

i deres desimaltallrepresentasjon, der alle a_{ij} er enten 0 eller 1. Vi lager nå et nytt desimaltall $a = 0, c_1c_2c_3c_4\ldots c_n\ldots \in I$ ved å definere

$$c_n = \begin{cases} 1, & \text{hvis } a_{nn} = 0 \\ 0, & \text{hvis } a_{nn} = 1. \end{cases}$$

Da vil den n-te desimalen til a være forskjellig fra den n-te desimalen til a_n, som er a_{nn}, slik at a ikke er i listen ovenfor, altså $a \notin I$, en selvmotsigelse. □

Følgesetning 7.9.11: Overtellbarhet av irrasjonale tall

Mengden $\mathbb{R} \setminus \mathbb{Q}$ av irrasjonale tall er ikke-tellbar.

Bevis

Vi argumenterer ved selvmotsigelse. Anta at $\mathbb{R} \setminus \mathbb{Q}$ er tellbar (den er opplagt ikke endelig). Da kan vi ordne elementene i en følge, det vil si vi kan skrive $\mathbb{R} \setminus \mathbb{Q} = \{x_1, x_2, x_3, \ldots, \}$. Ved Setning 7.9.8 er $\mathbb{Q}$ tellbar, slik at vi også kan skrive $\mathbb{Q} = \{y_1, y_2, y_3, \ldots\}$. Da er

$$\mathbb{R} = \mathbb{Q} \cup (\mathbb{R} \setminus \mathbb{Q}) = \{y_1, x_1, y_2, x_2, y_3, x_3, \ldots, \},$$

som viser at $\mathbb{R}$ er tellbar, men dette motsier Setning 7.9.10. □

Merknad 7.9.12. Siste setning viser, som allerede antydet i Merknad 7.2.3 og i introduksjonen til kapitlet, at det finnes "mange flere" irrasjonale tall enn rasjonale. Likevel, ved tetthetsegenskapen til rasjonale tall (Setning 7.2.2), finnes det alltid uendelig mange rasjonale tall mellom to irrasjonale tall.

Vi avslutter seksjonen med et par sider fra sagaen *The Tempest* fra tegneserien *The League of Extraordinary Gentlemen*[14], hvor våre helter møter skurken "Infinity" og vinner over ham med matematisk argumentasjon!

[14]Skrevet av Alan Moore, illustrert av Kevin O'Neill, og publisert av *Top Shelf Productions* og *Knockabout Comics*


TREMBLE, INSIGNIFICANT MORTALS, FOR YOU TRESPASS UPON MY DOMAIN!
I AM INFINITY, THE BIGGEST NUMBER THERE IS! MERELY TO CONTEMPLATE ME HAS DRIVEN MEN MAD!
THEY CRY, "BUT WHAT ABOUT INFINITY PLUS ONE?" STERNLY, I SAY UNTO THEM, "FOOLS!"
"YOU CANNOT HAVE INFINITY PLUS ONE!"
GREAT JOWLS OF JORGULAX! WE FACE INFINITY, THE LARGEST AMOUNT OF ALL!
TRULY, ZOM CANNOT IMAGINE A MORE NUMEROUS FOE!
HUH! SOUNDS LIKE SOME ULTIMATE QUANTITY NEEDS A GOOD TALKING TO! WHERE'S THE ESCAPE HATCH?
ZOM INSTALLED AN AIRLOCK IN THE FLASH AVENGER'S LEFT EAR.
JIM, DON'T SCRAP WITH INFINITY! REMEMBER WE'VE A DATE LATER!
SORRY, CAROL! THAT SWAGGERING SUM DESERVES A LESSON...
...FROM MY PYTHAGORAS-GRANTED POWER OF MATHEMATICS!


EMERGING FROM THE FLASH AVENGER'S AUDIAL CANAL, CAPTAIN UNIVERSE CONFRONTS HIS OMNI-NUMERICAL OPPONENT...
YOU DARE CHALLENGE THE ONE AND ONLY INFINITY?
DON'T MAKE ME LAUGH! THERE'S INFINITE INFINITIES, CHUM! HAVEN'T YOU HEARD OF CANTOR?
UM...YOU MEAN EDDIE CANTOR?
MATHEMATICIAN GEORGE CANTOR, YOU ILLIMITABLE BERK! INVENTOR OF SET THEORY, HE DEMONSTRATED MULTIPLE INFINITIES, SUCH AS THE INFINITE SUBDIVISIONS BETWEEN ANY ADJACENT WHOLE NUMBERS!
MODERN MATHEMATICIANS SUGGEST LOSING INFINITY ALTOGETHER!
B-BUT...WHERE WOULD I GO?
LUCKILY, CANTOR'S DISCIPLE DAVID HILBERT DESIGNED AN INFINITE HOTEL THAT COULD ACCOMMODATE YOU!
I'VE THEIR CARD RIGHT HERE!
WH-WHAT IF ANY OTHER INFINITIES SHOW UP?
THEN GUESTS MOVE TO TWICE THEIR ROOM-NUMBER, LEAVING INFINITE ODD-NUMBERED ROOMS AVAILABLE!
YOU'LL BE FINE!
SOON, BACK INSIDE THE FLASH AVENGER'S HEAD...
WELL DONE, CAPTAIN! THAT ABSTRACT CONCEPT LOOKS UTTERLY CONFOUNDED!
HA! DON'T THANK ME, VULL...THANK PYTHAGORAS!
NOW, LET'S GET THE FLASH AVENGER BACK ON TRACK FOR POSTHUMIA!
≠SIGH!≠ NOW THAT'S WHAT I CALL A SUPER-MARVEL KING OF THE SPACEWAYS!
THE RAKISH DIRIGIBLE CONTINUES THROUGH ECTO-SPACE...
UGGH! THOSE THINGS ARE ZITZING UNBELIEV-ABLE!
AS BRITISH CHILDREN'S ENTERTAINMENT? AYE, THEY'RE CALLED "THE UNBELIEVABLE ONES."
EVENTUALLY...
THE COSMIC AWARENESS OF GALILEO TELLS ME THIS MUST BE POSTHUMIA AHEAD!
JUMPING JETPACKS! LOOK AT THOSE INDESCRIBABLE COLOURS IN ITS

7.10. Algebraiske og transcendentale tall*

Vi har sett at $\sqrt{2}$ er irrasjonalt (Setning 1.3.5), og mer generelt gir Følgesetning 7.9.11 at det finnes overtellbart mange irrasjonale tall. På den annen side har for eksempel $\sqrt{2}$ følgende likhetstrekk med rasjonale tall: på samme måte som et rasjonalt tall q er rot/nullpunkt til et polynom i $\mathbb{Q}[x]$ (det vil si, jf. §4.1, et polynom med rasjonale koeffisienter), nemlig $x - q$, er $\sqrt{2}$ rot i polynomet $x^2 - 2$. Likeledes er enhver n-te rot $\sqrt[n]{q}$ til et rasjonalt tall q rot i et slikt polynom, nemlig $x^n - q$. Slike tall har fått et eget navn:

Definisjon 7.10.1

Et reelt tall kalles *algebraisk* dersom det er en rot til et polynom i $\mathbb{Q}[x]$.

Et reelt tall som ikke er algebraisk kalles *transcendentalt*[a].

[a]Fra latin *transcendere*, som betyr å overstige.

Ved det vi sa ovenfor er alle rasjonale tall algebraiske, men ikke alle algebraiske tall er rasjonale, som $\sqrt{2}$.

MERKNAD 7.10.2. I definisjonen kunne vi like godt krevd at polynomet skal være i $\mathbb{Z}[x]$, siden ethvert polynom i $\mathbb{Q}[x]$ kan omdannes til et polynom i $\mathbb{Z}[x]$ ved å gange med et tilstrekkelig stort heltall, uten at dette forandrer nullpunktene.

MERKNAD 7.10.3. Definisjonen av transcendentale tall går trolig tilbake til den sveitsiske matematikeren Leonhard Euler (1707–1783), selv om begrepet ble introdusert av den tyske filosofen og matematikeren Gottfried Wilhelm Leibniz (1646–1716) i en artikkel fra 1682 der han beviste at $\sin x$ ikke kunne uttrykkes som et polynom og kalte slike funksjoner *transcendentale*.

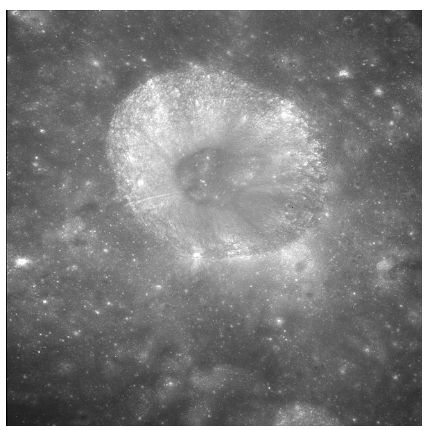
Fra det vi sa rett før definisjonen, får vi følelsen av at det finnes svært mange algebraiske tall (se også Oppgave 7.61). Det er langt i fra opplagt at det finnes transcendentale tall. Faktisk var dette ett av de viktigste uløste problemene i matematikken i første halvdel av 1800-tallet. Selv om grekerne hadde bevist at ikke alle tall er rasjonale, så kunne det likevel tenkes at alle tall er nullpunkter til polynomer med rasjonale koeffisienter. Den sveitsiske matematikeren Johann Heinrich Lambert (1728–1777) la frem en formodning om at både π og e er transcendentale i en artikkel fra 1768, hvor han beviste at π er irrasjonal. (At e er irrasjonalt viser du selv i Oppgave 7.42.)

Det var den franske matematikeren Joseph Liouville (1809-1882) som først i 1844 viste at det måtte finnes transcendentale tall og ga i 1851 det første konkrete eksemplet på et slikt tall, nemlig

$$\sum_{n=1}^{\infty} 10^{-n!} = 10^{-1} + 10^{-1} + 10^{-6} + 10^{-24} + 10^{-120} + 10^{-720} + \cdots$$

som i dag bærer navnet *Liouville-konstanten.* Liouville har også fått oppkalt et krater på månen etter seg.

Joseph Liouville (Kilde: Wikipedia, public domain)



Liouville-krateret på månen (Kilde: Wikipedia, public domain)

Den franske matematikeren Charles Hermite (1822–1901) viste i 1873 at e er transcendentalt, og den tyske matematikeren Carl von Lindemann (1852–1939) viste i 1882 at π er transcendentalt. Alle disse bevisene for at bestemte tall er transcendentale er temmelig kompliserte, så vi skal ikke gå inn på dem her. Med den kunnskapen vi har, er det imidlertid veldig enkelt å vise at det finnes svært mange transcendentale tall. Ideen kommer igjen fra Georg Cantor (1845–1918), som i 1874 viste at algebraiske tall er tellbare:

Setning 7.10.4: Tellbarhet av algebraiske tall

Mengden av algebraiske tall er tellbar.

Bevis
Ideen er overraskende enkel og genial: Skriv

$$\mathbb{Q}[x] \setminus \{0\} = \mathbb{Q}_0[x] \cup \mathbb{Q}_1[x] \cup \mathbb{Q}_2[x] \cup \mathbb{Q}_3[x] \cup \cdots ,$$

hvor $\mathbb{Q}_n[x]$ er mengden av polynomer av grad n. Vis så at hver $\mathbb{Q}_n[x]$ er tellbar ved induksjon; vi overlater dette til Oppgave 7.59(a). Mengden av algebraiske tall $\mathcal{A}$ kan skrives som

$$\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_3 \cup \cdots ,$$

hvor $\mathcal{A}_n$ er mengden av tall som er røtter til polynomer i $\mathbb{Q}_n[x]$, det vil si av rasjonale polynomer av grad n med rasjonale koeffisienter. Siden et slikt polynom kan ha høyst n røtter ved Følgesetning 4.1.10, er hver $\mathcal{A}_n$ tellbar. Ved Setning 7.9.9 er derfor $\mathcal{A}$ tellbar. □

Følgesetning 7.10.5: Overtellbarhet av transcendentale tall

Mengden av transcendentale tall er ikke-tellbar.

Bevis
Dette beviset er i prinsippet likt beviset for Følgesetning 7.9.11, siden vi kan skrive $\mathbb{R} = \mathcal{A} \cup (\mathbb{R} \setminus \mathcal{A})$, hvor $\mathcal{A}$ er mengden av alle algebraiske tall og $\mathbb{R} \setminus \mathcal{A}$ er mengden av alle transcendentale tall. Siden $\mathcal{A}$ er tellbar ved Setning 7.10.4 og $\mathbb{R}$ er ikke-tellbar ved Setning 7.9.10, må også $\mathbb{R} \setminus \mathcal{A}$ være ikke-tellbar. □

Merknad 7.10.6. Ett av *Hilberts 23 problemer* fra år 1900 nevnt i Merknad 7.9.2, problem nummer 7, dreide seg om transcendentale tall: Hvis a er et algebraisk tall og $a \neq 0$, $a \neq 1$, og b er et irrasjonalt algebraisk tall, er da a^b alltid transcendentalt?

Spørsmålet fikk et positivt svar i 1934, gjennom uavhengige arbeider av den sovjetiske matematikeren Alexander Gelfond (1906–1968) og den tyske matematikeren Theodor Schneider (1911–1988).

7.11. Appendiks: Dedekinds konstruksjon av de reelle tallene*

For å være så komplette som mulig (unnskyld ordspillet), vil vi i denne seksjonen gi et raskt overblikk over Dedekinds konstruksjon av de reelle tallene som vi nevnte i introduksjonen til kapitlet. Seksjonen trengs ikke for videre lesing.

I denne konstruksjonen representeres de reelle tallene med visse delmengder av $\mathbb{Q}$, kalt *dedekindske snitt*, som vi nå definerer:

Definisjon 7.11.1

Et *dedekindsk snitt*[a] er en delmengde $A \subset \mathbb{Q}$ som oppfyller:

(I) $A \neq \emptyset$ og $A \neq \mathbb{Q}$;
(II) Hvis $x \in A$, $y \in \mathbb{Q}$ og $y < x$, da er $y \in A$;
(III) Hvis $x \in A$, da finnes $y \in A$ slik at $x < y$.

[a] *(Dedekind) cut* på engelsk.

Egenskap (II) betegnes gjerne med å si at A er *lukket nedentil*. Egenskap (III) sier at A ikke har et største element.

Vi merker at et dedekindsk snitt A automatisk definerer en annen delmengde av $\mathbb{Q}$, nemlig *komplementet* $B = \mathbb{Q} \setminus A$, som oppfyller

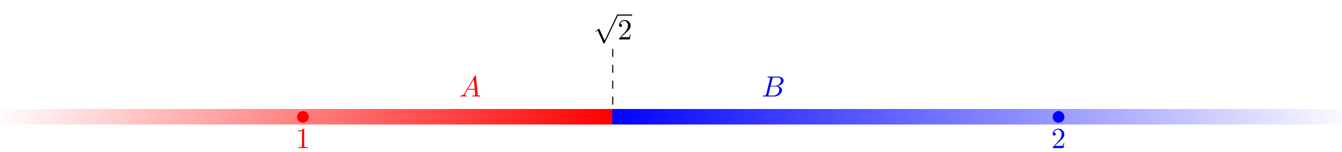
(I)' $B \neq \emptyset$ og $B \neq \mathbb{Q}$;
(II)' Hvis $x \in B$, $y \in \mathbb{Q}$ og $x < y$, da er $y \in B$,

samt

$$A \cup B = \mathbb{Q} \quad \text{og} \quad A \cap B = \emptyset.$$

Vi kan altså også se på et dedekindsk snitt som en partisjon av $\mathbb{Q}$ i to disjunkte delmengder A og B slik at ethvert element i A er mindre enn ethvert element i B, og A ikke har noe største element (mens B kan ha et minste element).

EKSEMPEL 7.11.2. (i) La $A = \{q \in \mathbb{Q} \mid q < 0 \ \text{ eller } \ q^2 < 2\}$. (Dette er lik mengden S vi har betraktet i Eksempel 6.1.4 og i beviset for Setning 6.1.9.) Da er det lett å verifisere at A et dedekindsk snitt og at komplementet til A er $B = \mathbb{Q} \setminus A = \{q \in \mathbb{Q} \mid q > 0 \ \text{ og } \ q^2 \geq 2\}$. Hvis vi aksepterer eksistensen av de reelle tallene, ser vi at $A = (-\infty, \sqrt{2}) \cap \mathbb{Q}$ og $B = (\sqrt{2}, \infty) \cap \mathbb{Q}$. I Dedekinds konstruksjon vil dette dedekindske snittet representere det reelle tallet $\sqrt{2}$, som rent intuitivt er "hullet" mellom mengdene A og B.



MENGDEN A I RØDT; MENGDEN $B = \mathbb{Q} \setminus A$ I BLÅTT.

(ii) La $A = \{q \in \mathbb{Q} \mid q < 1\}$. Da er det lett å verifisere at A et dedekindsk snitt, med komplement $B = \mathbb{Q} \setminus A = \{q \in \mathbb{Q} \mid q \geq 1\}$. I Dedekinds konstruksjon vil dette dedekindske snittet representere det rasjonale (og dermed reelle) tallet 1.

Som mengde definerer vi nå $\mathbb{R}$ som mengden av alle dedekindske snitt. Vi merker at vi har en avbildning

$$\begin{aligned} \mathbb{Q} &\longrightarrow \mathbb{R} \\ q &\longrightarrow \{x \in \mathbb{Q} \mid x < q\}, \end{aligned}$$

og denne avbildningen er opplagt injektiv, siden

$$\{x \in \mathbb{Q} \mid x < q\} \neq \{x \in \mathbb{Q} \mid x < q'\} \Longleftrightarrow q \neq q'.$$

Vi kan derfor identifisere $\mathbb{Q}$ med verdimengden til denne avbildningen og således betrakte $\mathbb{Q}$ som en delmengde av $\mathbb{R}$.

Vi må nå vise at mengden $\mathbb{R}$ oppfyller egenskapene i Teorem 7.1.2.

Først gjør vi $\mathbb{R}$ til en ordnet mengde: vi definerer, for to dedekindske snitt A og B:

$$A < B \Longleftrightarrow \ A \subsetneq B,$$

hvor vi minner om at $A \subsetneq B$ betyr at $A \subset B$ og $A \neq B$, det vil si, A er en *ekte delmengde* av B (jf. side 38 i §2.1). Vi må vise at dette definerer en ordensrelasjon på $\mathbb{R}$, det vil si, at betingelsene (i) og (ii) i Definisjon 3.6.9 er oppfylt: transitivitesegenskapen (ii) er enkel å sjekke, og litt mer arbeid må til for å sjekke trikotomiegenskapen (i), men det er heller ikke så vanskelig. Det er heller ikke så vanskelig å bevise at denne ordningen induserer den opprinnelige ordningen på delmengden $\mathbb{Q}$.

For å vise at $\mathbb{R}$ med denne ordningen har minste-øvre-skranke-egenskapen, la $X \subset \mathbb{R}$ være en oppad begrenset delmengde. Husk at X er en mengde bestående av dedekindske snitt. Det går an å vise at X har en minste øvre skranke gitt ved det dedekindske snittet

$$\{x \in \mathbb{Q} \mid x \in A \ \text{ for en } \ A \in X\},$$

som er med andre ord *unionen* av alle dedekindske snitt i X.

Så må vi definere to operasjoner $+$ og $\cdot$ på $\mathbb{R}$.

Gitt $A, B \in \mathbb{R}$, definerer vi

$$A + B = \{p + q \mid p \in A \ \text{ og } \ q \in B\}.$$

Altså: summen av to dedekindske snitt er mengden av alle rasjonale tall som kan skrives som en sum av ett element i hver av mengdene. Først må vi vise at $A + B$ igjen er et dedekindsk snitt, hvilket ikke er så vanskelig. Deretter må vi vise at kroppsaksiomene (A1)-(A4) er oppfylt: det nøytrale elementet 0^* er det dedekindske snittet $0^* = \mathbb{Q}^-$, mens det additive inverselementet $-A$ til et dedekindsk snitt A er mengden

$$-A = \{p \in \mathbb{Q} \mid \text{det finnes } q \in \mathbb{Q}^+ \text{ slik at } -p - q \notin A\}.$$

Dette tar litt tid å sjekke, men kan gjøres uten store komplikasjoner. Det er dessuten lett å verifisere at for $A, B, C \in \mathbb{R}$ har vi

$$B < C \Longrightarrow A + B < A + C,$$

slik at aksiom (i) i Definisjon 3.6.12 på en ordnet kropp er oppfylt.

Å definere multiplikasjonen $\cdot$ er noe mer komplisert. Vi må gjøre det i to trinn. Siden vi har gjort mengden av alle dedekindske snitt $\mathbb{R}$ til en ordnet mengde og har et nøytralt element 0^* for addisjon, kan vi som vanlig definere $\mathbb{R}^+ = \{A \in \mathbb{R} \mid A > 0^*\}$. Det er ikke vanskelig å se at

$$A > 0^* \Longleftrightarrow \mathbb{Q}^- \subsetneq A \Longleftrightarrow 0 \in A \Longleftrightarrow A \cap \mathbb{Q}^+ \neq \emptyset$$

og at

$$-A < 0^* \iff A > 0^*. \tag{172}$$

Vi definerer først multiplikasjon på $\mathbb{R}^+$: Gitt $A, B \in \mathbb{R}^+$, definerer vi

$$A \cdot B = \{p \in \mathbb{Q} \mid p \leq qr \text{ for en } q \in A \cap \mathbb{Q}^+ \text{ og en } r \in B \cap \mathbb{Q}^+\}.$$

Man kan vise at $A \cdot B$ igjen er et dedekindsk snitt, og ikke minst at $A \cdot B \in \mathbb{R}^+$, hvilket viser aksiom (ii) i Definisjon 3.6.12 på en ordnet kropp. Vi utvider operasjonen $\cdot$ til hele $\mathbb{R}$ ved å utnytte (172) og definere:

$$A \cdot B = \begin{cases} (-A) \cdot (-B), & \text{hvis } A, B < 0^* \\ -((-A) \cdot B), & \text{hvis } A < 0^*, B > 0^* \\ -(A \cdot (-B)), & \text{hvis } A > 0^*, B < 0^* \end{cases}$$

Så må vi vise at kroppsaksiomene (M1)-(M4) og (D) (den distributive lov) er oppfylt med disse definisjonene, hvilket er en god del arbeid. Det multiplikative identitetselementet er

$$1^* = \{q \in \mathbb{Q} \mid q < 1\}$$

og det multiplikative inverselementet til en vilkårlig $A \in \mathbb{R}^+$ er

$$A^{-1} = \left\{ \frac{1}{q} \;\middle|\; q \in \mathbb{Q} \setminus A \right\}.$$

Med dette har vi vist at $\mathbb{R}$ er en ordnet kropp.

Til slutt må vi vise at operasjonene $+$ og $\cdot$ induserer de sedvanlige operasjonene på $\mathbb{Q}$ under identifiseringen av $\mathbb{Q}$ med en delmengde av $\mathbb{R}$.

Oppgaver

Oppgaver til §7.1

Oppgave 7.1. Vis at $\sim$ i Definisjon 7.1.3 er en ekvivalensrelasjon på mengden av alle Cauchy-følger i $\mathbb{Q}$.

Oppgave 7.2. (a) Anta at $\{p_n\}$ er en Cauchy-følge av rasjonale tall slik at $p_n = 0$ for uendelig mange n. Vis at da må $p_n \sim 0$.

(b) Anta at $\{p_n\}$ er en Cauchy-følge av rasjonale tall slik at $p_n \not\sim 0$. Vis at det da finnes en Cauchy-følge $\{q_n\}$ av rasjonale tall slik at $q_n \sim p_n$ og $q_n \neq 0$ for alle n. (Hint: bruk (a).)

Oppgave 7.3. Bevis Hjelpesetning 7.1.5. Du får bruk for Oppgave 7.2(b) ovenfor.

Oppgave 7.4. Hvorfor kunne vi i (106) ikke definert relasjonen som

$$\{p_n\} < \{q_n\} \iff \text{det finnes en } N \in \mathbb{Z}^+ \text{ slik at } p_n < q_n \text{ for alle } n \geq N?$$

Oppgave 7.5. Bevis Hjelpesetning 7.1.7.

Oppgave 7.6. Bevis påstanden om $[C_{y_k}]$ i Lemma 7.1.9.

Oppgave 7.7. La $A \subset \mathbb{R}$ være en ikke-tom nedad begrenset mengde. Definér

$$-A = \{-a \mid a \in A\}.$$

Vis at $\inf -A = -\sup A$. (Dette gir et alternativt bevis for Teorem 6.1.14 for delmengder av de reelle tallene.)

Oppgave 7.8. Ola og Kari diskuterer Teorem 6.1.14. Ola mener å ha funnet et moteksempel. Han sier: "Ta intervallet $(0,1]$. Det oppfyller minste-øvre-skranke-egenskapen, siden $\mathbb{R}$ gjør det, og 1 er med i intervallet. Enhver oppad begrenset delmengde $A \subset (0,1]$ oppfyller derfor at $\sup A \in (0,1]$. Men $(0,1]$ oppfyller ikke største nedre-skranke-egenskapen, siden enhver delmengde $(0,r) \subset (0,1]$, med $r \leq 1$ oppfyller at $\inf(0,r) = 0 \notin (0,1]$."

Kari oppdager en feil i resonnementet. Hvor?

Oppgaver til §7.2–7.3

Oppgave 7.9. La $(a,b) \subset \mathbb{R}$. Vi skal i denne oppgaven fullføre beviset for at (a,b) inneholder et rasjonalt tall, jf. Setning 7.2.2. Vi har allerede funnet en $n \in \mathbb{Z}^+$ slik at $b - a > 1/n$.

(a) Argumentér at det finnes $m_1, m_2 \in \mathbb{Z}^+$ slik at

$$-m_2 < na < m_1.$$

(b) Argumentér at det finnes en $m \in \mathbb{Z}$ slik at

$$m - 1 \leq na < m.$$

(c) Vis at $a < m/n < b$.

Oppgave 7.10. La $(a,b) \subset \mathbb{R}$. Vi skal i denne oppgaven vise at (a,b) inneholder et irrasjonalt tall, jf. Setning 7.2.2. La $q \in (a,b)$ være det rasjonale tallet vi fant i Oppgave 7.9.

(a) Argumentér at det finnes $p \in \mathbb{Z}^+$ slik at $\sqrt{2} < p(b-q)$.

(b) Vis at $q + \frac{\sqrt{2}}{p}$ er irrasjonalt og ligger i (a,b).

Oppgave 7.11. (a) Vis at for enhver $x \in \mathbb{R}$ finnes nøyaktig én $n \in \mathbb{Z}$ slik at $n \leq x < n+1$. (Hint: Bruk Arkimedes' prinsipp og velordningsprinsippet.) Tallet n betegnes med $\lfloor x \rfloor$ (jf. Oppgave 2.18) og kalles *nedrundingen*[15] *av* x.

(b) Sjekk at definisjonen av *opprundingen* $\lceil x \rceil$ gitt for $x \in \mathbb{Q}^{\geq 0}$ som det minste heltallet n slik at $n \geq x$ gir mening for alle $x \in \mathbb{R}$. Begrunn at

$$\lceil x \rceil = \begin{cases} \lfloor x \rfloor, & x \in \mathbb{Z} \\ \lfloor x \rfloor + 1, & x \notin \mathbb{Z}. \end{cases}$$

[15] *Round-down* på engelsk.

Oppgave 7.12. (a) Vis identiteten

$$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1})$$

for alle $n \in \mathbb{Z}^+$ og $a, b \in \mathbb{R}$, for eksempel ved induksjon.

(b) Vis at for $0 < b < a$ gjelder

$$a^n - b^n < (a-b)na^{n-1}$$

Oppgave 7.13. Bruk induksjon til å vise *Bernoullis ulikhet*, oppkalt etter den sveitsiske matematikeren Jacob Bernoulli (1654–1705): For alle $n \in \mathbb{N}$ og reelle tall $x \geq -1$, gjelder

$$(1+x)^n \geq 1 + nx.$$

Oppgave 7.14. Fyll ut detaljene i Merknad 7.3.4.

Oppgave 7.15. Denne oppgaven er ganske vanskelig og fyller ut detaljene i definisjonen av potenser av positive reelle tall med rasjonale og reelle eksponenter i §7.3.

(a) Vis at dersom $m, m' \in \mathbb{Z}$, $n, n' \in \mathbb{Z}^+$ slik at $\frac{m}{n} = \frac{m'}{n'}$, og $x \in \mathbb{R}^+$, da er

$$(x^m)^{\frac{1}{n}} = \left(x^{m'}\right)^{\frac{1}{n'}}.$$

Vi definerer x^r for $x \in \mathbb{R}^+$ og $r \in \mathbb{Q}$ som i Definisjon 7.3.7.

For $x \in \mathbb{R}^+$ og $r \in \mathbb{R}$ betrakter vi delmengden $\{x^t \mid t \in \mathbb{Q}, t \leq r\} \subset \mathbb{R}$.

(b) Vis at $\sup\{x^t \mid t \in \mathbb{Q}, t \leq r\} = x^r$ når $r \in \mathbb{Q}$

Vi definerer x^r for $x \in \mathbb{R}^+$ og $r \in \mathbb{R}$ som i Definisjon 7.3.8.

(c) Bevis Setning 7.3.9. Vis først tilfellene der $r, s \in \mathbb{Q}$. Utvid så til alle $r, s \in \mathbb{R}$ ved å bruke Definisjon 7.3.7.

Oppgaver til §7.4–7.6

Oppgave 7.16. Vis følgende ekvivalente karakterisering av konvergens: *En følge av reelle tall $\{a_n\}$ konvergerer mot a hvis og bare hvis enhver omegn om a inneholder alle bortsett fra endelig mange a_n.*

Oppgave 7.17. (a) Bevis Setning 7.4.14.

(b) Vis at setningen ikke holder hvis vi bytter ut "$\leq$" med "$<$" alle steder.

Oppgave 7.18. La $\{a_n\}$ være en følge av reelle tall.
Vis at $\lim_{n\to\infty} a_n = 0 \iff \lim_{n\to\infty} |a_n| = 0$.

Oppgave 7.19. La $\{a_n\}$ være en følge av reelle tall.
Vis at $\lim_{n\to\infty} a_n = a \iff \lim_{n\to\infty} (a_n - a) = 0$.

Oppgave 7.20. Avgjør om følgene nedenfor konvergerer og finn eventuelt grenseverdien:

(a) $\frac{8n^4+2n^2}{3n^4-8n^3-9}$

(b) $\frac{3n^3+2}{7n^2-9n}$

(c) $\frac{n^2+2\sin n}{2n^2-9n} - \frac{3n^2}{9n^2-8n}$
(d) $\sqrt{n^2+n} - n$
(e) $\frac{1}{\sqrt{n+1}-\sqrt{n}}$
(f) $\frac{1}{\sqrt{n+\sqrt{n}}-\sqrt{n}}$

Oppgave 7.21. La $r \in \mathbb{R}$ slik at $|r| < 1$. Vis at:

(a) $\lim_{n\to\infty} r^n = 0$. (Hint: skriv $r = \frac{1}{1+x}$ for $x > 0$ og bruk Bernouillis ulikhet fra Oppgave 7.13.)
(b) $\lim_{n\to\infty} nr^n = 0$. (Hint: skriv $r = \frac{1}{1+x}$ for $x > 0$ og bruk *binomialteoremet* 3.8.7.)
(c) $\displaystyle\lim_{n\to\infty} \sum_{j=0}^{n} ar^j = \frac{a}{1-r}$, for enhver $a \in \mathbb{R}$. Dette kan også uttrykkes som $\displaystyle\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}$ med notasjonen for rekker fra siste del av §7.4, og denne rekken kalles en *geometrisk rekke*. Resultatet ble først bevist av Cauchy i 1821. (Hint: bruk Oppgave 5.10 og (a).)

Oppgave 7.22 (Akilles og skilpadden). Zenon fra Elea (ca. 490–430) var en gresk filosof som er mest kjent for sine paradokser om tid og bevegelse, blant annet den om *Akilles og skilpadden* som løper om kapp. Begge løper med konstant fart og Akilles løper (mye) fortere enn skilpadden; derfor får skilpadden et forsprang. Ifølge Zenon kan Akilles aldri ta igjen skilpadden, for når Akilles når frem til stedet der skilpadden starter, har skilpadden flyttet seg til et nytt sted, og når Akilles når frem til dette stedet, har skilpadden igjen flyttet seg til et nytt sted. Slik fortsetter argumentet til det uendelige, uten at Akilles klarer å ta igjen skilpadden.

Vis at paradokset ikke holder, ved å la a være skilpaddens forsprang (i meter) og r være forholdet mellom de to hastighetene, og beregne etter hvor mange meter Akilles faktisk tar igjen skilpadden. (Hint: Bruk Oppgave 7.21(c).)

Oppgave 7.23. La $x \in \mathbb{R}$. Finn $\displaystyle\lim_{n\to\infty} \frac{x^n}{n!}$.

Oppgave 7.24. La $r \in \mathbb{R}^+$. Vis at

(a) $\displaystyle\lim_{n\to\infty} \frac{1}{n^r} = 0$;
(b) $\lim_{n\to\infty} \sqrt[n]{r} = 1$.

Oppgave 7.25. La $\{a_n\}$ være en følge av *positive* reelle tall.
Vis at $\displaystyle\lim_{n\to\infty} a_n = \infty \iff \lim_{n\to\infty} \frac{1}{a_n} = 0$.
Hva skjer hvis vi dropper betingelsen at $a_n > 0$?

Oppgave 7.26. La X være en ikke-tom oppad begrenset delmengde av $\mathbb{R}$. La $M \in \mathbb{R}$. Vis at $M = \sup X$ hvis og bare hvis følgende er oppfylt: *$a \leq M$ for alle $a \in A$ og det finnes en følge $a_n \in X$ slik at* $\lim_{n\to\infty} a_n = M$.

OPPGAVE 7.27. Vis Observasjon 7.7.3.

OPPGAVE 7.28. En delmengde $A \subset \mathbb{R}$ kalles *åpen* dersom det for hver $a \in A$ finnes en omegn $(a-\delta, a+\delta) \subset A$, der $\delta > 0$.

En delmengde $A \subset \mathbb{R}$ kalles *lukket* dersom følgende holder: for enhver følge $\{x_n\}$ i A slik at x_n konvergerer mot $x \in \mathbb{R}$, da er $x \in A$.

Vis at A er åpen hvis og bare hvis $\mathbb{R} \setminus A$ er lukket.

(Mengden $\mathbb{R} \setminus A$ kalles gjerne *komplemementet*[16] *til* A.)

OPPGAVE 7.29. Vis at rekken $\displaystyle\sum_{n=1}^{\infty} a_n$ (i $\mathbb{R}$) konvergerer hvis og bare hvis det til enhver $\epsilon > 0$ finnes en $N \in \mathbb{Z}^+$ slik at

$$\left| \sum_{k=n}^{m} a_k \right| < \epsilon$$

for alle $m \geq n \geq N$.

OPPGAVE 7.30. Vis at dersom rekken $\displaystyle\sum_{n=1}^{\infty} a_n$ (i $\mathbb{R}$) konvergerer, da må $\displaystyle\lim_{n\to\infty} a_n = 0$. (Hint: Bruk resultatet fra Oppgave 7.29.)

OPPGAVE 7.31. Vi skal i denne oppgave vise at rekken $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$ divergerer. Denne rekken kalles den *harmoniske rekken*. Siden $\displaystyle\lim_{n\to\infty} \frac{1}{n} = 0$ (Hjelpesetning 6.2.7), viser dette at det motsatte av påstanden i Oppgave 7.30 ikke er sant.

Vi skal gjennomføre det første kjente beviset for divergens av den harmoniske rekken, som går tilbake til den franske matematikeren og filosofen Nicole Oresme (ca. 1322–1382).

La $s_n = \displaystyle\sum_{k=1}^{n} \frac{1}{n}$ være den $n$te delsummen til rekken.

Skriv

$$\begin{aligned}
s_{2^n} = \quad & 1 \\
+ \quad & \frac{1}{2} \\
+ \quad & \frac{1}{3} + \frac{1}{4} \\
+ \quad & \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \\
+ \quad & \frac{1}{9} + \frac{1}{10} + \frac{1}{11} + \cdots + \frac{1}{16} \\
& \cdots\cdots\cdots\cdots\cdots\cdots \\
+ \quad & \frac{1}{2^{n-1}+1} + \frac{1}{2^{n-1}+2} + \frac{1}{2^{n-1}+3} + \cdots + \frac{1}{2^n}.
\end{aligned}$$

[16] *Complement* på engelsk.

Vis at summen av linjen som ender på $\frac{1}{2^k}$ er $\geq \frac{1}{2}$. Bruk dette til å konkludere at rekken $\sum_{n\to\infty} \frac{1}{n}$ divergerer.

Oppgave 7.32. Bevis Setning 7.5.4(ii).

Oppgave 7.33. Avgjør i hvert av tilfellene nedenunder om den rekursivt definerte følgen $\{x_n\}$ av reelle tall konvergerer og finn eventuelt grensen.

(a) $x_1 = 0$ og $x_{n+1} = \frac{x_n}{2} + 1$;
(b) $x_1 = 1$ og $x_{n+1} = \frac{x_n^2+6}{5}$;
(c) $x_1 = 1$ og $x_{n+1} = \sqrt{2x_n + x_n^2}$;
(d) $x_1 = 1$ og $x_{n+1} = \sqrt{2x_n}$.

Oppgave 7.34. La følgen $\{x_n\}$ være gitt ved

$$x_1 = a \text{ hvor } a > 0, \quad x_{n+1} = \frac{x_n^2 + 2}{3}, \quad n \geq 1.$$

(a) Vis at hvis a er slik at $x_2 \geq x_1$, så er følgen voksende, og at hvis a er slik at $x_2 \leq x_1$, så er følgen avtagende.
(b) Bestem for hvilke positive verdier av a følgen er konvergent og finn grenseverdien.

Oppgave 7.35. Etter endt studietid har du fått deg en godt betalt jobb i industrien og har glemt alle idealene fra ungdomstiden. Fabrikken du jobber for, slipper hver kveld ulovlig ut et forurensende avfallstoff i et vassdrag. I løpet av ett døgn forsvinner 2/5 av stoffet fra vassdraget. Samtidig vet man at myndighetene ikke vil legge merke til avfallet så lenge mengden avfall i vassdraget holdes under 20 kg. Sjefen din ønsker å vite hvor mye fabrikken kan slippe ut hver kveld uten at denne grensen overskrides. Hva er svaret? (Hint: du kan få bruk for grensen fra Oppgave 7.21(c).)

Oppgave 7.36. (a) Gitt den reelle tallfølgen $\{a_n\}$ definert ved

$$a_1 = a > 0 \text{ og } a_{n+1} = \frac{1}{5}a_n + a \text{ for alle } n \in \mathbb{Z}^+.$$

Vis at

$$a_n = a + \frac{1}{5}a + \left(\frac{1}{5}\right)^2 + \cdots + \left(\frac{1}{5}\right)^n a$$

for alle $n \in \mathbb{Z}^+$ og bruk dette til å finne et uttrykk for a_n og grensen til $\{a_n\}$. (Hint: du kan få bruk for grensen fra Oppgave 7.21(c).)

(b) Gitt den reelle tallfølgen $\{b_n\}$ definert ved

$$b_1 = 0 \text{ og } b_{n+1} = \frac{1}{10}b_n + \frac{3}{5}a_n \text{ for alle } n \in \mathbb{Z}^+.$$

Vis at $\{b_n\}$ konvergerer.

(c) En høyskole planlegger et nytt, to-årig studium. Hvert år ønsker man å ta opp x nye studenter, og ett år senere venter man at 60% av studentene fra første klasse vil fortsette på det andre året, 20% vil ta det første året om igjen, mens 20% vil ha sluttet. Man regner også med at 10% i andre klasse vil ta det andre året på nytt.

Hvis man ønsker at det totale studenttallet skal stabilisere seg rundt 250, hvor mange nye studenter bør man da ta opp i året?

Oppgave 7.37. I denne oppgaven lar vi x_n være antallet dyr i en bestand målt ved måned n, der $n \in \mathbb{Z}^+$.

Den såkalte *Beverton-Holt rekrutteringsmodellen* forutsier at

$$x_{n+1} = \frac{Rx_n}{1 + \frac{R-1}{K} x_n},$$

der R og K er positive konstanter som kalles *vekstkonstanten* og *bærekapasiteten* til bestanden.

(a) Vis at følgen $\{x_n\}$ er voksende dersom $x_1 \leq K$ og avtagende dersom $x_1 \geq K$.
(b) Hva skjer med bestanden etter hvert som tiden går i henhold til modellen?

Oppgave 7.38. La X være en ikke-tom oppad begrenset delmengde av $\mathbb{R}$ som ikke har noe største element. Vis at det finnes en strengt voksende følge $\{x_n\}$ i X som konvergerer mot $\sup X$.

Oppgave 7.39. Vis at følgen i (104) generert av Herons metode konvergerer mot $\sqrt{K}$, ved å bruke

(a) resultatet fra Oppgave 6.12 og Setning 7.6.1;
(b) Setning 7.5.4.

Oppgave 7.40. Hva er tallet

$$1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cdots}}}}}},$$

dersom det finnes?

Oppgave 7.41. Vi skal i deloppgavene (a)-(b) i denne oppgaven vise at følgen $\left\{\left(1 + \frac{1}{n}\right)^n\right\}$ er voksende og oppad begrenset, som hevdet rett før definisjonen av eulertallet e i Definisjon 7.5.7.

(a) Vis at $\left(1 + \frac{1}{m}\right)^m < \left(1 + \frac{1}{n}\right)^n$ for $m, n \in \mathbb{Z}^+$, $m < n$. Spesielt er $\left(1 + \frac{1}{n}\right)^n < \left(1 + \frac{1}{n+1}\right)^{n+1}$ for alle $n \in \mathbb{Z}^+$.
(b) Vis at $\left(1 + \frac{1}{n}\right)^n < 3$ for alle $n \in \mathbb{Z}^+$. (Hint: bruk binomialteoremet og Oppgavene 3.27(b) og 5.10.)
(c) Vis at $\left(1 - \frac{1}{n}\right)^n = \frac{1}{e}$. Hint: bruk at

$$\left(1 - \frac{1}{n}\right)^n = 1 \Big/ \left(1 + \frac{1}{n-1}\right)^n = 1 \Big/ \left(1 + \frac{1}{n-1}\right)^{n-1} \left(1 + \frac{1}{n-1}\right).$$

Oppgave 7.42. Vi skal i denne oppgaven vise grensen

$$e = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{1}{k!} \tag{173}$$

(eller $e = \sum_{n=0}^{\infty} \frac{1}{n!}$ med notasjonen for rekker fra siste del av §7.4) allerede nevnt i Merknad 7.5.8, samt at e er irrasjonal.

(a) Vis at følgen $\{\sum_{k=0}^{n} \frac{1}{k!}\}$ (eller, om man foretrekker, *rekken* $\sum_{n=0}^{\infty} \frac{1}{n!}$) konvergerer.

La $a_n = \left(1+\frac{1}{n}\right)^n$ og $s_n = \sum_{k=0}^{n} \frac{1}{k!}$ i resten av oppgaven.

(b) Vis at $a_n = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} \leq s_n$ for alle $n \in \mathbb{Z}^+$.
(c) Vis at for $m > n$ er

$$a_m \geq 1 + 1 + \left(1 - \frac{1}{m}\right)\frac{1}{2!} + \cdots + \left(1 - \frac{1}{m}\right)\left(1 - \frac{2}{m}\right)\cdots\left(1 - \frac{n-1}{m}\right)\frac{1}{n!}.$$

(d) Vis (173).
(e) Vis at for $m > n$ er

$$s_m - s_n < \frac{1}{(n+1)!}\left(1 + \frac{1}{n+1} + \frac{1}{(n+1)^2} + \cdots + \frac{1}{(n+1)^{m-n}}\right).$$

(f) Utled at $0 < e - s_n \leq \frac{1}{n!n}$ for alle $n \in \mathbb{Z}^+$. (Hint: La $m \to \infty$ i (e) og bruk Oppgave 5.10.)
(g) Anta at $e = \frac{m}{n} \in \mathbb{Q}$. Vis at da er både $n!e \in \mathbb{Z}$ og $n!s_n \in \mathbb{Z}$ og utled en selvmotsigelse fra (f). Dette viser at e er irrasjonalt.

Oppgave 7.43. Ville det være nok i Definisjon 7.4.3 av en Cauchy-følge å kreve at det finnes en $N \in \mathbb{Z}^+$ slik at $|a_{n+1} - a_n| < \epsilon$ for alle $n \geq N$? (Hint: betrakt $a_n = \sqrt{n}$.)

Oppgaver til §7.7

Oppgave 7.44. Vis følgende ekvivalente karakterisering av konvergens av en delfølge: *En følge av reelle tall $\{a_n\}$ har en delfølge som konvergerer mot a hvis og bare hvis enhver omegn om a inneholder uendelig mange ledd i følgen $\{a_n\}$.*

(Sammenlign med resultatet fra Oppgave 7.16.)

Oppgave 7.45. Vi skal i denne oppgaven gi et alternativt bevis for *Bolzano-Weierstrass teoremet* 7.7.5

(a) La $\{a_n\}$ være en begrenset følge av reelle tall, og start som i beviset for Setning 7.6.1 med et intervall $[-M, M]$ av lengde $d = 2M$ slik at $a_n \in [-M, M]$ for alle $n \in \mathbb{Z}^+$. Argumentér at én av de to halvdelene $[-M, 0]$ og $[0, M]$ av intervallet må inneholde uendelig mange av leddene a_n. Velg dette intervallet og kall det $[x_1, y_1]$.

Fortsett prosedyren med å dele intervallet i to og konstruér dermed en kjede av intervaller

$$[-M, M] \supset [x_1, y_1] \supset [x_2, y_2] \supset [x_3, y_3] \supset \cdots$$

(som i beviset for Setning 7.6.1) slik at hvert intervall har lengde $y_k - x_k = \frac{d}{2^k}$ og inneholder uendelig mange av leddene a_n.

(b) Konstruér en delfølge $\{z_n\}$ av $\{a_n\}$ ved å velge $z_n \in [x_n, y_n]$ for alle $n \in \mathbb{Z}^+$.

(c) Vis at $\{z_n\}$ er Cauchy (jf. beviset for Setning 7.6.1) og fullfør beviset.

Oppgave 7.46. La $\{a_n\}$ være en Cauchy-følge av reelle tall. Vis at dersom $\{a_n\}$ har en delfølge som konvergerer mot a, da konvergerer også $\{a_n\}$ selv mot a.

Oppgave 7.47. Bevis "hvis"-retningen ("$\Leftarrow$") av Følgesetning 7.7.8. (Hint: utfør et kontrapositivt bevis; hvis I er ubegrenset eller hvis I er begrenset, men ikke inneholder begge endepunktene, finn en følge i I som ikke konvergerer i I. Bruk så Observasjon 7.7.3.)

Oppgave 7.48. Gi et alternativt bevis for "hvis"-delen av Setning 7.6.1 ved å kombinere Setning 7.4.12, *Bolzano-Weierstrass teoremet* 7.7.5 og resultatet fra Oppgave 7.46.

Oppgave 7.49. Vis at følgen $\{\sin n\}$ har en delfølge som konvergerer mot 0. (Faktisk har den delfølger som konvergerer mot en vilkårlig $r \in [-1, 1]$.)

Oppgaver til §7.8

Oppgave 7.50. Bevis Hjelpesetning 7.8.11.

Oppgave 7.51. Bevis Setning 7.8.13.

Oppgave 7.52. Finn de 5 første desimalene i desimaltallsrepresentasjonen av $\sqrt{3}$ ved å kvadrere (jf. Eksempel 7.8.10).

Oppgave 7.53. Finn desimaltallsrepresentasjonen til $\frac{6}{13}$.

Oppgave 7.54. Finn det rasjonale tallet som har desimaltallsrepresentasjon $1,2\overline{377}$ (som fullt forkortet brøk).

Oppgave 7.55. La r være det reelle tallet som representeres ved desimaltallene $d_0, d_1 \ldots d_{N-1} d_N$ og $d_0, d_1 \ldots d_{N-1}\overline{9}$ (jf. Setning 7.8.8(ii)). Hvilken av desimaltallsrepresentasjonene for r får vi ved å kjøre algoritmen i beviset for Setning 7.8.8(i)?

(Hint: vis at $r = a_N + \frac{1}{10^N}$, med samme notasjon som i beviset for Setning 7.8.8(ii).)

Oppgaver til §7.9–7.10

Oppgave 7.56. Finn det presise uttrykket for bijeksjonen mellom $\mathbb{Z}$ og $\mathbb{Z}^+$ i Eksempel 7.9.7 og mellom $\mathbb{Q}$ og $\mathbb{Q}^+$ i beviset for Setning 7.9.8.

Oppgave 7.57. Er det kartesiske produktet av to tellbare mengder tellbart? Hva med det kartesiske produktet av endelig mange tellbare mengder?

Oppgave 7.58. Bevis Setning 7.9.9. Hint: La $A_1, A_2, A_3, \ldots$ være de tellbare mengdene som er med i unionen og list opp elementene på følgende måte:

$$\begin{array}{llllllllll}
A_1: & a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} & a_{18} & a_{19} & \ldots \\
A_2: & a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} & a_{28} & a_{29} & \ldots \\
A_3: & a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} & a_{37} & a_{38} & a_{39} & \ldots \\
A_4: & a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & a_{47} & a_{48} & a_{49} & \ldots \\
A_5: & a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & a_{57} & a_{58} & a_{59} & \ldots \\
\ldots
\end{array}$$

Oppgave 7.59. (a) Vis at mengden $\mathbb{Q}_n[x]$ av polynomer av grad n med rasjonale koeffisienter er tellbar. (Hint: bruk induksjon)

(b) Vis at mengden $\mathbb{Q}[x]$ av polynomer med rasjonale koeffisienter er tellbar.

Oppgave 7.60. Vis at en uendelig delmengde av en tellbar mengde er tellbar.

Oppgave 7.61. Vis at hvis $q \in \mathbb{Q}^+$ og $r \in \mathbb{Q}$, da er q^r algebraisk.

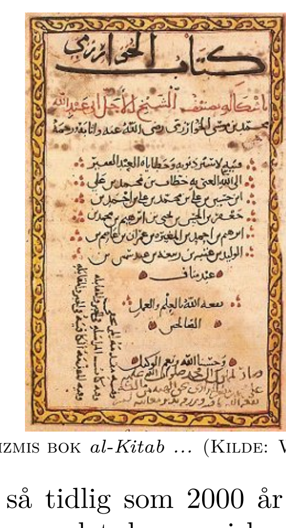
KAPITTEL 8

De komplekse tallene

Det er instruktivt å se tilbake på de tallmengdene vi har sett på så langt og hvordan "mangler" i hver av dem har ført til at vi har funnet det naturlig å utvide mengdene:

- $\mathbb{N}$ mangler inverser for operasjonen $+$ (jf. observasjonen rett etter Setning 3.2.12); ved å legge til disse får vi mengden $\mathbb{Z}$. Det gjør for eksempel at vi i $\mathbb{Z}$ kan løse ligningen $x + 1 = 0$, noe vi ikke kan gjøre i $\mathbb{N}$.
- $\mathbb{Z}$ mangler inverser for operasjonen $\cdot$ (jf. observasjonen etter (57)); ved å legge til disse får vi mengden $\mathbb{Q}$. Det gjør for eksempel at vi i $\mathbb{Q}$ kan løse ligningen $2x = 1$, noe vi ikke kan gjøre i $\mathbb{Z}$.
- $\mathbb{Q}$ oppfyller ikke kompletthetsprinsippet (Setning 6.1.9); ved å legge til ekvivalensklasser av Cauchy-følger får vi mengden $\mathbb{R}$. Det gjør for eksempel at vi i $\mathbb{R}$ kan løse ligningen $x^2 = 2$, noe vi ikke kan gjøre i $\mathbb{Q}$.

Hva så med $\mathbb{R}$? Mangler det noe i $\mathbb{R}$? Vi trenger ikke tenke lenge før vi innser at svaret er "Ja". For eksempel har ligningen $x^2 = 1$ ingen løsning. Mer generelt, vet vi fra skolematematikken at annengradsligninger $ax^2 + bx + c = 0$ ikke har løsninger når $b^2 - 4ac < 0$.



En side fra al-Khwarizmis bok *al-Kitab* ... (Kilde: Wikipedia, public domain)

Selv om Babylonerne så tidlig som 2000 år f.Kr. kunne løse 1. og 2. gradsligninger geometrisk, var det den persiske matematikeren Muhammad ibn Musa al-Khwarizmi (ca. 780–ca. 850) som først publiserte den generelle løsningen for annengradsligninger i boken *al-Kitab al-mukhtasar fi hisab al-jabr wa'l-muqabala.*

La oss minne oss om hvordan man utleder løsningsformelen for en annengradsligning:

$$\begin{array}{ccccc} ax^2+bx+c & = & 0 & \cdot\, 4a \\ 4a^2x^2+4abx+4ac & = & 0 & -\, 4ac \\ 4a^2x^2+4abx & = & -4ac & +\, b^2 \\ 4a^2x^2+4abx+b^2 & = & b^2-4ac & \\ (2ax+b)^2 & = & b^2-4ac & \\ 2ax+b & = & \pm\sqrt{b^2-4ac} & \\ x & = & \frac{-b\pm\sqrt{b^2-4ac}}{2a}. & \end{array}$$

Vi merker allerede i linje 5 at den opprinnelige ligningen ikke kan ha noen løsning blant de reelle tallene hvis $b^2-4ac<0$. Hvis vi imidlertid ser på siste linje og tenker oss at det finnes en kvadratrot av -1, så kan vi tenke oss at vi ved å følge de aritmetiske reglene vi er vant til fra $\mathbb{R}$ kunne skrive:

$$\begin{aligned} x &= \frac{-b\pm\sqrt{b^2-4ac}}{2a} = \frac{-b\pm\sqrt{(-1)(4ac-b^2)}}{2a} = \frac{-b\pm\sqrt{-1}\sqrt{4ac-b^2}}{2a} \\ &= -\frac{b}{2a}+\sqrt{-1}\frac{\sqrt{4ac-b^2}}{2a}. \end{aligned}$$

Dette ser unektelig veldig fantasifullt ut. Imidlertid var det nettopp slik komplekse tall så dagens lys, og det var i forbindelse med et annet problem, som vi nå beskriver.

Det tok nemlig lang tid før matematikere var i stand til å utvikle løsningsformler for tredje- og fjerdegradsligninger. Dette ble gjort, med bidrag av litt forskjellig type, av de italienske matematikerne Scipione del Ferro (1465-1526), Niccoló Fontana Tartaglia (1499-1557), Girolamo Cardano (1501-76) og Ludovico Ferrari (1522-65). Løsningsformlene er temmelig kompliserte i forhold til formelen for annengradsligninger, og ble i utgangspunktet utviklet for ligninger av forskjellig type, som kunne ha 1, 2 eller 3 reelle løsninger. Den italienske matematikeren Rafaele Bombelli (1526-1572) innså at løsningsprosedyren kunne fremstilles på en mer enhetlig måte dersom man i mellomregninger "løste" naturlige annengradsligninger som dukket opp ved å utføre "trikset" vi så ovenfor med å "trekke ut" det "oppdiktede" (eller *imaginære*) tallet $\sqrt{-1}$ og få ut en "løsning" på formen $A+B\sqrt{-1}$, med $A,B\in\mathbb{R}$. Så jobbet han videre med dette uttrykket som om det oppfylte "vanlige regneregler". Bombelli viste at i tilfeller med tre reelle løsninger, så ville "tallet" $\sqrt{-1}$ kanselleres, slik at man satt tilbake med de reelle løsningene, som var riktige. Han innførte altså tallet $\sqrt{-1}$ for å lettere hanskes med mellomregninger og viste at "trikset" ledet frem til riktige løsninger. Han publiserte reglene som uttrykkene skulle oppfylle i sin bok *l'Algebra* i 1572. Denne boken var for øvrig det første europeiske verket som viste hvordan man skulle regne med *negative tall*, som inntil da fremdeles var betraktet som ganske mystiske.


L ALGEBRA
PARTE MAGGIORE
DELL'ARIMETICA
DIVISA IN TRE LIBRI
DI RAFAEL BOMBELLI
DA BOLOGNA.
Nouamente poſta in luce.
IN BOLOGNA
Nella ſtamperia di Giouanni Roſſi
M D L X X I I.
Con Licentia delli RR. VV. del Veſc. & Inquiſit.

FRONTEN PÅ BOMBELLIS BOK *l'Algebra* FRA 1572 (KILDE: WIKIPEDIA, PUBLIC DOMAIN)

Hva er så de "vanlige regnereglene" Bombelli brukte? Vi kan selv tenke oss frem til dette, siden reglene skal være de samme når man jobber med reelle tall, det vil si uttrykk på formen $a+b\sqrt{-1}$ med $b=0$. La oss for enkelhets skyld betegne uttrykket $\sqrt{-1}$ med i, som er den vanlige skrivemåten idag. Altså er i et symbol som oppfyller $i^2=-1$.

Den naturlige måten å addere to uttrykk på formen $a+bi$ og $c+di$ er å addere dem som vi adderer (lineære) polynomer i variabelen i (altså elementer i ringen $\mathbb{R}[i]$, jf. §4.1), nemlig å addere alle ledd med og uten i hver for seg:

$$(a+bi)+(c+di)=(a+c)+(b+d)i. \tag{174}$$

Likeledes for differanse:

$$(a+bi)-(c+di)=(a-c)+(b-d)i. \tag{175}$$

Vi bruker samme tankegang når vi multipliserer. Her bruker vi først den distributive loven (som vi bare "later som om" er oppfylt!), alternativt de vanlige regnereglene for multiplikasjon av polynomer i variabelen i. Dessuten bruker vi tilleggsopplysningen om at $i^2=1$:

$$\begin{aligned}(a+bi)(c+di) &= ac+bci+adi+bdi^2=ac+bci+adi-bd \\ &= (ac-bd)+(bc+ad)i.\end{aligned} \tag{176}$$

Vi merker at vi får tilbake vanlig multiplikasjon på $\mathbb{R}$ når $b=d=0$.

Kan vi så dele to uttrykk $a+bi$ og $c+di$ på hverandre? For å ende opp med gangeregler som oppfyller det vi er vant til, må $\frac{c+di}{a+bi}$ være det samme som $(c+di)\cdot(a+bi)^{-1}$, hvor $(a+bi)^{-1}$ er det multiplikative inverselementet til $a+bi$, altså noe som oppfyller $(a+bi)^{-1}\cdot(a+bi)=1$. For å se om et slikt uttrykk finnes, løser vi for de ukjente x og y og bruker regelen i (176):

$$\begin{aligned} (x+yi)\cdot(a+bi) &= 1 \\ xa-yb+(xb+ya)i &= 1 \\ xa-yb=1 \ \text{ og } \ xb+ya=0 \\ x=\frac{a}{a^2+b^2} \ \text{ og } \ y=\frac{-b}{a^2+b^2}. \end{aligned}$$

Altså er

$$(a+bi)^{-1}=\frac{a}{a^2+b^2}+\frac{-b}{a^2+b^2}i=\frac{1}{a^2+b^2}\,(a-bi)\,, \tag{177}$$

slik at

$$\begin{aligned} \frac{c+di}{a+bi} &= (c+di)\cdot(a+bi)^{-1}=\frac{1}{a^2+b^2}\,(c+di)\,(a-bi) \\ &= \frac{1}{a^2+b^2}\,(ac+bd+(ad-bc)i)\,. \end{aligned} \tag{178}$$

Vi har nå utledet "regnereglene" som Bombelli brukte. Det ble etter hvert godtatt i matematiske kretser at man kunne operere med kvadratrøtter av negative tall på denne måten og bruke dem i mellomregninger. Det var imidlertid mange som følte seg ukomfortable med dette, og det bør også vi være på det nåværende tidspunkt. Vi har utledet naturlige regneregler, men kan vi være sikre på at disse ikke inneholder motsigelser? Og kan vi være sikre på at vi i fremtiden ikke vil komme frem til feil svar i noen beregninger, selv om prosedyren har vist seg å fungere bra så langt?

Med språket vi har utviklet hittil, ser vi at det riktige spørsmålet er om elementene på formen $a+bi$ med de regnereglene vi har definert danner en kropp som inneholder $\mathbb{R}$ som underkropp. Svaret er ja, men det tok matematikere flere hundre år etter Bombelli før dette ble stadfestet ordentlig. Et viktig bidrag stod den norsk-danske matematikeren og landmåleren Caspar Wessel (1745–1818) for, da han i et arbeid i 1799 representerte uttrykkene $a+bi$ ved parene $(a,b)\in\mathbb{R}^2$, som han tolket som vektorer i planet. Dessverre fikk hans arbeid ikke den oppmerksomheten det fortjente i første omgang. Den sveitsiske bokholderen og hobbymatematikeren Jean-Robert Argand (1768–1822) publiserte lignende idéer i 1813, men det var først da den tyske matematikeren Johann Carl Friedrich Gauss (1777–1855) publiserte et arbeid i 1831 at matematikere virkelig ble oppmerksomme på denne representasjonen. Ved representasjonen som punkter eller vektorer i planet kan formlene for addisjon, subtraksjon, multiplikasjon og divisjon, og invers i (174)-(178) uttrykkes som:

$$\begin{aligned} (a,b)+(c,d) &= (a+c,b+d) \\ (a,b)-(c,d) &= (a-c,b-d) \\ (a,b)\cdot(c,d) &= (ac-bd,ad+bc) \\ \frac{(c,d)}{(a,b)} &= \left(\frac{ac+bd}{a^2+b^2},\frac{ad-bc}{a^2+b^2}\right) \end{aligned}$$

$$(a,b)^{-1} = \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right).$$

Deretter gikk utviklingen raskt, og et stort bidrag ble gjort av Cauchy (1789–1857), som utviklet læren om komplekse funksjoner.

Det er for øvrig Gauss som introduserte symbolet i for $\sqrt{-1}$ og navnet *komplekse tall* for mengden av alle uttrykk på formen $a + bi$. Den franske matematikeren, filosofen og vitenskapsmannen René Descartes (1596–1650) introduserte navnet *imaginært* for uttrykket $\sqrt{-1}$ i 1637.

Et høydepunkt i teorien om de komplekse tallene er *Fundamentalteoremet i algebra*, et av de fremste teoremene i matematikken, som sier at over de komplekse tallene er det ikke bare ligningen $x^2 + 1 = 0$ som har en løsning, men alle polynomielle ligninger med koeffisienter i $\mathbb{C}$. Det første helt korrekte beviset for teoremet ble gitt av Argand i 1814, etter to bevis av vidt forskjellig natur med mangler gitt av den engelske matematikeren James Wood (1760–1839) i 1798 (som for øvrig ble fullstendig ignorert) og Gauss i 1799.

Vi skal i neste seksjon gi den formelle definisjonen av komplekse tall. Deretter vil vi i resten av seksjonen studere egenskaper ved komplekse tall, deriblant fremstillingen i det komplekse planet i §8.2. *Fundamentalteoremet i algebra* presenteres i §8.4. Beviset vil bli gitt i §8.6, etter å ha oppsummert resultater om følger av komplekse tall i §8.5.

Vi skal ikke gjennomgå noe av det svært interessante stoffet om funksjoner mellom mengder av komplekse tall, selv om vi skal si noe om begrepet kontinuitet også for slike funksjoner i de neste to kapitlene. Dette er grunnlaget for fagfeltet *kompleks analyse* og studeres i kursene *MAT213–Komplekse funksjoner* og *MAT214–Kompleks analyse.*

8.1. Definisjonen av kroppen av de komplekse tall

Vi definerer de komplekse tallene som ordnede par (a, b) av reelle tall, det vil si elementer $(a, b) \in \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$, som oppfyller regnereglene vi fant i introduksjonen til kapitlet. Denne definisjonen kommer fra den irske matematikeren og astronomen William Rowan Hamilton (1805–1865):

Definisjon 8.1.1: Komplekse tall

Mengden av *komplekse tall*[a], betegnet med $\mathbb{C}$, er lik mengden $\mathbb{R}^2$ sammen med de to operasjonene

$$(a,b) + (c,d) = (a+c, b+d) \tag{179}$$

$$(a,b) \cdot (c,d) = (ac-bd, ad+bc). \tag{180}$$

Vi identifiserer $\mathbb{R}$ med delmengden $\mathbb{R} \times \{0\} \subset \mathbb{C}$, det vil si delmengden av punkter på formen $(a, 0)$.

[a] *Complex numbers* på engelsk.

Setning 8.1.2

Operasjonene ovenfor gjør $\mathbb{C}$ til en kropp som inneholder $\mathbb{R}$ som underkropp; nærmere bestemt har vi at

- $(0,0)$ er det nøytrale element;
- $(1,0)$ er den multiplikative identitet;
- $(-a,-b)$ er det additive inverselementet til (a,b);
- $\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right)$ er det multiplikative inverselementet til elementet $(a,b) \neq (0,0)$.

Bevis
Overlates til Oppgave 8.3. □

Merknad 8.1.3. Merk at $\mathbb{C}$ ikke er en ordnet kropp (jf. Oppgave 8.4).

Vi definerer

$$i = (0,1)$$

og kaller i *den imaginære enheten*[1]. Vi skriver dessuten a istedenfor $(a,0)$. Da får vi skrivemåten

$$(a,b) = (a,0) + (0,b) = a + ib,$$

og det er denne vi pleier å bruke for de komplekse tallene. Med denne notasjonen får vi tilbake regnereglene vi fant i introduksjonen til kapitlet.

Vi kaller det reelle tallet a for *realdelen*[2] til det komplekse tallet $z = a+ib$ og betegner det med

$$a = \mathrm{Re}(z).$$

Det reelle tallet b kaller vi for *imaginærdelen*[3] til det komplekse tallet $z = a + ib$ og betegner det med

$$b = \mathrm{Im}(z).$$

Vi betrakter $\mathbb{R}$ som en underkropp av $\mathbb{C}$ gjennom identifiseringen av $a \in \mathbb{R}$ med $a + 0 \cdot i \in \mathbb{C}$. Alle tallene på formen $0 + ib = ib$ kalles *imaginære* eller *rent imaginære*.

Kompleks konjugasjon. Det finnes en viktig regneoperasjon for komplekse tall som ikke har noe motstykke hos de reelle tallene:

Definisjon 8.1.4: Kompleks konjugert

La $z = a+ib \in \mathbb{C}$. Det *(kompleks) konjugerte tallet*[a] til z er $\overline{z} = a - ib$.

[a] *(Complex) conjugate* på engelsk.

[1] *Imaginary unit* på engelsk.
[2] *Real part* på engelsk.
[3] *Imaginary part* på engelsk.

Legg merke til at det kompleks konjugerte tallet til et reelt tall er tallet selv: hvis $z = a = a + i \cdot 0$ er $\overline{z} = a$. Konjugasjon oppfyller følgende naturlige regler:

Setning 8.1.5: Regneregler for konjugasjon

La $z, w \in \mathbb{C}$. Da gjelder:

(i) $\overline{z+w} = \overline{z} + \overline{w}$
(ii) $\overline{z-w} = \overline{z} - \overline{w}$
(iii) $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$
(iv) $\overline{\frac{z}{w}} = \frac{\overline{z}}{\overline{w}}$

Bevis
Overlates til Oppgave 8.5. □

Én av grunnene til at konjugasjon er sentralt er at

$$z\overline{z} = (a+ib)(a-ib) = a^2 + b^2,$$

som er et reelt tall. Vi kan bruke denne egenskapen som huskeregel når vi skal regne ut brøker som i (178), som unektelig er en vanskelig formel å huske: hvis $z = a + ib$ og $w = c + id$, da er

$$\frac{w}{z} = \frac{w\overline{z}}{z\overline{z}} = \frac{w\overline{z}}{a^2+b^2} = \frac{(c+di)(a-bi)}{a^2+b^2} = \frac{1}{a^2+b^2}\left(ac+bd+(ad-bc)i\right).$$

Spesielt er

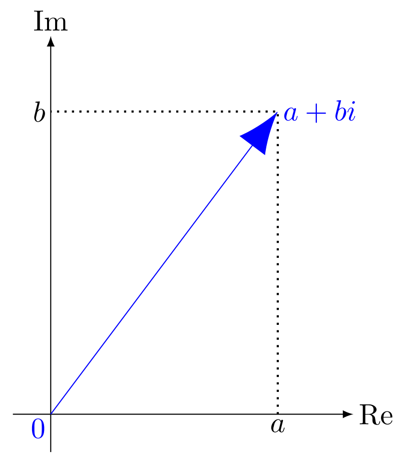
$$z^{-1} = \frac{1}{z} = \frac{\overline{z}}{z\overline{z}} = \frac{\overline{z}}{a^2+b^2}.$$

Eksempel 8.1.6. For å regne ut $\frac{3+6i}{2-3i}$ skriver vi:

$$\frac{3+6i}{2-3i} = \frac{(3+6i)(2+3i)}{(2-3i)(2+3i)} = \frac{6+12i+9i-18}{4+9} = \frac{-12+21i}{13}$$

8.2. Det komplekse planet

Som nevnt i introduksjonen til kapitlet kan vi tolke et komplekst tall $z = a + ib$ som et punkt eller en vektor (a, b) i planet. De komplekse tallene utgjør altså ikke en tallinje, men et "tallplan", som vi kaller *det komplekse planet*. (Noen ganger blir det også kalt et *Argand plan* eller *Argand diagramm*.) Den horisontale aksen kalles *den reelle aksen* og er ikke annet enn den gode gamle tallinjen som representerer alle reelle tall, siden alle punktene her er på formen $(a, 0)$ eller $z = a \in \mathbb{R}$. Den vertikale aksen kalles *den imaginære aksen*.



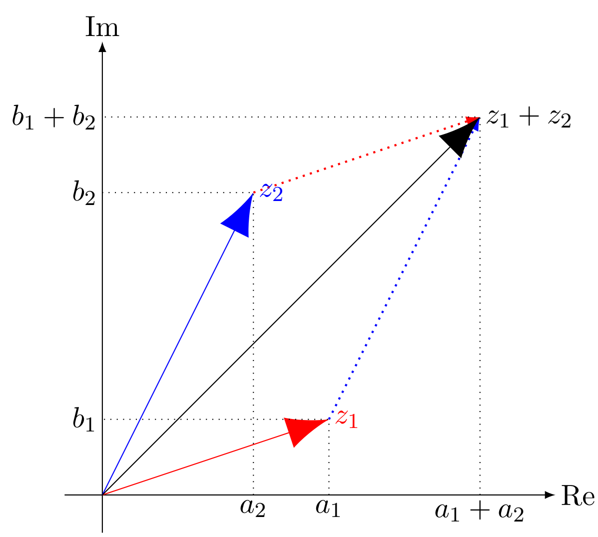
TALLET $z = a + bi$ MARKERT I DET KOMPLEKSE PLANET

Denne geometriske visualiseringen av de komplekse tallene har som fordel at vi kan tolke mye av teorien til de komplekse tallene geometrisk, som vi nå skal forklare.

Addisjon og konjugasjon i det komplekse plan. Vi legger med en gang merke til at operasjonen addisjon

$$z_1 + z_2 = (a_1 + ib_1) + (a_2 + ib_2) = (a_1 + a_2) + i(b_1 + b_2)$$

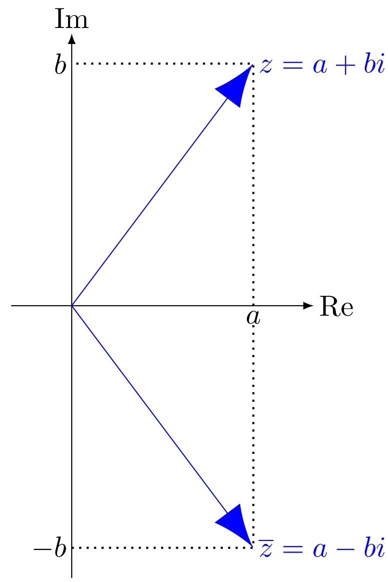
gitt i (179) er det samme som *vektoraddisjon* av vektorene (a_1, b_1) og (a_2, b_2):



ADDISJON AV $z_1 = a_1 + ib_1$ OG $z_2 = a_2 + ib_2$ I DET KOMPLEKSE PLANET

(For de som har glemt vektoraddisjon fra skolen, se Definisjon 10.1.7.)

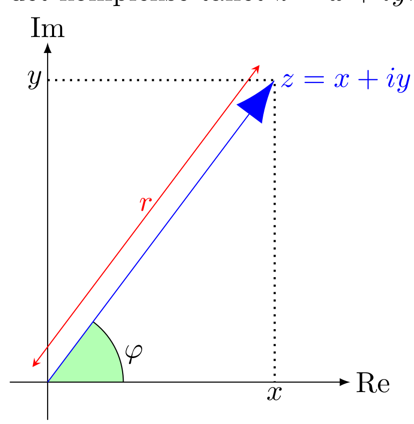
Vi legger også med en gang merke til at kompleks konjugasjon av et tall $z = a + ib$ tilsvarer å speile den tilsvarende vektoren om den reelle aksen:



Kompleks konjugasjon av $z = a + ib$ i det komplekse planet

Vi vil om litt se at også multiplikasjonen gitt i (180) har en fin geometrisk tolkning. Til det må vi imidlertid innføre litt notasjon.

Polarformen til et komplekst tall. Et punkt i planet $\mathbb{R}^2$ forskjellig fra origo kan representeres med de *kartesiske koordinatene* (x, y) til endepunktet, men også ved hjelp av lengden r på vektoren og størrelsen φ på vinkelen den danner med den horisontale aksen, som vi så i Eksempel 2.3.14. Paret (r, φ) kalles *polarkoordinatene* til vektoren eller punktet (x, y). Det samme gjelder da for det komplekse tallet $z = x + iy$:



Modulus r og argument φ til det komplekse tallet $z = x + iy$

Fra figuren ser vi at

$$r = \sqrt{x^2 + y^2}, \quad x = r\cos\varphi, \quad y = r\sin\varphi.$$

slik at vi kan skrive

$$z = x + iy = r\cos\varphi + ir\sin\varphi.$$

Uttrykket til høyre kalles *polarformen* til z, mens uttrykket til venstre kalles *standardformen* til z. Tallet r kalles *modulus* (også kalt *norm, absoluttverdi, tallverdi*) til z og betegnes med $|z|$. Altså

$$|z| = \sqrt{x^2 + y^2} \in \mathbb{R}. \tag{181}$$

Dette er rent geometrisk avstanden mellom tallet z representert i det komplekse planet og origo. Dette tallet spiller altså samme rolle som absoluttverdien til et reelt tall. Spesielt merker vi at

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

(som er egenskapen tilsvarende den i Setning 3.6.16(i)) og at *trekantulikheten* (jf. Setning 3.6.16(ii)) fremdeles holder:

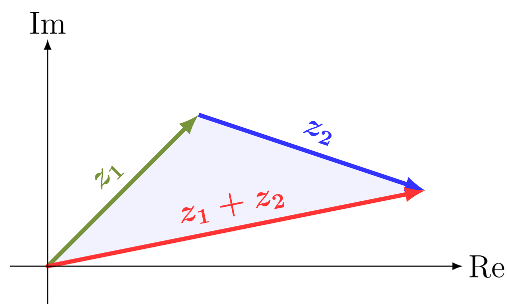
Setning 8.2.1: Trekantulikheten for komplekse tall

Dersom $z_1, z_2 \in \mathbb{C}$, da er

$$|z_1 + z_2| \leq |z_1| + |z_2|.$$

Bevis

Overlates til Oppgave 8.9. Det enkleste er å argumentere geometrisk ved å betrakte figuren



Trekantulikheten for komplekse tall

□

Trekantulikheten kan generaliseres til vilkårlig mange komplekse tall, jf. Oppgave 8.10.

Legg også merke til at

$$|z|^2 = x^2 + y^2 = z\overline{z}, \tag{182}$$

som er en formel som viser seg å være svært nyttig.

Merknad 8.2.2. Vi merker oss også at størrelsen $|z_1 - z_2| \in \mathbb{R}$ ikke er annet enn avstanden mellom punktene som representerer z_1 og z_2 i det komplekse planet. På samme måte som for de reelle tallene har vi altså et avstandsbegrep på mengden $\mathbb{C}$. Mer presist definerer funksjonen

$$\begin{aligned} \mathbb{C} \times \mathbb{C} &\longrightarrow \mathbb{R} \\ (z_1, z_2) &\mapsto |z_1 - z_2| \end{aligned}$$

en metrikk på $\mathbb{C}$, og gjør derfor $\mathbb{C}$ til et metrisk rom (Definisjon 7.2.4). Vi overlater verifiseringen av dette til Oppgave 8.13.

Vi merker oss også at tallet $|z|$ er entydig, mens vinkelen φ kan erstattes med en hvilken som helst vinkel på formen $\varphi + 2\pi n$ for $n \in \mathbb{Z}$. En hvilken som helst av disse vinklene kalles et *argument* for z, og betegnes med $\arg z$.

Polarformen til z kan altså skrives som

$$z = |z|\cos\varphi + i|z|\sin\varphi = |z|\left(\cos\varphi + i\sin\varphi\right), \quad \varphi \text{ et argument for } z. \tag{183}$$

Vi merker også at origo har modulus 0, mens alle φ er et argument for origo.

Vi oppsummerer:

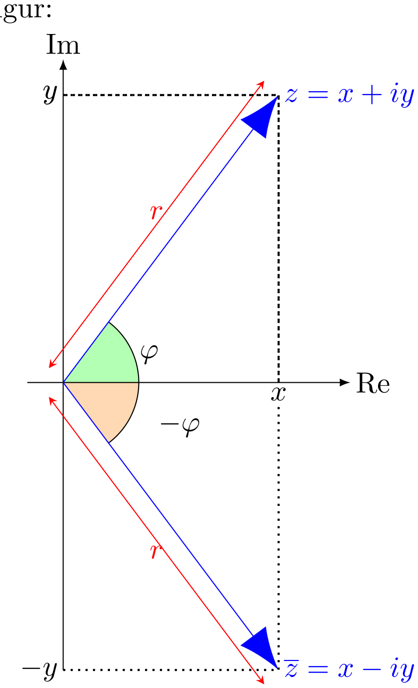
Observasjon 8.2.3

To komplekse tall forskjellig fra null er like hvis og bare hvis modulusene er like og differansen av argumentene er $2\pi n$ for en $n \in \mathbb{Z}$.

Legg ellers merke til at

$$|\overline{z}| = |z| \quad \text{og} \quad \arg \overline{z} = -\arg z, \tag{184}$$

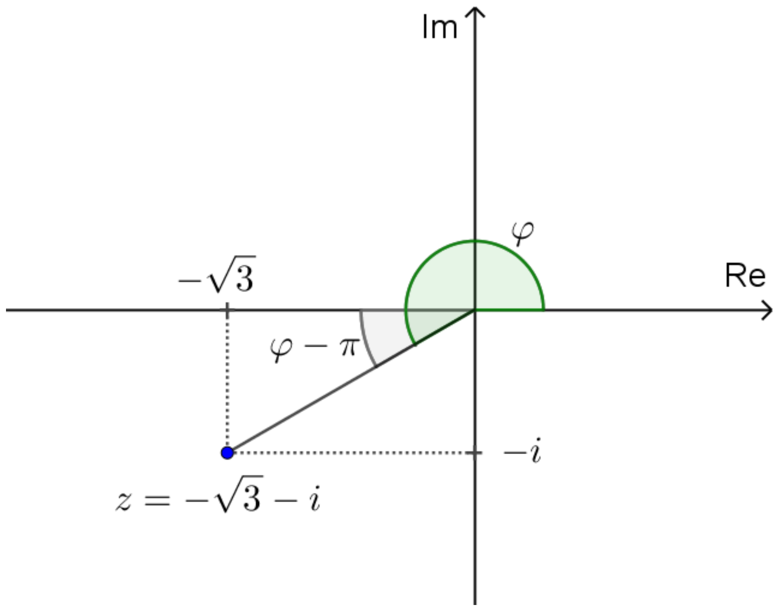
som vist i følgende figur:



Modulus og argument til kompleks konjugerte tall

Eksempel 8.2.4. Finn polarformen til det komplekse tallet $z = -\sqrt{3} - i$.

Løsning. Vi starter med å legge merke til at tallet ligger i tredje kvadrant i det komplekse planet. Vi plotter tallet inn i en figur:



TALLET $z = -\sqrt{3} - i$ I DET KOMPLEKSE PLAN.

Vi regner ut avstanden fra origo $|z| = \sqrt{(\sqrt{3})^2 + 1^2} = \sqrt{4} = 2$. På figuren har vi markert et argument φ for z. Ved elementær trigonometri ser vi at

$$\tan(\varphi - \pi) = \frac{1}{\sqrt{3}}.$$

Siden $\varphi - \pi \in [0, \frac{\pi}{2}]$, får vi $\varphi - \pi = \frac{\pi}{6}$. Altså er $\varphi = \frac{7\pi}{6}$, slik at

$$z = 2\left(\cos\frac{7\pi}{6} + i\sin\frac{7\pi}{6}\right).$$

Vi kan erstatte $\frac{7\pi}{6}$ med et hvilket som helst tall på formen $\frac{7\pi}{6} + 2\pi n$ for $n \in \mathbb{Z}$. Det naturlige valget er nok å velge $\frac{7\pi}{6}$, slik vi har gjort, eller $\frac{7\pi}{6} - 2\pi = -\frac{5\pi}{6}$. □

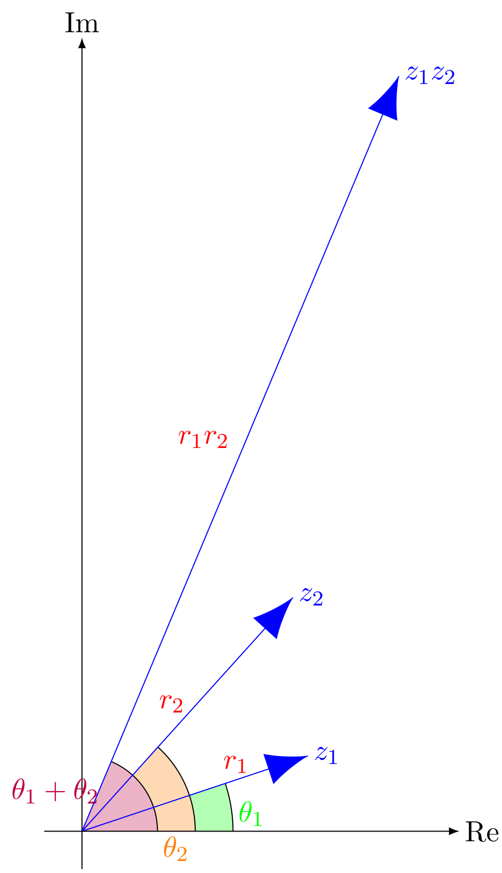
EKSEMPEL 8.2.5. Finn standardformen til det komplekse tallet z som oppfyller $|z| = 3$ og $\varphi = \arg z = \frac{7\pi}{3}$.

LØSNING. Vi har

$$\begin{aligned} z &= |z|(\cos\varphi + i\sin\varphi) = 3\left(\cos\frac{7\pi}{3} + i\sin\frac{7\pi}{3}\right) \\ &= 3\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right) = 3\left(\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) = \frac{3}{2} + i\frac{3\sqrt{3}}{2}. \end{aligned}$$

□

Multiplikasjon i det komplekse plan. La oss nå se på hvordan vi kan tolke multiplikasjonen gitt i (180) geometrisk. Wessels store oppdagelse var at vi *multipliserer to komplekse tall ved å multiplisere modulusene og addere argumentene*, som vist i følgende figur:



Multiplikasjon av komplekse tall: multipliser modulusene og adder argumentene

Det presise resultatet er:

Setning 8.2.6: Multiplikasjon av to komplekse tall på polarform

La $z_1 = r_1 (\cos\theta_1 + i \sin\theta_1)$ og $z_2 = r_2 (\cos\theta_2 + i \sin\theta_2)$ være to komplekse tall på polarform. Da er

$$z_1 z_2 = r_1 r_2 (\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)) .$$

Bevis

Dette vises ved å gange ut og bruke formlene for cosinus og sinus til en sum og overlates til Oppgave 8.16(a). □

Setningen kan generaliseres til produktet av vilkårlig mange komplekse tall:

Setning 8.2.7: Multiplikasjon av komplekse tall på polarform

La $z_k = r_k(\cos\theta_k + i\sin\theta_k)$, $k = 1, \ldots, n$, være komplekse tall på polarform. Da er

$$z_1 \cdots z_n = r_1 \cdots r_n \left(\cos(\theta_1 + \cdots + \theta_n) + i\sin(\theta_1 + \cdots + \theta_n)\right).$$

Bevis

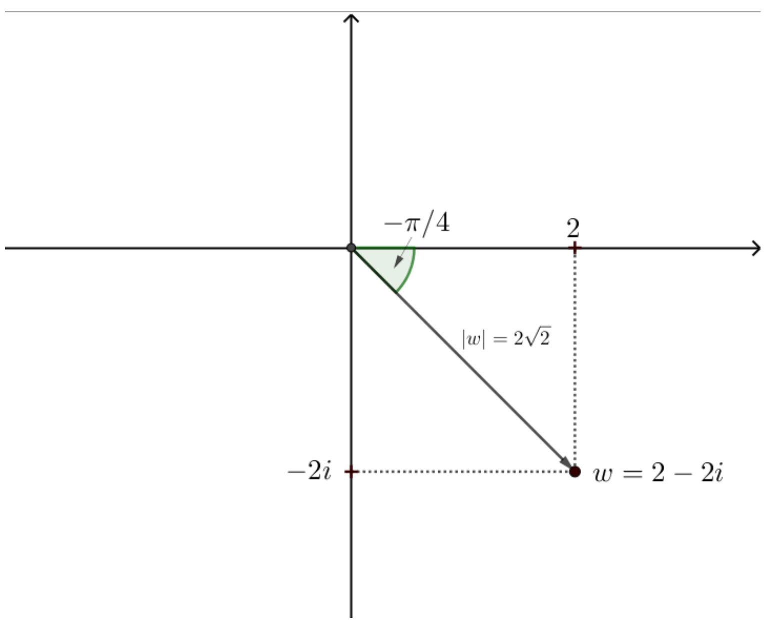
En opplagt induksjon viser setningen ved å bruke Setning 8.2.6 i basistilfellet og induksjonstrinnet (jf. Oppgave 8.16(b)). $\square$

Eksempel 8.2.8. Finn w^8 når $w = 2 - 2i$.

Løsning. Vi har

$$|w| = \sqrt{2^2 + 2^2} = \sqrt{8} = \sqrt{2^2 \cdot 2} = 2\sqrt{2}.$$

Tegnet inn i det komplekse plan, ligger tallet i fjerde kvadrant:



Tallet $w = 2 - 2i$ i det komplekse plan.

Tangens til vinkelen mellom den reelle aksen aksen og vektoren som tilsvarer tallet er $\frac{2}{2} = 1$, derfor er vinkelen $\frac{\pi}{4}$. Et argument for w er da $-\frac{\pi}{4}$. Derfor har vi

$$w = 2\sqrt{2}\Big(\cos\left(-\frac{\pi}{4}\right) + i\sin\left(-\frac{\pi}{4}\right)\Big)$$

på polarform. Dermed har vi

$$\begin{aligned} w^8 &= \left(2\sqrt{2}\right)^8\left(\cos 8\left(-\frac{\pi}{4}\right) + i\sin 8\left(-\frac{\pi}{4}\right)\right) \\ &= \left(2\sqrt{2}\right)^8\left(\cos(-2\pi) + i\sin(-2\pi)\right) \\ &= 2^8 \cdot \sqrt{2}^8\left(\cos 0 + i\sin 0\right) \\ &= 2^8 \cdot 2^4\left(1 + i\cdot 0\right) = 2^{12} = 4096. \end{aligned}$$

□

Spesialtilfellet av Setning 8.2.7 der $r_k = 1$ og $\theta_k = \theta$ (fast) for alle k, gir oss en formel som er oppkalt etter den franske matematikeren Abraham de Moivre (1667–1754), selv om den visstnok ikke finnes i noen av hans arbeider:

Følgesetning 8.2.9: De Moivres formel

For alle $n \in \mathbb{Z}$ er

$$(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta).$$

Eksponensialnotasjonen. Resultatene i Setningene 8.2.6-8.2.7 gir opphavet til følgende definisjon, hvor e er eulertallet (Definisjon 7.5.7):

Definisjon 8.2.10: Kompleks eksponensialfunksjon

Dersom $z = x + iy \in \mathbb{C}$ definerer vi

$$e^z = e^x(\cos y + i\sin y). \tag{185}$$

Med andre ord er $e^z \in \mathbb{C}$ med $|e^z| = e^x$ og $\arg e^z = y$.

Funksjonen $\mathbb{C} \to \mathbb{C}$ gitt ved $z \mapsto e^z$ kalles *den komplekse eksponensialfunksjonen.*

Vi ser spesielt at dersom $z = x \in \mathbb{R}$, det vil si $y = 0$, så er (185) den vanlige definisjonen på e^x. Grunnen til at denne definisjonen er fornuftig er at Setning 8.2.6 oversettes til en generalisering av regelen i Setning 7.3.9(i) for eksponensialfunksjoner:

Setning 8.2.11

For alle $z_1, z_2 \in \mathbb{C}$ gjelder

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

Bevis
Lar vi $z_1 = x_1 + iy_1$ og $z_2 = x_2 + iy_2$ og bruker Definisjon 8.2.10 og Setning 8.2.6, får vi

$$\begin{aligned} e^{z_1}e^{z_2} &= e^{x_1+iy_1}e^{x_2+iy_2} \\ &= (e^{x_1}(\cos(y_1) + i\sin(y_2))) \cdot (e^{x_2}(\cos(y_2) + i\sin(y_2))) \\ &= e^{x_1+x_2}(\cos(y_1+y_2) + \sin(y_1+y_2)) \\ &= e^{(x_1+x_2)+i(y_1+y_2)} = e^{z_1+z_2}. \end{aligned}$$

□

Som konsekvens kan vi uttrykke de Moivres formel (Følgesetning 8.2.9) på en måte som er lett å huske:

$$\left(e^{i\theta}\right)^n = e^{in\theta} \quad \text{(De Moivres formel).} \tag{186}$$

Dessuten gjør Definisjon 8.2.10 det mulig å skrive polarformen til et komplekst tall på en mer kompakt måte:

$$z = |z|(\cos\theta + i\sin\theta) = |z|e^{i\theta}.$$

Vi merker oss også at

$$|e^z| = e^x \text{ for } z = x + iy \in \mathbb{C}$$

slik at spesielt

$$|e^{i\theta}| = 1 \text{ for } \theta \in \mathbb{R}.$$

Vi vil bruke denne mer kompakte notasjonen i neste avsnitt når vi skal trekke røtter av komplekse tall.

Merknad 8.2.12. Vi har egentlig jukset litt rent historisk i fremstillingen av stoffet. I utgangspunktet ble e^z for $z \in \mathbb{C}$ definert ved hjelp av en grense av en følge som

$$e^z = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{z^k}{k!}$$

eller

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

med notasjonen for rekker nevnt mot slutten av §7.4. (For $z = 1$, er dette Merknad 7.5.8). Man kan vise at følgen/rekken konvergerer for alle z. Med denne definisjonen *beviste* Leonhard Euler (1707–1783) formelen (185) i 1748, som i dag bærer navnet *Eulers formel*. Dette skjedde ca. 50 år før Wessels fremstilling av komplekse tall i det komplekse plan.

Vi merker oss at tilfellet $z = i\pi$ i Eulers formel gir oss identiteten

$$e^{i\pi} + 1 = 0.$$

Denne formelen er kjent som *Eulers identitet* og er en av de mest berømte formlene i matematikken, siden den relaterer de fem kanskje viktigste matematiske konstantene: 0, 1, i, π og e.

En meningsmåling utført blant leserne til det matematiske populærvitenskapelige tidsskriftet *The Mathematical Intelligencer* i 1990 kåret Eulers identitet til “det vakreste teoremet i matematikk”. I en annen meningsmåling blant leserne til tidsskriftet *Physics World* i 2004 ble Eulers identitet kåret sammen med Maxwells elektromagnetiske ligninger til “den største ligningen noensinne”. Matematikkprofessor Keith J. Devlin (1947–) har sagt “like a Shakespearean sonnet that captures the very essence of love, or a painting that brings out the beauty of the human form that is far more than just skin deep, Euler’s equation reaches down into the very depths of existence”. Matematikeren Benjamin Peirce (1809–1880) skal ha sagt følgende under en forelesning, etter å ha bevist Eulers identitet: “it is absolutely paradoxical; we cannot understand it, and we don’t know what it means, but we have proved it, and therefore we know it must be the truth".

Fysikeren Richard Feynman (1918–1988) kalte Eulers formel “our jewel” og “the most remarkable formula in mathematics”.

8.3. Røtter av komplekse tall

Vi vet at likningen $y^n = x$, der $x \in \mathbb{R}$ er et fiksert tall, har nøyaktig én løsning i de reelle tallene når n er et oddetall, og nøyaktig to løsninger såfremt $x > 0$ når n er et partall. Vi har vist hvordan dette følger av kompletthetsprinsippet for $\mathbb{R}$ i Setning 7.3.1 og Merknad 7.3.4. Vi kaller den entydige løsningen i tilfellet n oddetall for *n-te roten* til x, og betegner den med $\sqrt[n]{x}$. Det samme gjør vi med den positive løsningen i tilfellet n er partall, og da er $-\sqrt[n]{x}$ den andre løsningen.

Over de komplekse tallene skal vi nå se at situasjonen er annerledes: vi har alltid n løsninger, eller n “røtter”. Med andre ord har ligningen

$$w^n = z \tag{187}$$

der $z \in \mathbb{C}$ er et fiksert tall (mens w er den ukjente) n forskjellige løsninger i $\mathbb{C}$, med mindre $z = 0$ hvor vi bare har én løsning. Vi behandler alle disse løsningene likeverdig og kaller dem *n-te røttene til z*. Er $n = 2$, kaller vi dem *kvadratrøtter*, og er $n = 3$ kaller vi dem *kubikkrøtter*.

Ved å skrive tallet z på polarform og bruke resultatene for multiplikasjon fra forrige seksjon, kan vi gi en presis formel for alle n-te røttene:

Setning 8.3.1

La $z = |z|e^{i\theta} \in \mathbb{C} \setminus \{0\}$. Da har ligningen $w^n = z$ nøyaktig n løsninger $w = w_k$ for $k \in \{0, \ldots, n-1\}$, som er gitt ved

$$w_k = \sqrt[n]{|z|}e^{i\left(\frac{\theta}{n} + k\cdot\frac{2\pi}{n}\right)}, \quad k = 0, 1, 2, \ldots, n-1.$$

Bevis
Vi skriver både w og z på polarform:

$$w = |w|e^{i\varphi} \tag{188}$$

og

$$z = |z|e^{i\theta}. \tag{189}$$

Ved å sette inn (188) og (189) i ligningen (187) og bruke De Moivres formel (186), får vi

$$\begin{aligned} w^n &= z \\ |w|^n e^{in\varphi} &= |z|e^{i\theta}. \end{aligned}$$

Så bruker vi at to komplekse tall forskjellig fra null er like hvis og bare hvis modulusene er like og argumentene er like opptil å addere multipler av 2π (Observasjon 8.2.3). Da får vi at:

$$|w|^n = |z| \quad \text{og} \quad n\varphi = \theta + k \cdot 2\pi, \quad k \in \mathbb{Z},$$

med andre ord:

$$|w| = \sqrt[n]{|z|} \quad \text{og} \quad \varphi = \frac{\theta}{n} + k \cdot \frac{2\pi}{n}, \quad k \in \mathbb{Z},$$

som gir oss (når vi setter inn i (188)):

$$w = \sqrt[n]{|z|}e^{i\left(\frac{\theta}{n} + k \cdot \frac{2\pi}{n}\right)}, \quad k \in \mathbb{Z}.$$

Vi ser at alle disse løsningene har modulus $\sqrt[n]{|z|}$, det vil si de ligger alle på sirkelen med sentrum i origo og radius $\sqrt[n]{|z|}$ i det komplekse plan. Siden to argumenter med differanse et multippel av 2π gir oss samme komplekse tall, ser vi at vi får for eksempel samme løsning for $k = 0$ og $k = n$, og samme for $k = -1$ og $k = n - 1$, osv. Dette betyr at vi kun har forskjellige løsninger for $k = 0, \ldots, n - 1$. Dette gir oss de n løsningene w_k som vi hevdet. □

Vi merker at forskjellen i argument mellom to nærliggende løsninger er $\frac{2\pi}{n}$ og at vi kan beregne alle løsningene med utgangspunkt i

$$w_0 = \sqrt[n]{|z|}e^{i\left(\frac{\theta}{n}\right)}, \tag{190}$$

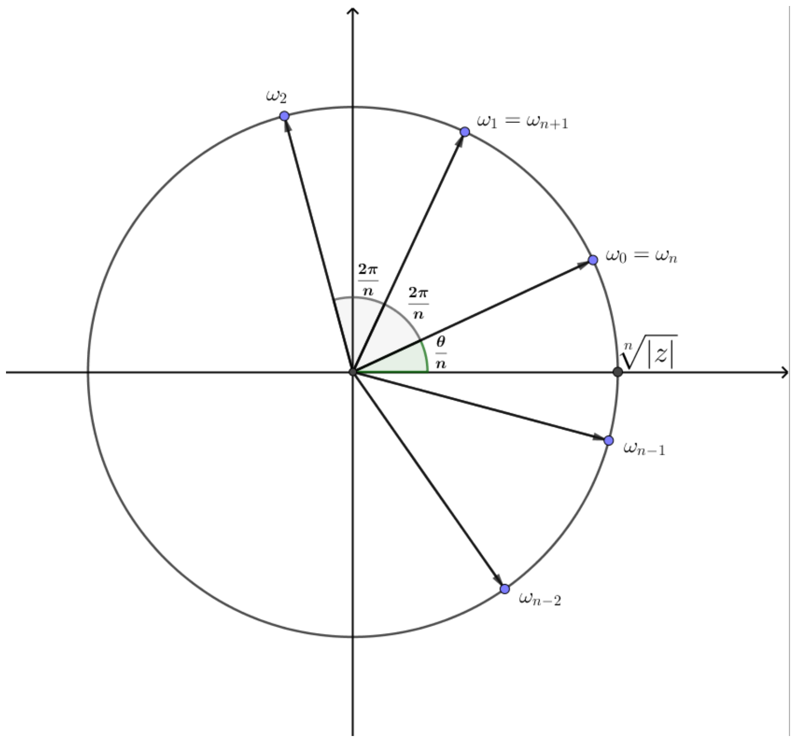
nemlig

$$\begin{aligned} w_1 &= \sqrt[n]{|z|}e^{i\left(\frac{\theta}{n} + \frac{2\pi}{n}\right)} = \sqrt[n]{|z|}e^{i\left(\frac{\theta}{n}\right)} \cdot e^{i\left(\frac{2\pi}{n}\right)} = w_0 \cdot e^{i\left(\frac{2\pi}{n}\right)} \\ w_2 &= \sqrt[n]{|z|}e^{i\left(\frac{\theta}{n} + 2 \cdot \frac{2\pi}{n}\right)} = \sqrt[n]{|z|}e^{i\left(\frac{\theta}{n} + \frac{2\pi}{n}\right)} \cdot e^{i\left(\frac{2\pi}{n}\right)} = w_1 \cdot e^{i\left(\frac{2\pi}{n}\right)} \\ &\text{osv...} \end{aligned}$$

Med andre ord har vi generelt at

$$w_k = w_{k-1} \cdot e^{i\left(\frac{2\pi}{n}\right)} = w_0 \cdot e^{i\left(k \cdot \frac{2\pi}{n}\right)}, \quad k = 1, \ldots, n - 1. \tag{191}$$

Dersom vi plotter inn alle løsningene $w_0, \ldots, w_{n-1}$ i det komplekse plan, ser vi at forskjellen i argument mellom to nærliggende løsninger er $\frac{2\pi}{n}$ og at alle ligger symmetrisk fordelt på sirkelen av radius $\sqrt[n]{|z|}$ i det komplekse planet:



Følgesetning 8.3.2

Annengradslikningen

$$w^2 = z, \quad \text{der} \quad z = |z|e^{i\theta} \in \mathbb{C}$$

har løsningene

$$w = \pm\sqrt{|z|}e^{i\frac{\theta}{2}}.$$

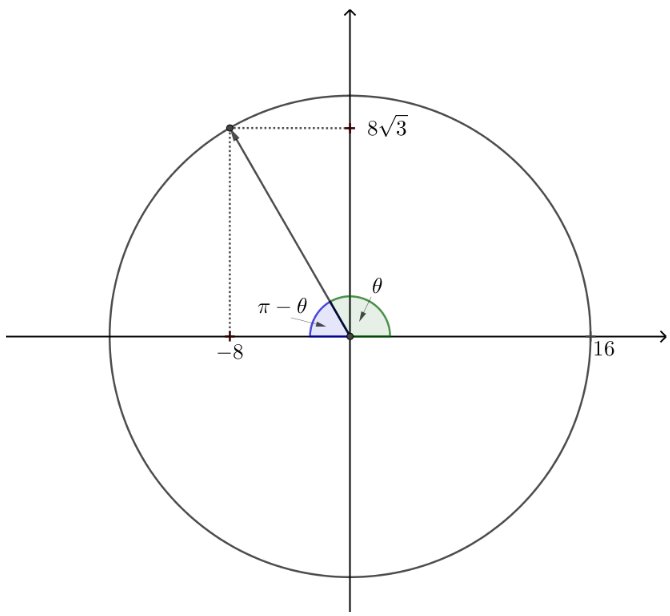
Vi ser spesielt at dersom $z \in \mathbb{R}^{\geq 0}$, det vil si $\theta = 0$, da får vi tilbake de velkjente reelle løsningene $w = \pm\sqrt{z}$.

Eksempel 8.3.3. Finn fjerderøttene til $-8 + 8\sqrt{3}i$.

Løsning. Vi skal løse ligningen

$$w^4 = -8 + 8\sqrt{3}i$$

for w. Vi setter $z = -8+8\sqrt{3}i$ og tegner inn z i det komplekse plan (i andre kvadrant, pga. fortegnene til den reelle og den imaginære delen):



TALLET $z=-8+8\sqrt{3}i$ I DET KOMPLEKSE PLAN.

Vi beregner modulus til z:

$$|z|=\sqrt{(-8)^2+(8\sqrt{3})^2}=\sqrt{8^2(1+3)}=16.$$

Argumentet θ til z beregnes lettest ved å bruke elementær trigonometri:

$$\tan(\pi-\theta)=\frac{8\sqrt{3}}{8}=\sqrt{3}\Longrightarrow \pi-\theta=\frac{\pi}{3}\Longrightarrow \theta=\frac{2\pi}{3}.$$

Dette betyr at vi kan skrive z på polarform som:

$$z=16e^{i\left(\frac{2\pi}{3}\right)}.$$

I henhold til Setning 8.3.1 blir de fire løsningene på ligningen:

$$w_k=2e^{i\left(\frac{\pi}{6}+k\cdot\frac{\pi}{2}\right)},\quad k=0,1,2,3.$$

Disse kan vi alle regne ut ved å regne ut cosinus og sinus til de forskjellige argumentene hver gang (gjerne kombinert med forskjellige regler for cosinus og sinus av summer og differanser av vinkler):

$$\begin{aligned}
w_0 &= 2e^{i\left(\frac{\pi}{6}\right)}=2\left[\cos\left(\frac{\pi}{6}\right)+i\sin\left(\frac{\pi}{6}\right)\right]=2\left[\frac{\sqrt{3}}{2}+i\cdot\frac{1}{2}\right]=\sqrt{3}+i\\
w_1 &= 2e^{i\left(\frac{\pi}{6}+\frac{\pi}{2}\right)}=2\left[\cos\left(\frac{\pi}{6}+\frac{\pi}{2}\right)+i\sin\left(\frac{\pi}{6}+\frac{\pi}{2}\right)\right]=2\left(-\frac{1}{2}+i\cdot\frac{\sqrt{3}}{2}\right)=-1+\sqrt{3}i\\
w_2 &= 2e^{i\left(\frac{\pi}{6}+2\cdot\frac{\pi}{2}\right)}=2\left[\cos\left(\frac{\pi}{6}+\pi\right)+i\sin\left(\frac{\pi}{6}+\pi\right)\right]=2\left(-\frac{\sqrt{3}}{2}-i\cdot\frac{1}{2}\right)=-\sqrt{3}-i\\
w_3 &= 2e^{i\left(\frac{\pi}{6}+3\cdot\frac{\pi}{2}\right)}=2\left[\cos\left(\frac{\pi}{6}+\frac{3\pi}{2}\right)+i\sin\left(\frac{\pi}{6}+\frac{3\pi}{2}\right)\right]=2\left(-\frac{1}{2}-i\cdot\frac{\sqrt{3}}{2}\right)\\
&\qquad =1-\sqrt{3}i.
\end{aligned}$$

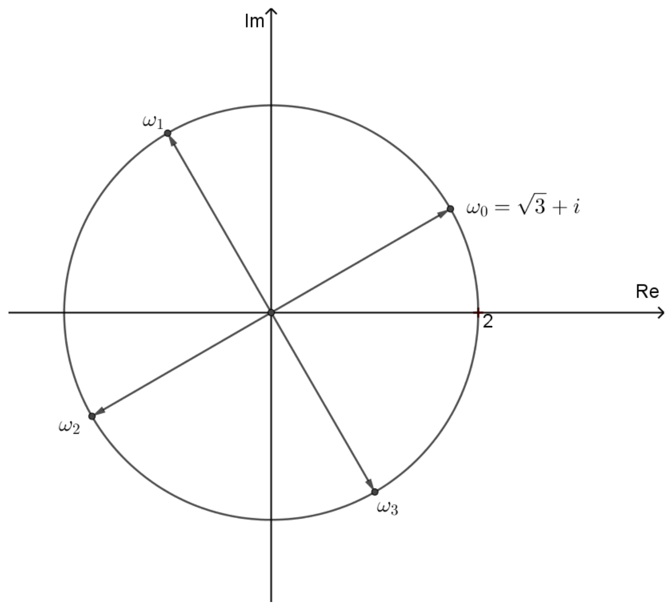
Eller vi kan bruke (191) etter å ha regnet ut w_0 og finne w_1, w_2, w_3 ved å multiplisere hver rot suksessivt med

$$e^{i\left(\frac{2\pi}{4}\right)} = e^{i\left(\frac{\pi}{2}\right)} = \cos\left(\frac{\pi}{2}\right) + i\sin\left(\frac{\pi}{2}\right) = i.$$

Da får vi:

$$\begin{aligned} w_1 &= w_0 \cdot i = \left(\sqrt{3}+i\right) \cdot i = \sqrt{3}i - 1 = -1 + \sqrt{3}i \\ w_2 &= w_1 \cdot i = \left(-1+\sqrt{3}i\right) \cdot i = -i - \sqrt{3} = -\sqrt{3} - i \\ w_3 &= w_2 \cdot i = \left(-\sqrt{3}-i\right) \cdot i = -\sqrt{3}i + 1 = 1 - \sqrt{3}i, \end{aligned}$$

som før. Tegner vi inn $w_0, \ldots, w_3$ i det komplekse plan, ser vi at de alle ligger symmetrisk fordelt på sirkelen av radius 2 om origo og med vinkeldifferanse $\frac{\pi}{2}$ mellom to påfølgende røtter:



FJERDERØTTENE TIL $-8 + 8\sqrt{3}i$ I DET KOMPLEKSE PLAN.

$\square$

EKSEMPEL 8.3.4. La oss vise at prosedyren fra introduksjonen til kapitlet for å finne løsninger på annengradslikninger av typen

$$ax^2 + bx + c = 0, \quad \text{med} \quad a, b, c \in \mathbb{R} \;\text{ og }\; b^2 - 4ac < 0$$

kan rettferdiggjøres matematisk. Mer generelt, la oss vise at alle annengradslikninger

$$az^2 + bz + c = 0, \quad \text{med} \quad a, b, c \in \mathbb{C},$$

har løsninger i $\mathbb{C}$.

Vi starter med samme prosedyren som i introduksjonen:

$$\begin{array}{rcll} az^2+bz+c & = & 0 & \cdot\, 4a \\ 4a^2z^2+4abz+4ac & = & 0 & -\,4ac \\ 4a^2z^2+4abz & = & -4ac & +\,b^2 \\ 4a^2z^2+4abz+b^2 & = & b^2-4ac & \\ (2az+b)^2 & = & b^2-4ac & \end{array}$$

Nå skriver vi $b^2-4ac = |b^2-4ac|e^{i\theta}$ på polarform og bruker Følgesetning 8.3.2, som sier at

$$2az+b = \pm\sqrt{|b^2-4ac|}e^{i\frac{\theta}{2}}.$$

Dette gir

$$z = \frac{-b \pm \sqrt{|b^2-4ac|}e^{i\frac{\theta}{2}}}{2a}. \tag{192}$$

Vi ser at siste formel spesialiserer til løsningsformelen vi kjenner når $a,b,c \in \mathbb{R}$ og $b^2-4ac \geq 0$; da er nemlig $\theta = 0$, og vi får:

$$z = \frac{-b \pm \sqrt{b^2-4ac}}{2a}.$$

Når $a,b,c \in \mathbb{R}$ og $b^2-4ac < 0$ er $\theta = \pi$, slik at $e^{i\frac{\theta}{2}} = e^{i\frac{\pi}{2}} = i$, og

$$z = \frac{-b \pm i\sqrt{b^2-4ac}}{2a},$$

som rettferdiggjør det vi gjorde i introduksjonen til kapitlet.

8.4. Algebraens fundamentalteorem

Innføringen av komplekse tall gjør oss i stand til å løse ligningen $x^2 = -1$, som vi ikke kunne løse over $\mathbb{R}$. Vi har også sett at vi kan løse *alle* annengradsligninger i $\mathbb{C}$ i Eksempel 8.3.4. Faktisk viser det seg at vi få mye mer på kjøpet, nemlig at vi kan løse *alle* polynomielle likninger over $\mathbb{C}$. Dette er en direkte konsekvens av ett av de viktigste resultatene innenfor matematikk:

Teorem 8.4.1: Algebraens fundamentalteorem

La

$$P(z) = c_nz^n + c_{n-1}z^{n-1} + \cdots + c_1z + c_0$$

være et polynom med koeffisienter $c_i \in \mathbb{C}$, $i \in \{0,\ldots,n\}$, $c_n \neq 0$. Da finnes entydige $r_1,\ldots,r_n \in \mathbb{C}$ slik at P kan faktoriseres som

$$P(z) = c_n(z-r_1)(z-r_2)\cdots(z-r_n). \tag{193}$$

Tallet n ovenfor, som altså er eksponenten til den største potensen av variabelen z som dukker opp i polynomet (siden vi krever $c_n \neq 0$), er det vi kaller *graden* til polynomet, jf. §4.1.

Beviset er ganske komplisert og vi vil vente med beviset til §8.6, siden vi også vil trenge noen resultater om følger i $\mathbb{C}$ i beviset, som vi skal utvikle i §8.5. I denne seksjonen vil vi diskutere noen konsekvenser av algebraens fundamentalteorem.

Det første vi merker er at nullpunktene (også kalt *røttene*, jf. Definisjon 4.1.8) til polynomet ovenfor er nøyaktig tallene $r_1, \ldots, r_n$: det er på den ene siden opplagt ved innsetting i høyresiden av (193) at $P(r_i) = 0$ for alle i; motsatt, dersom $P(r) = 0$ for en $r \in \mathbb{C}$, da må (igjen ved innsetting i høyresiden av (193)) $r - r_i = 0$ for en i (husk egenskapen til kropper i Setning 3.6.3(x)), og dermed $r = r_i$. Altså er $\{r_1, \ldots, r_n\}$ lik mengden av røttene til P. Noen av $r_i$ene kan være sammenfallende; vi kaller antallet ganger roten r_i gjentas for *multiplisiteten til roten*. I dette språket kan vi reformulere algebraens fundamentalteorem som:

Teorem 8.4.2

Ethvert kompleks polynom av grad n har nøyaktig n røtter telt med multiplisitet.

Faktorisering av reelle polynomer. Hva så med reelle polynomer, det vil si uttrykk på formen

$$P(x) = c_n x^n + c_{n-1} x^{n-1} + \cdots + c_1 x + c_0$$

med koeffisienter $c_i \in \mathbb{R}$, $i \in \{0, \ldots, n\}$, $c_n \neq 0$? Har vi et tilsvarende resultat om røtter og faktoriseringer av reelle polynomer over $\mathbb{R}$? Reelle polynomer er spesialtilfeller av komplekse polynomer, slik at resultatene i Teoremene 8.4.1 og 8.4.2 fremdeles gjelder, men med alle $r_i \in \mathbb{C}$, hvilket betyr at faktoriseringen gjelder i $\mathbb{C}$, og ikke i $\mathbb{R}$.

Vi observerer at:

Setning 8.4.3

La P være et reelt polynom. Dersom r er en rot i P, da er også den kompleks konjugerte $\overline{r}$ en rot i P.

Bevis

Vi skriver

$$P(z) = c_n z^n + c_{n-1} z^{n-1} + \cdots + c_1 z + c_0$$

som ovenfor. Ved egenskapene til konjugasjon (Setning 8.1.5) har vi

$$\begin{aligned} 0 &= \overline{0} = \overline{P(r)} = \overline{c_n r^n + c_{n-1} r^{n-1} + \cdots + c_1 r + c_0} \\ &= \overline{c_n}\,\overline{r}^n + \overline{c_{n-1}}\,\overline{r}^{n-1} + \cdots + \overline{c_1}\,\overline{r} + \overline{c_0}. \end{aligned}$$

Siden alle $c_i \in \mathbb{R}$ har vi $\overline{c_i} = c_i$, og dermed får vi at

$$0 = c_n \overline{r}^n + c_{n-1} \overline{r}^{n-1} + \cdots + c_1 \overline{r} + \overline{c_0} = P(\overline{r}),$$

som viser setningen. □

MERKNAD 8.4.4. Merk at setningen sier ingenting i tilfellet der $r \in \mathbb{R}$, for da er $\overline{r} = r$. Altså er den kun interessant for *ikke-reelle røtter*, det vil si $r \in \mathbb{C} \setminus \mathbb{R}$.

Dette resultatet sier at alle de ikke-reelle lineære faktorene i faktoriseringen (193) (over $\mathbb{C}$) av et reelt polynom kommer i par av formen $(x-r)(x-\overline{r})$. Multipliserer vi ut og husker at $r\overline{r} = |r|^2 \in \mathbb{R}$ ved (182), finner vi

$$(x-r)(x-\overline{r}) = x^2 - (r+\overline{r})x + r\overline{r} = x^2 - 2(\operatorname{Re} r)x + |r|^2,$$

der begge koeffisientene $-2\operatorname{Re} r \in \mathbb{R}$ og $|r|^2 \in \mathbb{R}$. Dette er altså et reelt annengradspolynom av grad 2, også kalt et (reelt) *kvadratisk polynom*. Dette betyr at vi klarer å pare sammen to og to alle ikke-reelle lineære faktorer i (193) til kvadratiske reelle faktorer. De resterende lineære faktorene er reelle. Vi får derfor den reelle varianten av Teorem 8.4.1:

Teorem 8.4.5: Faktorisering av reelle polynomer

Ethvert reelt polynom kan skrives som et produkt av reelle lineære og kvadratiske faktorer.

8.5. Følger av komplekse tall

Vi kan definere avstanden mellom to komplekse tall z_1 og z_2 som avstanden mellom de tilsvarende punktene i det komplekse planet $|z_1 - z_2|$, og dette gjør $\mathbb{C}$ til et metrisk rom, jf. Merknad 8.2.2. Dermed gir det like mye mening å snakke om konvergens av en følge i $\mathbb{C}$ som i $\mathbb{R}$. (Vi minner om Definisjon 6.2.1, som gjelder for vilkårlige mengder.) Igjen betyr intuitivt at en følge $\{z_n\}$ av komplekse tall konvergerer mot $z \in \mathbb{C}$ dersom vi kan få avstanden mellom leddene z_n og z så liten vi vil for stor nok n. Det viser seg at de aller fleste resultatene for følger i $\mathbb{R}$ gjelder videre for følger i $\mathbb{C}$. Unntaket er resultater som bruker ordensrelasjonen på $\mathbb{R}$, siden $\mathbb{C}$ ikke er en ordnet kropp (jf. Oppgave 8.4), og resultater om divergens mot $\pm\infty$. Vi starter med å gjengi definisjonen:

Definisjon 8.5.1: Konvergens av følge

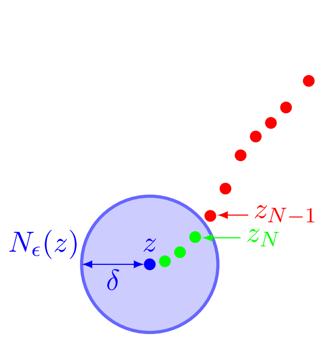
Følgen $\{z_n\}$ av komplekse tall *konvergerer* mot det komplekse tallet z, dersom det for ethvert reelt tall $\epsilon > 0$ finnes et tall $N \in \mathbb{Z}^+$ slik at $|z_n - z| < \epsilon$ for alle $n \geq N$.

Vi skriver i så fall $\lim\limits_{n\to\infty} z_n = z$ eller "$z_n \to z$ når $n \to \infty$" og kaller L for *grensen* til følgen.

Hvis følgen $\{z_n\}$ ikke konvergerer, sier vi at den *divergerer*.

Siden en følge er en funksjon $\mathbb{Z}^+ \to \mathbb{C}$, er det ikke så lett å visualisere grafen til følgen, som er en delmengde av $\mathbb{Z}^+ \times \mathbb{C}$. Vi kan imidlertid visualisere leddene i følgen i det komplekse planet, og da betyr konvergens $z_n \to z$

at uansett hvilken sirkel vi slår om z, så finnes en $N \in \mathbb{Z}^+$ slik at alle ledd i følgen z_n med $n \geq N$ havner på innsiden av sirkelen, som vist på følgende figur:



Konvergens av en følge $z_n \to z$ i det komplekse planet $\mathbb{C}$.

Vi kan skrive $z_n = x_n + iy_n$ med $x_n, y_n \in \mathbb{R}$, slik at $\{x_n\}$ og $\{y_n\}$ er følger i $\mathbb{R}$. Følgende resultat gir oss akkurat det vi håper på:

Setning 8.5.2

Følgen $\{z_n = x_n + iy_n\}$ konvergerer mot $z = x + iy$ hvis og bare hvis $\{x_n\}$ konvergerer mot x og $\{y_n\}$ konvergerer mot y i $\mathbb{R}$.

Bevis
Overlates til Oppgave 8.27. □

Som en umiddelbar konsekvens av dette generaliseres en del resultater for følger i $\mathbb{R}$ til $\mathbb{C}$:

Setning 8.5.3: Entydighet av grenser

Dersom en følge av komplekse tall konvergerer mot z og z', da er $z = z'$.

Bevis
Dette følger direkte av Setningene 7.4.6 og 8.5.2. Alternativt fungerer samme bevis som for Setning 6.2.6 utledet i Oppgave 6.8. □

Setning 8.5.4: Regneregler for grenser av følger

La $\{z_n\}$ og $\{w_n\}$ være konvergente følger av komplekse tall. La $z = \lim_{n\to\infty} z_n$ og $w = \lim_{n\to\infty} w_n$. Da gjelder:

(i) $\lim_{n\to\infty}(z_n + w_n) = z + w$;

(ii) $\lim_{n\to\infty}(z_n - w_n) = z - w$;

(iii) $\lim_{n\to\infty}(cz_n) = cz$ for alle $c \in \mathbb{C}$;

(iv) $\lim_{n\to\infty}(z_n w_n) = zw$;

(v) $\lim_{n\to\infty}\dfrac{z_n}{w_n} = \dfrac{z}{w}$, dersom $w \neq 0$ og $w_n \neq 0$ for alle $n \in \mathbb{Z}^+$.

Bevis
Dette følger direkte av Setningene 7.4.8 og 8.5.2. Alternativt fungerer samme bevis som for Setning 6.2.14. □

Følgesetning 8.5.5

La P være et polynom av én variabel med komplekse koeffisienter (det vil si $P \in \mathbb{C}[x]$ med notasjonen fra 4.1) og $\{z_n\}$ en følge av komplekse tall som konvergerer mot z. Da er

$$\lim_{n\to\infty} P(z_n) = P(z).$$

Bevis
Likt beviset for Følgesetning 7.4.9. □

Også egenskapen av å være begrenset har en naturlig utvidelse fra $\mathbb{R}$ til $\mathbb{C}$:

Definisjon 8.5.6: Begrenset følge

En følge $\{z_n\}$ av komplekse tall er *begrenset* dersom det finnes en $M \in \mathbb{R}$ slik at $|z_n| \leq M$ for alle $n \in \mathbb{Z}^+$.

En følge er *ubegrenset* dersom den ikke er begrenset.

Det er lett å se at:

Setning 8.5.7

Følgen $\{z_n = x_n + iy_n\}$, med $x_n, y_n \in \mathbb{R}$, er begrenset hvis og bare hvis følgene $\{x_n\}$ og $\{y_n\}$ er begrenset.

Bevis
Overlates til Oppgave 8.28. □

En umiddelbar konsekvens er:

Setning 8.5.8

En konvergent følge av komplekse tall er begrenset.

Bevis
Dette er en umiddelbar konsekvens av Setningene 7.4.10 og 8.5.7. Alternativt fungerer samme bevis som for Setning 6.2.11. □

Også egenskapen av å være Cauchy har en naturlig utvidelse fra $\mathbb{R}$ til $\mathbb{C}$:

Definisjon 8.5.9: Cauchy-følge

En følge $\{z_n\}$ av komplekse tall kalles en *Cauchy-følge* dersom det for ethvert reellt tall $\epsilon > 0$ finnes et tall $N \in \mathbb{Z}^+$ slik at $|z_n - z_m| < \epsilon$ for alle $m, n \geq N$.

Igjen er sammenhengen mellom følgen av komplekse tall og de to følgene av de reelle og imaginære delene så pen som vi kunne ønske:

Setning 8.5.10

Følgen $\{z_n = x_n + iy_n\}$, med $x_n, y_n \in \mathbb{R}$, er en Cauchy-følge hvis og bare hvis følgene $\{x_n\}$ og $\{y_n\}$ er Cauchy-følger.

Bevis
Overlates til Oppgave 8.29. □

Som en umiddelbar konsekvens får vi:

Setning 8.5.11

En Cauchy-følge av komplekse tall er begrenset.

Bevis
Dette følger av Setningene 7.4.12 og 8.5.10. Alternativt fungerer samme bevis som i Oppgave 6.13. □

Følgende resultat viser at $\mathbb{C}$ har samme pene egenskap som $\mathbb{R}$ med hensyn på Cauchy-følger:

Setning 8.5.12: Cauchy-kriteriet for konvergens

En følge av komplekse tall konvergerer hvis og bare hvis den er en Cauchy-følge.

Bevis
Beviset for at en konvergent følge av komplekse tall er Cauchy er likt beviset for Setning 6.3.3.

Motsatt, la $\{z_n = x_n + iy_n\}$, med $x_n, y_n \in \mathbb{R}$, være en Cauchy-følge av komplekse tall. Ved Setning 8.5.10 er $\{x_n\}$ og $\{y_n\}$ Cauchy-følger av reelle tall, og er dermed konvergente ved Setning 8.5.12. Men da konvergerer også $\{z_n\}$ ved Setning 8.5.2. □

Merknad 8.5.13. På samme måte som man danner en rekke ut i fra en følge av reelle tall, som forklart mot slutten av §7.4, kan man danne en rekke ut i fra en følge av komplekse tall. Notasjonen og definisjonene er de samme som tidligere.

8.6. Bevis for algebraens fundamentalteorem*

Dette er et bevis av det mer krevende slaget. Underveis vil vi trenge stoffet i §7.7 om delfølger og Bolzano-Weierstrass Teoremet. Det som er vanskeligst er å vise at ethvert komplekst polynom $P(z)$ har en rot, det vil si at det finnes en $r \in \mathbb{C}$ slik at $P(r) = 0$. *Algebraens fundamentalteorem* (Teorem 8.4.1) er en konsekvens av:

Setning 8.6.1

La
$$P(z) = c_n z^n + c_{n-1} z^{n-1} + \cdots + c_1 z + c_0$$
være et polynom med koeffisienter $c_i \in \mathbb{C}$, $i \in \{0, \ldots, n\}$, $c_n \neq 0$, $n > 0$. Da finnes en $r \in \mathbb{C}$ slik at $P(r) = 0$.

La oss se på hvorfor *algebraens fundamentalteorem*, som vi gjentar, følger fra dette:

Teorem 8.4.1: Algebraens fundamentalteorem

La
$$P(z) = c_n z^n + c_{n-1} z^{n-1} + \cdots + c_1 z + c_0$$
være et polynom med koeffisienter $c_i \in \mathbb{C}$, $i \in \{0, \ldots, n\}$, $c_n \neq 0$. Da finnes entydige $r_1, \ldots, r_n \in \mathbb{C}$ slik at P kan faktoriseres som
$$P(z) = c_n (z - r_1)(z - r_2) \cdots (z - r_n).$$

Bevis (gitt Setning 8.6.1)

Vi viser først eksistensdelen ved induksjon på graden til polynomet, som er n. Vi vil altså vise påstanden: *Ethvert komplekst polynom $P(z)$ av grad $n \geq 1$ kan skrives som et produkt $P(z) = c(z - r_1) \cdots (z - r_n)$, der $r_1, \ldots, r_n \in \mathbb{C}$.* (At konstanten c da er lik koeffisienten c_n foran z^n i polynomet er automatisk, så det trenger vi ikke bevise.)

Påstanden er opplagt sann for $n = 1$, siden ethvert førstegradspolynom $P(z) = c_1 z + c_0$ kan skrives som $P(z) = c_1(z - r_1)$, med $r_1 = -c_0/c_1$.

Anta nå at påstanden er sann for et heltall $n \geq 1$ og at $P(z)$ er et polynom av grad $n + 1$. Ved Setning 8.6.1 har $P(z)$ en rot r. Ved Faktorteoremet (Følgesetning 4.1.9) er P delelig med $z - r$, det vil si det finnes et polynom Q slik at
$$P(z) = Q(z)(z - r).$$
Nå kan vi bruke induksjonshypotesen på Q, som har grad n, og beviset for eksistensen følger.

Anta nå at det finnes to forskjellige mengder av røtter $\{r_1, \ldots, r_n\}$ og $\{r'_1, \ldots, r'_n\}$ slik at
$$P(z) = c(z - r_1) \cdots (z - r_n) = c'(z - r'_1) \cdots (z - r'_n).$$

Vi kan etter å omrokkere faktorene, anta at de $n-k$ siste faktorene, for et heltall $1 \leq k \leq n$, er like, og at de k første er ulike. Ved å dele på faktorene som er felles på hver side vil vi ende opp med likheten

$$c(z-r_1)\cdots(z-r_k) = c'(z-r_1')\cdots(z-r_k').$$

Setter vi inn r_1 får vi 0 på grunn av venstresiden, men da må minst én av $r_1 - r_j'$, for $j \in \{1,\ldots,k\}$, være null ved Setning 3.6.3(x). Dette betyr at $r_1 = r_j'$ for en $j \in \{1,\ldots,k\}$, en motsigelse, siden vi hadde antatt vi hadde delt ut alle felles faktorer. □

Vi vil nå gå i gang med beviset for Setning 8.6.1. Beviset er todelt:

- vi viser først at det finnes et punkt $a \in \mathbb{C}$ der $|P(z)| \in \mathbb{R}$ er minst mulig (i Hjelpesetning 8.6.3);
- vi viser så at $P(a) = 0$, ved å vise at dersom dette ikke var sant, ville vi kunne bevege oss i en retning fra a i det komplekse planet og finne z slik at $|P(z)| < |P(a)|$, som motsier forrige punkt.

Vi viser først:

Hjelpesetning 8.6.3

Funksjonen

$$\begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{R} \\ z & \mapsto & |P(z)| \end{array}$$

antar en minimalverdi i et punkt $a \in \mathbb{C}$, det vil si

$$|P(a)| \leq |P(z)| \text{ for alle } z \in \mathbb{C}. \tag{194}$$

Bevis

Verdimengden til den ovennevnte funksjonen er

$$V = \{|P(z)| \mid z \in \mathbb{C}\},$$

som er en nedad begrenset delmengde av $\mathbb{R}$ (siden 0 er en nedre skranke). Mengden har derfor en største nedre skranke $m = \inf V$. Vi vil vise at det finnes en $a \in \mathbb{C}$ slik at $|P(a)| = m$, med andre ord at $m \in V$.

Ved Oppgave 8.18 finnes en $R \in \mathbb{R}^+$ slik at

$$|P(z)| \geq m+1 \text{ for alle } z \in \mathbb{C} \text{ som oppfyller } |z| > R. \tag{195}$$

Vi kan altså begrense vår søken etter a til den begrensede mengden

$$S = \{z \in \mathbb{C} \mid |z| \leq R\}$$

(som er det indre av en sirkel inkludert randen), siden (195) sier at

$$|P(z)| \geq m+1 \text{ for alle } z \notin S. \tag{196}$$

For enhver $n \in \mathbb{Z}^+$ er $m + \frac{1}{n}$ per definisjon *ikke* en nedre skranke for V. Det betyr at det finnes et punkt $z_n \in \mathbb{C}$ slik at

$$m \leq |P(z_n)| < m + \frac{1}{n} \leq m+1;$$

spesielt må $z_n \in S$, ved (196), som betyr at følgen $\{z_n\}$ er begrenset. Ved *Bolzano-Weierstrass teoremet for komplekse følger* (Oppgave 8.31(b)) har $\{z_n\}$ en konvergent delfølge, det vil si (Definisjon 7.7.2) at vi kan finne en følge av positive heltall

$$n_1 < n_2 < n_3 < \cdots < n_k < \cdots$$

slik at følgen $\{z_{n_k}\}_{k=1}^{\infty}$ konvergerer. La $a \in \mathbb{C}$ være grensen til følgen. Ved Følgesetning 8.5.5 vil $P(z_{n_k})$ konvergere mot $P(a)$, og ved Oppgave 8.30 vil $|P(z_{n_k})|$ derfor konvergere mot $|P(a)|$. Siden

$$m \leq |P(z_{n_k})| \leq m + \frac{1}{n_k},$$

vil $|P(z_{n_k})|$ konvergere mot m, ved *Skvisesetningen* (Setning 7.4.13). Ved entydigheten av grenser (Setning 8.5.3), er $m = |P(a)|$, og beviset er fullført. □

Vi vil nå fullføre beviset for *algebraens fundamentalteorem* ved å vise at $P(a) = 0$ for a som i hjelpesetningen.

Bevis for Setning 8.6.1

La a være punktet i Hjelpesetning 8.6.3 slik at $|P(z)|$ er minimal for $z = a$. Vi vil vise at $P(a) = 0$. Vi antar dermed at $P(a) = A \neq 0$ og viser at dette fører til en selvmotsigelse.

Vi skriver først om:

$$P(z) = c_n((z-a)+a)^n + c_{n-1}((z-a)+a)^{n-1} + \cdots + c_1((z-a)+a) + c_0.$$

Multipliserer vi ut alle parentesene, men beholder potensene av $z - a$, får vi et uttrykk på formen

$$P(z) = d_n(z-a)^n + d_{n-1}(z-a)^{n-1} + \cdots + d_1(z-a) + d_0,$$

med $d_i \in \mathbb{C}$, $i \in \{0, \ldots, n\}$ og $d_n = c_n \neq 0$. Vi ser at $d_0 = P(a) = A$, som vi har antatt ikke er 0. La

$$k = \min\{i \in \mathbb{Z}^+ \mid d_i \neq 0\},$$

hvor uttrykket min betyr "det minste av tallene", det vil si at d_k er den nest laveste koeffisienten forskjellig fra null. Da har vi

$$P(z) = d_n(z-a)^n + d_{n-1}(z-a)^{n-1} + \cdots + d_k(z-a)^k + A,$$

med $A \neq 0$ og $d_k \neq 0$. Ved å definere

$$Q(z) = d_n(z-a)^{n-k} + d_{n-1}(z-a)^{n-1-k} + \cdots + d_{k+1}(z-a), \tag{197}$$

kan vi skrive

$$P(z) = Q(z)(z-a)^k + d_k(z-a)^k + A. \tag{198}$$

(Her er $Q = 0$ dersom $k = n$.)

Vi stopper opp et øyeblikk for å antyde tankegangen i resten av beviset. Målet er nå å vise at det finnes et punkt z nær a slik at $|P(z)| < |A|$, som vil motsi Hjelpetning 8.6.3. Idéen er å velge z slik at det komplekse tallet $d_k(z-a)^k$ peker i motsatt retning av A (sett som vektor i det komplekse planet). Da vil $|d_k(z-a)^k + A| < |A|$, og hvis z er nær nok a, blir leddet $Q(z)(z-a)^k$ i (198) så lite at også $|P(z)| < |A|$. For å finne riktig z skriver vi $z = a+\epsilon w$ for en liten nok $\epsilon \in \mathbb{R}^+$ som vi bestemmer senere og en passe $w \in \mathbb{C}$. Da kan vi skrive $d_k(z-a)^k = d_k(\epsilon w)^k = \epsilon^k d_k w^k$. Vi ønsker at denne skal peke i motsatt retning av A, hvilket betyr at $d_k w^k$ skal være (positiv) proporsjonal med $-A$. Siden proporsjonalitetskonstanten ikke spiller noen rolle (vi har jo multiplisert $d_k w^k$ med ϵ^k, som vi skal justere etter hvert), kan vi velge w slik at $d_k w^k = -A$, det vil si vi velger w til å være en k-te rot av $\frac{-A}{d_k}$, som vi vet eksisterer ved Setning 8.3.1.

For å oppsummere, lar vi

$$z = a + \epsilon w \quad \text{med} \quad w^k = \frac{-A}{d_k}. \tag{199}$$

Innsatt i (198) gir dette

$$\begin{aligned} P(z) &= P(a+\epsilon w) = Q(a+\epsilon w)(\epsilon w)^k + d_k(\epsilon w)^k + A \\ &= Q(a+\epsilon w)\cdot \epsilon^k \cdot \left(\frac{-A}{d_k}\right) + \epsilon^k d_k \cdot \left(\frac{-A}{d_k}\right) + A \\ &= A\left(1 - \epsilon^k - \epsilon^k \frac{Q(a+\epsilon w)}{d_k}\right). \end{aligned}$$

Ved trekantulikheten får vi

$$\begin{aligned} |P(z)| &= |A|\left|(1-\epsilon^k) - \epsilon^k \frac{Q(a+\epsilon w)}{d_k}\right| \\ &\le |A|\left(|1-\epsilon^k| + \epsilon^k \left|\frac{Q(a+\epsilon w)}{d_k}\right|\right). \end{aligned} \tag{200}$$

Velger vi

$$0 < \epsilon < 1, \tag{201}$$

ser vi at $|1-\epsilon^k| = 1-\epsilon^k$. Klarer vi å velge $\epsilon \in \mathbb{R}^+$ slik at også

$$\left|\frac{Q(a+\epsilon w)}{d_k}\right| < \frac{1}{2}, \tag{202}$$

da gir (200) at

$$|P(z)| \le |A|\left|1 - \epsilon^k + \epsilon^k \cdot \frac{1}{2}\right| = |A|\left(1 - \frac{\epsilon^k}{2}\right) < |A|,$$

siden $\epsilon > 0$, og vi har fått ulikheten vi er på jakt etter.

Det gjenstår altså bare å vise at vi kan velge $0 < \epsilon < 1$ slik at (202) er oppfylt. Hvis $k = n$ er som sagt tidligere $Q = 0$, slik at (202) er automatisk oppfylt. Hvis $k < n$ bruker vi (197) og skriver

$$\begin{aligned} |Q(a+\epsilon w)| &= |d_n(\epsilon w)^{n-k} + d_{n-1}(\epsilon w)^{n-1-k} + \cdots + d_{k+1}(\epsilon w)| \\ &= \epsilon \left| d_n \epsilon^{n-k-1} w^{n-k} + d_{n-1}\epsilon^{n-k-2} w^{n-1-k} + \cdots + d_{k+1} w \right| \\ &\leq \epsilon \left(|d_n| \epsilon^{n-k-1} |w|^{n-k} + |d_{n-1}| \epsilon^{n-k-2} |w|^{n-1-k} + \cdots + |d_{k+1}||w| \right) \\ &\leq \epsilon \left(|d_n||w|^{n-k} + |d_{n-1}|\ |w|^{n-1-k} + \cdots + |d_{k+1}||w| \right), \end{aligned}$$

hvor vi har brukt trekantulikheten (i den generaliserte versjonen av Oppgave 8.10) og at $\epsilon < 1$. Altså ser vi at (202) er oppfylt dersom

$$0 < \epsilon < \min\left\{1, \frac{|d_k|}{2\left(|d_n||w|^{n-k} + |d_{n-1}|\ |w|^{n-1-k} + \cdots + |d_{k+1}||w|\right)}\right\} \tag{203}$$

(husk at alle $|d_j|$ er faste tall, bestemt av de opprinnelige koeffisientene c_j, og at $|w| = \left(\frac{|A|}{|d_k|}\right)^{\frac{1}{k}}$ fra (199)).

Som oppsummering: for z som i (199) med ϵ som i (203), da har vi $|P(z)| < |A|$, og vi har fått motsigelsen vi søkte. $\square$

Oppgaver

Oppgaver til §8.1–8.2

Oppgave 8.1. Regn ut

$$\overline{(3-2i) + i\frac{3+5i}{2-3i}}.$$

Oppgave 8.2. Løs ligningen

$$\frac{z-a}{z-b} = ci, \quad \text{der} \ \ a, b, c \in \mathbb{R}.$$

Oppgave 8.3. Bevis Setning 8.1.2.

Oppgave 8.4. Vis at det ikke finnes noen ordensrelasjon på $\mathbb{C}$ som gjør $\mathbb{C}$ til en ordnet kropp (jf. Merknad 8.1.3). (Hint: $-1 = i^2$.)

Oppgave 8.5. Bevis Setning 8.1.5

Oppgave 8.6. La $z, w \in \mathbb{C}$ slik at $z + w \in \mathbb{R}$ og $zw \in \mathbb{R}$. Vis at enten er $z, w \in \mathbb{R}$ eller så er z og w konjugerte av hverandre.

Oppgave 8.7. Gjør $\mathbb{C}$ til en ordnet mengde ved å identifisere $\mathbb{C}$ med $\mathbb{R} \times \mathbb{R}$ og bruke *leksikografisk orden* (jf. Eksempel 3.6.11 og Oppgave 3.17): dette betyr at vi definerer $a + bi < c + di$ hvis $a < c$ eller $a = c$ og $b < d$. Oppfyller $\mathbb{C}$ da kompletthetsprinsippet?

Oppgave 8.8. Vis at $\mathbb{C}$ med sine operasjoner gjør $\mathbb{C}$ til et $\mathbb{R}$-vektorrom (jf. Definisjon 4.5.5). Vis mer generelt at $\mathbb{C}^n$ er et vektorrom med naturlige operasjoner.

Oppgave 8.9. Bevis Setning 8.2.1

(a) ved å argumentere geometrisk og bruke figuren i beviset;
(b) ved utregning. (Hint: bruk (182) og Setning 8.1.5.)

Oppgave 8.10. Generalisér trekantulikheten ved å vise at

$$|z_1 + \cdots + z_n| \leq |z_1| + \cdots + |z_n|$$

for $z_1, \ldots, z_n \in \mathbb{C}$. (Hint: bruk induksjon.)

Oppgave 8.11. Vis at $||z_1| - |z_2|| \leq |z_1 - z_2|$ for alle $z_1, z_2 \in \mathbb{C}$.

Oppgave 8.12. Vis at for enhver $z \in \mathbb{C}$ er $|\operatorname{Re} z| \leq |z|$ og $|\operatorname{Im} z| \leq |z|$.

Oppgave 8.13. Vis at $\mathbb{C}$, med metrikken angitt i Merknad 8.2.2, er et metrisk rom.

Oppgave 8.14. Finn alle komplekse tall som oppfyller ulikheten

$$|z - 1| \geq |z - i|.$$

Oppgave 8.15. Regn ut

$$(1 + i)^{400}.$$

Oppgave 8.16. (a) Bevis Setning 8.2.6 ved å utnytte at

$$\cos(\theta_1 + \theta_2) = \cos(\theta_1)\cos\theta_2 - \sin(\theta_1)\sin(\theta_2)$$

og

$$\sin(\theta_1 + \theta_2) = \cos(\theta_1)\sin\theta_2 + \sin(\theta_1)\cos(\theta_2).$$

(b) Bevis Setning 8.2.7 ved å bruke induksjon og (a).

Oppgave 8.17. Vis den velkente *Cauchy-Schwarz-ulikheten*: dersom $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{C}$, da er

$$\left|\sum_{i=1}^{n} x_i \overline{y_i}\right|^2 \leq \sum_{i=1}^{n} |x_i|^2 \sum_{i=1}^{n} |y_i|^2.$$

Ulikheten er oppkalt etter Cauchy, som publiserte et bevis i 1821, og den tyske matematikeren Hermann Amandus Schwarz (1843–1921), som beviste en litt annen versjon i 1888.

Oppgave 8.18 (Polynomer av grad > 0 er ubegrenset). La

$$P(z) = c_n z^n + c_{n-1} z^{n-1} + c_1 z + c_0,$$

være et komplekst polynom av grad n, det vil si alle $c_i \in \mathbb{C}$ og $c_n \neq 0$. Vis at uansett gitt $M \in \mathbb{R}^+$, så finnes en $R \in \mathbb{R}^+$ slik at $|P(z)| > M$ for alle $z \in \mathbb{C}$ slik at $|z| > R$.

(Hint: la $z = re^{i\theta}$ og bruk resultatet fra Oppgave 8.11 til å konkludere at

$$|P(z)| = |P(re^{i\theta})| \geq |r^n| \left| |c_n| - \left| \frac{c_{n-1}e^{in\theta}}{r} + \cdots + \frac{c_1 e^{i\theta}}{r^{n-1}} + \frac{c_0}{r^n} \right| \right|.$$

Bruk så resultatet fra Oppgave 8.10 til å vise at

$$\left| \frac{c_{n-1}e^{in\theta}}{r} + \cdots + \frac{c_1 e^{i\theta}}{r^{n-1}} + \frac{c_0}{r^n} \right| \leq \frac{|c_{n-1}|}{r} + \frac{|c_1|}{r^{n-1}} + \frac{|c_0|}{r^n} \leq \frac{|c_n|}{2}$$

for tilstrekkelig stor r.)

Oppgaver til §8.3–8.4

Oppgave 8.19. Finn alle tredjerøttene til $z = -2\sqrt{2} + 2\sqrt{2}i$.

Oppgave 8.20. Løs annengradsligningen

$$2z^2 + 3z + 2 = 0.$$

Oppgave 8.21. La $w_1, \ldots, w_n$ være n-te røttene til et komplekst tall. Vis at

$$w_1 + \cdots + w_n = 0.$$

Oppgave 8.22. "Regnestykket"

$$1 = \sqrt{1} = \sqrt{(-1)\cdot(-1)} = \sqrt{(-1)} \cdot \sqrt{(-1)} = i \cdot i = i^2 = -1$$

skapte lenge forvirring hos matematikere. Hvor ligger feilen?

Oppgave 8.23. Finn den komplekse og den reelle faktoriseringen av polynomet $z^3 + 8$.

Oppgave 8.24. Finn alle komplekse løsninger av ligningen

$$z^3 + z^2 + z + 1 = 0.$$

(Hint: jf. Oppgave 5.10.)

Oppgave 8.25. Vis at et reelt polynom av odde grad alltid har en reell rot.

Oppgaver til §8.5

Oppgave 8.26. Vis at konklusjonene i Oppgavene 7.18 og 7.19 også gjelder for følger av komplekse tall.

Oppgave 8.27. Bevis Setning 8.5.2.
(Hint: $|z_n - z|^2 = |x_n - x|^2 + |y_n - y|^2$.)

Oppgave 8.28. Bevis Setning 8.5.7.

Oppgave 8.29. Bevis Setning 8.5.10.
(Hint: $|z_n - z_m|^2 = |x_n - x_m|^2 + |y_n - y_m|^2$.)

Oppgave 8.30. Vis at dersom følgen $\{z_n\}$ av komplekse tall konvergerer mot z, da konvergerer følgen $\{|z_n|\}$ (av reelle tall) mot $|z|$. (Hint: du kan få bruk for resultatet fra Oppgave 8.11.)

Oppgave 8.31. Bruk resultatene fra §8.5 til å gi korte bevis for de tilsvarende resultatene i Oppgavene 7.27 og 7.46, samt *Bolzano-Weierstrass-teoremet* 7.7.5, for følger av komplekse tall. (Husk Definisjon 7.7.2 av en delfølge.)

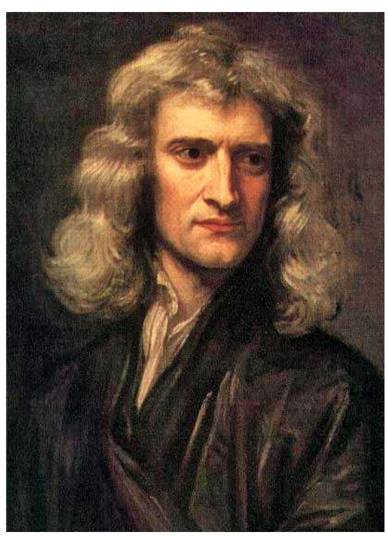
(a) En følge $\{z_n\}$ av komplekse tall konvergerer mot z hvis og bare hvis enhver delfølge av $\{z_n\}$ konvergerer mot z.
(b) (*Bolzano-Weierstrass-teoremet for komplekse følger*): Enhver begrenset følge av komplekse tall har en konvergent delfølge.
(c) Dersom en Cauchy-følge $\{z_n\}$ av komplekse tall har en delfølge som konvergerer mot z, da konvergerer $\{z_n\}$ mot z.

KAPITTEL 9

Kontinuitet

Som nevnt i Eksempel 2.1.16 ble grunnlaget for den moderne *analytiske geometrien* lagt i 1637 gjennom innføringen av koordinatsystemer i to arbeider av Pierre de Fermat (1601–1665) og René Descartes (1596–1650). Grunnlaget for moderne matematikk og funksjonsbegrepet, samt konsepter som kontinuitet, derivasjon og integrasjon var dermed lagt, og utviklingen gikk fort, men uten all den nødvendige teorien på plass.

Både Descartes og Fermat hadde metoder for å beregne tangenter til kurver, med regneteknikker som har klare likhetstrekk med derivasjon. En rimelig fullstendig derivasjonsteori med alle regler og teknikker vi kjenner i dag kom litt senere, nemlig hos den engelske matematikeren, astronomen og fysikeren Isaac Newton (1642–1727) og den tyske filosofen og matematikeren Gottfried Wilhelm Leibniz (1646–1716). Sett med dagens krav til matematisk rigorøsitet var det en del problemer med Newtons og Leibniz' fremgangsmåte, idet de regnet med såkalte *infinitesimaler* (som vi kan tenke på som "bittesmå størrelser") som noen ganger var null og noen ganger ikke: Leibniz selv innførte notasjonen $\frac{dy}{dx}$ for den deriverte og så på dette som en brøk av to bittesmå (eller *infinitesimale*) størrelser dy og dx. Intuitivt kan dette som oftest fungere bra, men det var svært uklart hva disse infinitesimalene var. Newton og Leibniz utviklet også grunnlaget for integrasjonsteorien.



Isaac Newton (Kilde: Wikipedia, public domain)

Gottfried Wilhelm Leibniz (Kilde: Wikipedia, public domain)

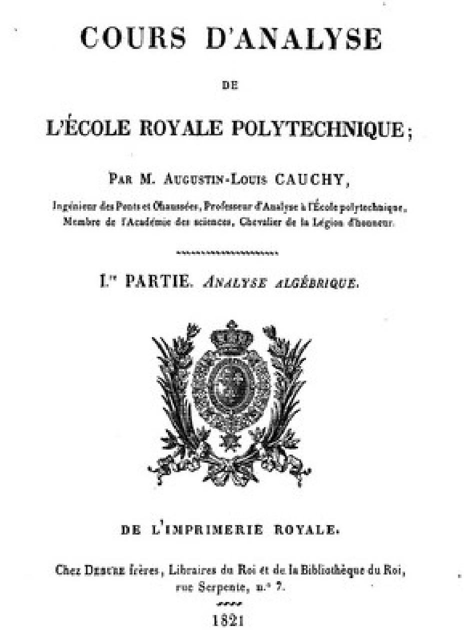
Interessant nok, som nevnt i Merknad 2.2.9, regner man at funksjonsbegrepet ble innført noe senere, av Leonhard Euler (1707–1783), som likevel i begynnelsen ikke godtok alle typer funksjoner som er akseptert idag, som for eksempel funksjoner med delt forskrift. Funksjonsbegrepet som vi kjenner det i dag (Definisjon 2.2.1) ble innført av den tyske matematikeren Peter Gustav Lejeune Dirichlet (1805–1859). Det gjorde han da han ga et bevis for at en metode utviklet av den franske matematikeren og fysikeren Jean-Baptiste Joseph Fourier (1768–1830) for å beskrive varmeledning var matematisk riktig. Fourier klarte aldri å bevise at metoden han hadde utviklet fungerte, og hans hovedverk *Théorie Analytique de la chaleur* ble først publisert i 1822. Fourier brydde seg mest om de praktiske anvendelsene av sine resultater, og mindre om metodene var matematisk korrekte. Teknikkene hans fikk en rask utbredelse etter at de var publisert, for de viste seg å fungere. Dirichlet var interessert i å bruke Fouriers teknikker på problemer i teoretisk matematikk, og måtte derfor bevise at teknikkene var riktige.

Jean-Baptiste Joseph Fourier (Kilde: Wikipedia, public domain)

Først på 1800-tallet fikk vi et matematisk rigorøst grunnlag for matematisk analyse gjennom arbeidene til Bolzano, Cauchy og Weierstrass. Sentralt var begrepet kontinuitet.

En første formell og temmelig rigorøs definisjon av kontinuitet ble gitt av den bohemske matematikeren, logikeren, filosofen, teologen og katolske presten Bernard Bolzano (1781–1848), men hans arbeider forble ukjent for matematikere i mange år.

Konseptet bak den formelle definisjonen av kontinuitet går tilbake til den franske matematikeren og fysikeren (og baronen) Augustin-Louis Cauchy (1789–1857). Han ga riktignok ikke noen formell ϵ–δ-definisjon av kontinuitet (se Definisjon 9.1.1) i sin berømte bok *Cours d'Analyse de l'École Royale Polytechnique* fra 1821, men brukte ϵ–δ-argumenter i bevis. Spesielt viktig var det at hans fremstilling unngikk de logiske problemene som infinitesimalene til Newton og Leibniz ledet til. Interessant falt hans teoretiske fremstillinger ikke i smak hos ingeniørstudentene han underviste ved *École Polytechnique* i Paris, og skolens ledelse påla Cauchy å vende tilbake til den tradisjonelle, om enn upresise, fremstillingen av stoffet!



COURS D'ANALYSE

DE

L'ÉCOLE ROYALE POLYTECHNIQUE;

PAR M. AUGUSTIN-LOUIS CAUCHY,

Ingénieur des Ponts et Chaussées, Professeur d'Analyse à l'École polytechnique, Membre de l'Académie des sciences, Chevalier de la Légion d'honneur.

I.re PARTIE. *ANALYSE ALGÉBRIQUE.*

DE L'IMPRIMERIE ROYALE.

Chez Debure frères, Libraires du Roi et de la Bibliothèque du Roi, rue Serpente, n.° 7.

1821

CAUCHYS BOK *Cours d'Analyse de l'École Royale Polytechnique* (KILDE: WIKIPEDIA, PUBLIC DOMAIN)

Cauchys forelesninger var et enormt fremskritt i utviklingen av en rigorøs fremstilling av matematikken, men løste ikke alle problemer. Fremdeles skulle det gå ca. 40 år før man virkelig fikk på plass en god forståelse av de reelle tallene og kompletthetsprinsippet. I denne perioden var kontinuitetsbegrepet fremdeles ikke helt forstått, blant annet trodde (og "beviste") mange matematikere at kontinuerlige funksjoner var deriverbare bortsett fra endelig mange knekkpunkter. Dette ble motbevist av den tyske matematikeren Karl Weierstrass (1815–1897) gjennom å publisere i 1872 et eksempel på en kontinuerlig, ingensteds deriverbar funksjon, som vi nevnte i §1.4. Den moderne og utvetydige definisjonen av kontinuitet slik vi kjenner den idag, ble gitt av Weierstrass. I sine berømte forelesninger i Berlin la han vekt på hvordan man kunne bygge opp analysen på en fullstendig rigorøs og utvetydig måte ved hjelp av logikk. Weierstrass har mye av æren for grunnlaget for kalkulus og den matematiske analysen slik vi kjenner dem i dag. I tillegg hadde Weierstrass mange doktorstudenter som kom til å få viktige plasser i matematikkens historie. Spesielt må vi nevne den russiske matematikeren Sofya Kovalevskaya (1850–1891), blant annet berømt for sine arbeider innenfor differensialligninger, som ble den andre kvinnelige matematikkprofessoren i historien etter Maria Gaetana Agnesi (1718–1799), da hun ble professor ved Stockholms Universitet i 1889.

Med begrepet kontinuitet på plass, samt konstruksjonen av de reelle tallene, mot slutten av 1800-tallet, falt også alle brikker på plass i teorien om derivasjon og integrasjon utviklet av Netwon og Leibniz.

Vi vil i dette kapitlet se på den formelle definisjonen av kontinuitet, samt en del viktige konsekvenser (§9.1–9.5). Blant annet skal vi se på et par viktige teorememer med mange anvendelser i §9.6–9.7: *Skjæringssetningen* (Teorem 9.6.1) og *ekstremalverdisetningen* (Teorem 9.7.2). Kompletthetsprinsippet til de reelle tallene er en helt sentral ingrediens i begge bevisene. Seksjon 9.8 gjennomgår en av de klassiske og viktigste numeriske metodene for å finne tilnærmingsverdier for nullpunkter til funksjoner: her bruker vi virkelig hele stoffet vi har lært om kompletthet, følger og kontinuitet. Vi avslutter med en seksjon om uniform kontinuitet (§9.9), et noe sterkere begrep enn kontinuitet, og om grensebegrepet (§9.10).

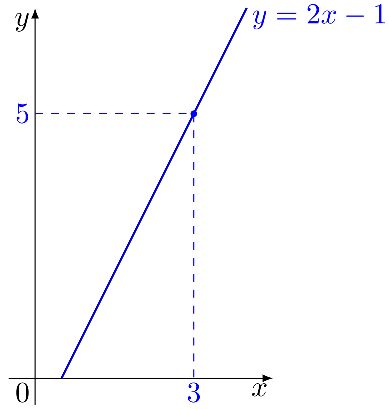
9.1. Definisjon av kontinuitet

Utgangspunktet for kontinuitetsbegrepet er at vi er interessert i å sile ut en viktig egenskap til en funksjon som går ut på at den ikke skal gjøre "plutselige endringer" i et punkt i definisjonsmengden. Mer presist, la $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$ og la $a \in D(f)$. At f er kontinuerlig i punktet a betyr intuitivt at funksjonsverdiene $f(x)$ nærmer seg $f(a)$ når x nærmer seg a. I mange lærebøker definerer man først hva som menes med at funksjonsverdiene $f(x)$ nærmer seg et tall L (dette uttrykkes som $\lim_{x \to a} f(x) = L$), for deretter å definere kontinuitet til å være at $\lim_{x \to a} f(x) = f(a)$. Dette er helt i orden, men siden grensebegrepet ikke vil spille en så stor rolle for oss i første omgang, vil vi definere kontinuitetsbegrepet direkte, som for eksempel i [**Li**]. Vi vil så definere grensebegrepet i siste seksjon i dette kapitlet.

Som et konkret eksempel, la oss starte med å betrakte en relativt enkel funksjon, nemlig en lineær funksjon $f : \mathbb{R} \to \mathbb{R}$ gitt ved uttrykket

$$f(x) = 2x - 1,$$

og punktet $x = 3$. Vi har $f(3) = 5$. Vi spør oss om $f(x)$ nærmer seg 5 når x nærmer seg 3.



GRAFEN TIL f GITT VED $x \mapsto 2x - 1$.

Studerer vi grafen til f, virker det temmelig klart hva som skjer: grafen går mot punktet $(3, 5)$ slik at $f(x)$ nærmer seg $f(3) = 5$ når x nærmer seg 3. Med dette mener vi at vi kan sørge for at

$$f(x) \text{ er så nær vi vil } f(3) \tag{204}$$

så lenge

$$x \text{ er nær nok } 3. \tag{205}$$

Matematisk kan vi uttrykke (204) og (205) henholdsvis ved hjelp av avstander som

$$|f(x) - f(3)| \text{ er så liten vi vil} \tag{206}$$

så lenge

$$|x - 3| \text{ er liten nok.} \tag{207}$$

La oss sjekke om funksjonen vår virkelig oppfyller dette.

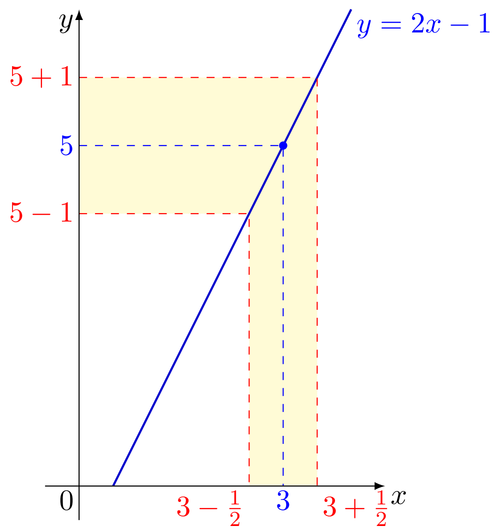
Hvis "så nær vi vil" betyr i avstand < 1 (eller, sagt med andre ord, med "feilmargin" 1), kan vi skrive:

$$\begin{array}{rcl} |f(x) - f(3)| & < & 1 \\ & \Updownarrow & \\ |2x - 1 - 5| & < & 1 \\ & \Updownarrow & \\ |2x - 6| & < & 1 \\ & \Updownarrow & \\ 2|x - 3| & < & 1 \\ & \Updownarrow & \\ |x - 3| & < & \dfrac{1}{2} \end{array}$$

Dette betyr at

Vi kan sørge for at $|f(x) - f(3)| < 1$ (eller "f er nær $f(3)$ med feilmargin 1") så lenge $|x - 3| < \frac{1}{2}$ (eller "x er i avstand mindre enn $\frac{1}{2}$ fra 3").

Grafisk kan det vi har gjort visualiseres i følgende figur:



$f(x)$ er i avstand høyst 1 fra 5 når x er i avstand høyst $\frac{1}{2}$ fra 3.

La oss forsøke om dette fungerer om vi bytter ut "avstand < 1" med noe mindre, si "avstand $< \frac{1}{2}$". Samme utregning med 1 byttet ut med $\frac{1}{2}$ viser:

$$\begin{array}{rcl} |f(x) - f(3)| & < & \dfrac{1}{2} \\ & \Updownarrow & \\ |2x - 1 - 5| & < & \dfrac{1}{2} \\ & \Updownarrow & \\ |2x - 6| & < & \dfrac{1}{2} \\ & \Updownarrow & \\ 2|x - 3| & < & \dfrac{1}{2} \\ & \Updownarrow & \\ |x - 3| & < & \dfrac{1}{4} \end{array}$$

Dette betyr at

Vi kan sørge for at $|f(x) - f(3)| < \frac{1}{2}$ (eller "f er nær $f(3)$ med feilmargin $\frac{1}{2}$") så lenge $|x - 3| < \frac{1}{4}$ (eller "x er i avstand høyst $\frac{1}{4}$ fra 3").

Går det også bra hvis vi krever knøttliten avstand, si "avstand $< \frac{1}{10^6}$"? Ja, fordi vi kan skrive

$$\begin{array}{rcl} |f(x)-f(3)| & < & \dfrac{1}{10^6} \\ & \Updownarrow & \\ |2x-1-5| & < & \dfrac{1}{10^6} \\ & \Updownarrow & \\ |2x-6| & < & \dfrac{1}{10^6} \\ & \Updownarrow & \\ 2|x-3| & < & \dfrac{1}{10^6} \\ & \Updownarrow & \\ |x-3| & < & \dfrac{1}{2\cdot 10^6} \end{array}$$

Dette betyr at

Vi kan sørge for at $|f(x)-f(3)| < \frac{1}{10^6}$ (eller "f er nær $f(3)$ med feilmargin $\frac{1}{10^6}$") så lenge $|x-3| < \frac{1}{2\cdot 10^6}$ (eller "x er i avstand høyst $\frac{1}{2\cdot 10^6}$ fra 3")

At f er kontinuerlig i punktet 3 betyr nettopp at dette fungerer videre uansett hvilken enda mindre avstand vi bytter ut 1, $\frac{1}{2}$ og $\frac{1}{10^6}$ med. Vi skal med andre ord klare å vise generelt hvor liten $|x-3|$ må gjøres for å sørge for at $|f(x)-f(3)|$ blir så liten som ønsket. Den generelle "avstanden som $|f(x)-f(3)|$ må være mindre enn" betegner vi med den greske bokstaven ϵ ("epsilon"), som i Definisjon 7.4.1 for konvergens av følger. Igjen symboliserer ϵ et "vilkårlig lite positivt reellt tall". Beregningene med 1, $\frac{1}{2}$ og $\frac{1}{10^6}$ erstattet med ϵ blir da:

$$\begin{array}{rcl} |f(x)-f(3)| & < & \epsilon \\ & \Updownarrow & \\ |2x-1-5| & < & \epsilon \\ & \Updownarrow & \\ |2x-6| & < & \epsilon \\ & \Updownarrow & \\ 2|x-3| & < & \epsilon \\ & \Updownarrow & \\ |x-3| & < & \dfrac{\epsilon}{2} \end{array}$$

Dette betyr at

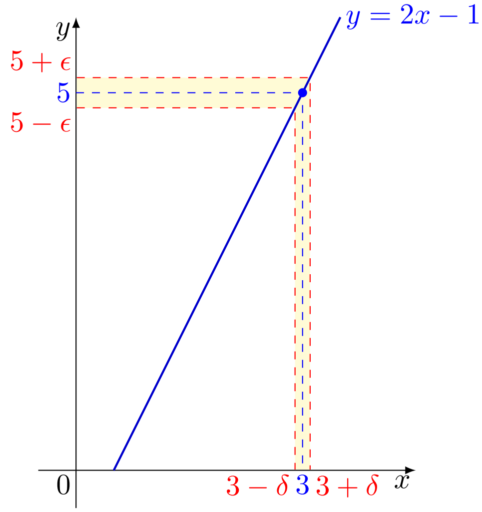
Uansett $\epsilon > 0$, kan vi sørge for at $|f(x)-f(3)| < \epsilon$ så lenge $|x-3| < \frac{\epsilon}{2}$.

Eller, omformulert litt mer matematisk:

For enhver $\epsilon > 0$*, gjelder*

$$|f(x) - f(3)| < \epsilon \text{ når } |x-3| < \frac{\epsilon}{2}.$$

Grafisk kan det vi har gjort igjen visualiseres i følgende figur:



$f(x)$ ER I AVSTAND HØYST ϵ FRA 5 NÅR x ER I AVSTAND HØYST δ FRA 3.

At $|x-3| < \frac{\epsilon}{2}$ betyr at $x \in (3-\frac{\epsilon}{2}, 3+\frac{\epsilon}{2})$, som er en δ-omegn om 3 for $\delta = \frac{\epsilon}{2}$ (se Definisjon 2.1.22). At $|f(x)-f(3)| < \epsilon$ betyr at $f(x) \in (f(3)-\epsilon, f(3)+\epsilon)$, som er en ϵ-omegn om $f(3)$. Vi har altså konkret funnet hvilken δ-*omegn* om 3 som x må ligge i for at $f(x)$ skal ligge i en gitt ϵ-*omegn* om $f(3)$.

Så igjen omformulert, ser vi at vi kan si at:

For enhver $\epsilon > 0$*, finnes en* $\delta > 0$ *slik at*

$$|f(x) - f(3)| < \epsilon \text{ når } |x-3| < \delta,$$

nemlig $\delta = \frac{\epsilon}{2}$. Vi liker å huske dette som:

$$\boxed{\text{gitt en vilkårlig } \epsilon\text{, så klarer vi å finne } \delta.}$$

Vi er nå klare for å gi den *formelle definisjonen av kontinuitet*:

Definisjon 9.1.1: Kontinuitet i et punkt

La $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$ og la $a \in D(f)$. Funksjonen f er *kontinuerlig*[a] i a dersom det til ethvert reelt tall $\epsilon > 0$ finnes et reelt tall $\delta > 0$ slik at

$$|f(x) - f(a)| < \epsilon \text{ når } |x-a| < \delta \text{ og } x \in D(f). \tag{208}$$

Dersom f ikke er kontinuerlig i a, sier vi at f er *diskontinuerlig*[b] i a, eller at f *har en diskontinuitet*[c] i a.

[a] *Continuous* på engelsk.
[b] *Discontinuous* på engelsk.
[c] *Discontinuity* på engelsk.

Merknad 9.1.2. Vi vil se senere, i Setning 9.10.5, at Definisjon 9.1.1 sier det samme som at $\lim_{x \to a} f(x) = f(a)$, når a er et såkalt *opphopningspunkt* i definisjonsmengden til f (jf. Definisjon 9.10.1).

Merknad 9.1.3. Det er ikke noe krav om at f skal være definert "på begge sider av punktet a"; definisjonsmengden $D(f)$ kan for eksempel godt være et intervall med punktet a som endepunkt, eller for den del ikke engang et intervall eller union av slike, som for eksempel $\mathbb{Z}$ og $\mathbb{Q}$.

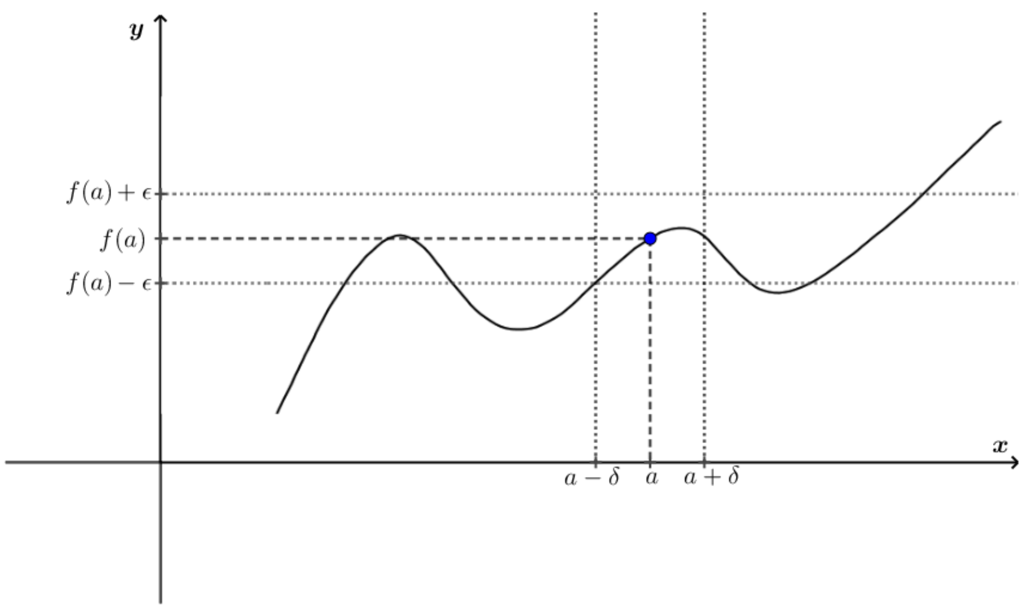
Merk at betingelsen "$|x - a| < \delta$ og $x \in D(f)$" også kan skrives som $x \in (a - \delta, a + \delta) \cap D(f)$, slik at den betyr at x ligger i en δ-omegn om a innenfor definisjonsmengden til f. Betingelsen $|f(x) - f(a)| < \epsilon$ kan skrives som $f(x) \in (f(a) - \epsilon, f(a) + \epsilon)$, som er en ϵ-omegn om $f(a)$. Hele betingelsen (208) kan derfor også uttrykkes som

$$f(x) \in (f(a) - \epsilon, f(a) + \epsilon) \text{ når } x \in (a - \delta, a + \delta) \cap D(f), \tag{209}$$

eller, med notasjonen fra Definisjon 2.2.4:

$$f\left((a - \delta, a + \delta) \cap D(f)\right) \subset (f(a) - \epsilon, f(a) + \epsilon), \tag{210}$$

og visualiseres i følgende figur:



$f(x)$ ligger i intervallet $(f(a) - \epsilon, f(a) + \epsilon)$ såfremt x ligger i intervallet $(a - \delta, a + \delta)$.

Ved å betrakte (210) ser vi at vi kan uttrykke den formelle definisjonen for kontinuitet som at

f er kontinuerlig i a hvis og bare hvis det til enhver omegn V om $f(a)$ finnes en omegn om a som f avbilder inn i V.

Dette er en beskrivelse av kontinuitet uten bruk av greske bokstaver og absoluttverdier (jf. Oppgave 9.5). For enda en ekvivalent formulering, se Setning 9.4.1 nedenfor.

Figuren ovenfor viser samspillet mellom ϵ og δ: uansett hvor liten ϵ vi får oppgitt, skal vi klare å finne en $\delta > 0$ slik at hele grafen til funksjonen over intervallet $(a-\delta, a+\delta)$ ligger mellom de to vannrette linjene $y = f(a) + \epsilon$ og $y = f(a) - \epsilon$. Vi tenker ofte på ϵ som en feilmargin som anslår hvor mye forskjellen mellom $f(x)$ og den ønskede verdien $f(a)$ er. Følgende analogi kan være til hjelp: Om du har en maskin du har laget for en kunde, som skal gi ut en bestemt verdi $f(a)$ bare kunden putter inn a, så er den ganske verdiløs om den gir helt andre verdier om kunden gir a med en ørliten feil. Det du i virkeligheten må garantere er at for enhver feilmargin (gitt ved ϵ) kunden kan tolerere, så finnes det en nøyaktighet ved input (gitt ved δ) som garanterer denne nøyaktigheten i output.

La oss gå tilbake til eksemplet vi har sett på, for å se hvordan vi i få ord *fører inn* løsningen.

EKSEMPEL 9.1.4. La $f : \mathbb{R} \to \mathbb{R}$ være definert ved $f(x) = 2x - 1$. Vis at f er kontinuerlig i 3.

LØSNING. Gitt $\epsilon > 0$. La $\delta = \frac{\epsilon}{2}$. Da, hvis $|x-3| < \delta = \frac{\epsilon}{2}$, er

$$|f(x) - f(3)| = |2x - 1 - 5| = |2x - 6| = 2|x-3| < 2 \cdot \frac{\epsilon}{2} = \epsilon,$$

som vi skulle bevise. □

Vi merker oss at løsningen er temmelig kort og starter med å angi δ'en. For å finne δ som vi gjorde over startet vi imidlertid med uttrykket $|f(x) - f(3)|$. Dette viser igjen at de "pene" løsningene vi skriver (og som vi finner i lærebøker) ikke helt viser den naturlige fremgangsmåten i slike typer bevis, som grovt sagt kan sies å være:

Jobb med uttrykket $|f(x) - f(a)|$ til du får ut uttrykket $|x - a|$.

Vi skal i neste seksjon se på flere eksempler på bruk av definisjonen.

Vi bemerker også at definisjonen av kontinuitet fungerer like bra for funksjoner med definisjonsmengde og verdimengde i $\mathbb{C}$, siden $\mathbb{C}$ har et avstandsbegrep mellom punkter på lik linje med $\mathbb{R}$ (jf. Merknad 8.2.2). Vi skriver ned definisjonen for ordens skyld og nøyer oss videre å legge inn merknader når resultater og eksempler som vi gir over $\mathbb{R}$ også fungerer over $\mathbb{C}$.

Definisjon 9.1.5: Kontinuitet i et punkt (komplekse funksjoner)

La $f : D(f) \to \mathbb{C}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{C}$ og la $a \in D(f)$. Funksjonen f er *kontinuerlig* i a dersom det til ethvert reelt tall $\epsilon > 0$ finnes et reelt tall $\delta > 0$ slik at

$$|f(z) - f(a)| < \epsilon \;\text{ når }\; |z - a| < \delta \;\text{ og }\; z \in D(f). \tag{211}$$

9.2. Flere eksempler på bruk av definisjonen

Det skal ikke mye til for å innse at teknikken brukt i Eksempel 9.1.4 fungerer like godt for alle *lineære funksjoner* og for alle punkter i definisjonsmengden:

Eksempel 9.2.1. La $f : \mathbb{R} \to \mathbb{R}$ være definert ved $f(x) = Ax + B$, for konstanter $A, B \in \mathbb{R}$. Vis at f er kontinuerlig i alle $a \in \mathbb{R}$.

Løsning (lang, resonnerende versjon). Vi skal, til enhver $\epsilon > 0$, finne en $\delta > 0$ slik at

$$|f(x) - f(a)| < \epsilon \;\text{ når }\; |x - a| < \delta.$$

Vi bruker prosedyren antydet ovenfor og starter med å omforme uttrykket $|f(x) - f(a)|$:

$$|f(x) - f(a)| = |(Ax + B) - (Aa + B)| = |A(x - a)| = |A||x - a|.$$

At dette er $< \epsilon$ er alltid oppfylt for $A = 0$. Hvis $A \neq 0$, ser vi at

$$|A||x - a| < \epsilon \iff |x - a| < \frac{\epsilon}{|A|}.$$

Dette sier oss "hva vi må gjøre $|x - a|$ mindre enn for å sørge for at $|f(x) - f(a)| < \epsilon$", nemlig $\frac{\epsilon}{|A|}$, som vi da tar som vår δ hvis $A \neq 0$, mens en vilkårlig $\delta > 0$ fungerer hvis $A = 0$. $\square$

Løsning ("Lærebokversion"). Gitt $\epsilon > 0$. Hvis $A \neq 0$, la $\delta = \frac{\epsilon}{|A|}$, mens hvis $A = 0$, la δ være et vilkårlig positivt reelt tall. Da, hvis $|x-a| < \delta$, er i begge tilfeller $|A||x - a| < \epsilon$, slik at

$$|f(x) - f(a)| = |(Ax + B) - (Aa + B)| = |A(x - a)| = |A||x - a| < \epsilon,$$

som vi skulle bevise. $\square$

Merknad 9.2.2. Som vi ser i tilfellet $A \neq 0$ av siste eksempel, og som vi allerede så i Eksempel 9.1.4, vil tallet δ generelt være avhengig av ϵ: jo mindre ϵ er, desto mindre må δ være. Dette betyr at x må nærmere a for at $f(x)$ skal komme nærmere $f(a)$. (Dette er et fenomen vi allerede har støtt på når vi har studert konvergens av følger, jf. Merknad 6.2.8.)

Avhengigheten av ϵ gjør at noen matematiske tekster betegner δ i Definisjon 9.1.1 med $\delta(\epsilon)$. Avhengigheten betyr imidlertid ikke at δ *må* variere når ϵ varierer: i tilfellet $A = 0$ siste eksempel, der $f(x) = B$, det vil si en konstant funksjon, har vi for eksempel at en hvilken som helst $\delta \in \mathbb{R}^+$ fungerer for alle ϵ.

Vi ser på noen eksempler til på bruk av Definisjon 9.1.1:

EKSEMPEL 9.2.3. La $f : \mathbb{R} \to \mathbb{R}$ være absoluttverdifunksjonen definert ved $f(x) = |x|$. Vis at f er kontinuerlig i alle $a \in \mathbb{R}$.

LØSNING (LANG, RESONNERENDE VERSJON). Vi skal, til enhver $\epsilon > 0$, finne en $\delta > 0$ slik at

$$|f(x) - f(a)| = ||x| - |a|| < \epsilon \text{ når } |x - a| < \delta.$$

Ved Oppgave 3.19 har vi (som konsekvens av trekantulikheten):

$$||x| - |a|| \leq |x - a|. \tag{212}$$

At dette er $< \epsilon$ er alltid oppfylt nå $|x-a| < \epsilon$. Vi kan derfor la $\delta = \epsilon$. □

LØSNING ("LÆREBOKVERSJON"). La $a \in \mathbb{R}$. Gitt $\epsilon > 0$. La $\delta = \epsilon$. Da, hvis $|x - a| < \delta$, er

$$|f(x) - f(a)| = ||x| - |a|| \leq |x - a| < \epsilon,$$

som viser at f er kontinuerlig i $a \in \mathbb{R}$. □

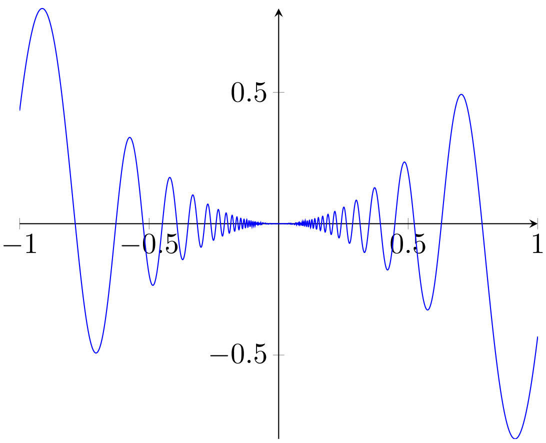
MERKNAD 9.2.4. De to siste eksemplene fungerer ordrett[1] like godt om vi betrakter $f : \mathbb{C} \to \mathbb{C}$ definert med samme uttrykk. Det sentrale er at *trekantulikheten* (Setning 8.2.1), samt konsekvensen (212) (fra Oppgave 8.11), også gjelder i $\mathbb{C}$.

EKSEMPEL 9.2.5. Vis at funksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & x \neq 0 \\ 0, & x = 0 \end{cases}$$

er kontinuerlig i 0.

[1] I matematikk brukes noen ganger det latinske uttrykket *ad verbatim* aller bare *verbatim* for "ordrett". Hvis imidlertid en tekst, for eksempel et bevis, fungerer nesten ordrett med kun "de nødvendige endringene", brukes det latinske uttrykket *mutatis mutandis*, som betyr "endret det som må endres".



Grafen til f i Eksempel 9.2.5

Løsning (lang, resonnerende versjon). Vi skal, til enhver $\epsilon > 0$, finne en $\delta > 0$ slik at

$$|f(x) - f(0)| = |f(x)| < \epsilon \text{ når } |x - 0| = |x| < \delta.$$

Dette er opplagt oppfylt for $x = 0$, siden $f(0) = 0$. Hvis $x \neq 0$, har vi:

$$|f(x) - f(0)| = \left|x^2 \sin\left(\frac{1}{x}\right) - 0\right| = |x|^2 \left|\sin\left(\frac{1}{x}\right)\right|, \quad x \neq 0.$$

Dette ønsker vi å få $< \epsilon$, uansett $\epsilon > 0$, ved å velge $|x - 0| = |x|$ liten nok. Vi har lyktes med, som i de forrige eksemplene, å faktorisere ut $|x - 0| = |x|$, men vi har også faktoren $\left|\sin\left(\frac{1}{x}\right)\right|$ denne gangen. Når vi får slike "tilleggsfaktorer", så må vi prøve å begrense disse også. I dette tilfellet vet vi at sinusfunksjonen kun tar verdier mellom -1 og 1, og dermed vet vi at $\left|\sin\left(\frac{1}{x}\right)\right| \leq 1$ for $x \neq 0$. Da får vi

$$|f(x) - f(0)| = \left|x^2 \sin\left(\frac{1}{x}\right)\right| = |x|^2 \left|\sin\left(\frac{1}{x}\right)\right| \leq |x|^2 \cdot 1 = |x|^2, \quad x \neq 0.$$

Vi har:

$$|x|^2 < \epsilon \iff |x| < \sqrt{\epsilon}.$$

Vi har dermed funnet at så lenge $|x| < \sqrt{\epsilon}$, da er $|f(x) - 0| < \epsilon$, og vi har funnet vår δ. □

Løsning ("Lærebokversjon"). Gitt $\epsilon > 0$. La $\delta = \sqrt{\epsilon}$. Da, hvis $|x - 0| = |x| < \delta = \sqrt{\epsilon}$, er

$$\left|x^2 \sin\left(\frac{1}{x}\right) - 0\right| = |x|^2 \left|\sin\left(\frac{1}{x}\right)\right| \leq |x|^2 \cdot 1 = |x|^2 < \sqrt{\epsilon}^2 = \epsilon,$$

som vi skulle bevise. □

Merk hvordan vi i den innførte løsningen tar med kjeden av likheter og ulikheter som gjør det klart for leseren hvordan vi bruker $|x-a|<\delta$ til å konkludere at $|f(x)-f(a)|<\epsilon$.

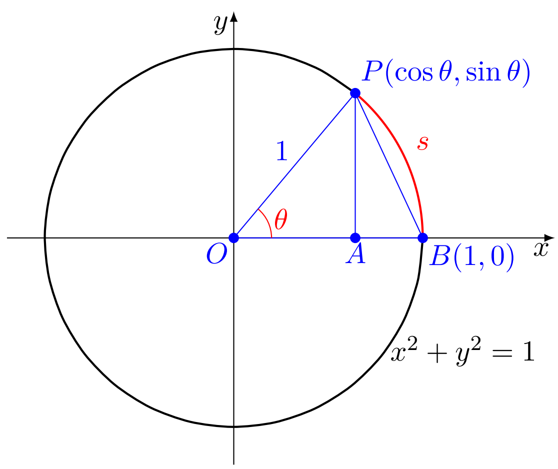
Dessverre er det ikke alltid slik at δ'en faller ut så lett som i eksemplene ovenfor, som de tre neste eksemplene viser:

EKSEMPEL 9.2.6. Vis at funksjonen $f:\mathbb{R}\to\mathbb{R}$ gitt ved $f(x)=\sin x$ er kontinuerlig i 0.

LØSNING (LANG, RESONNERENDE VERSJON). Vi skal, til enhver $\epsilon>0$, finne en $\delta>0$ slik at

$$|f(x)-f(0)|=|\sin x-\sin 0|=|\sin x|<\epsilon \text{ når } |x-0|=|x|<\delta.$$

Som i de forrige eksemplene, kunne vi ønsket oss at $|f(x)-f(0)|$ inneholdt uttrykket $|x-0|=|x|$, men det gjorde det ikke denne gangen. Vi vil bruke trigonometriske betraktninger til å konkludere at $|\sin x|\leq|x|$ for liten x. Betrakt figuren:



hvor sirkelen er en enhetssirkel. Da er per definisjon buelengden $s=\theta$. Vi ser da at

$$0\leq\sin\theta=PA\leq PB\leq s=\theta \text{ for } \theta\in\left[0,\frac{\pi}{2}\right).$$

Samme figur med trekantene OAP og OPB speilet om x-aksen gir at

$$0\leq-\sin(\theta)\leq-\theta \text{ for } \theta\in\left(-\frac{\pi}{2},0\right].$$

Vi konkluderer at

$$|\sin\theta|\leq|\theta| \text{ for alle } \theta\in\left(-\frac{\pi}{2},\frac{\pi}{2}\right).$$

Hvis $|x|<\frac{\pi}{2}$, har vi derfor

$$|f(x)-f(0)|=|\sin x|\leq|x|.$$

Denne ulikheten er imidlertid oppfylt for alle $x \in \mathbb{R}$, siden vi alltid har $|\sin x| \leq 1$. Dette betyr at vi kan sørge for at $|f(x) - f(0)| < \epsilon$ dersom $|x| < \epsilon$. Vi har derfor funnet at $\delta = \epsilon$ gjør jobben. □

LØSNING ("LÆREBOKVERSJON"). Gitt $\epsilon > 0$. La $\delta = \epsilon$. Ved trigonometriske betraktninger (som ovenfor) har vi $|\sin x| \leq |x|$. Derfor, hvis $|x| < \delta$, er

$$|\sin x - \sin 0| = |\sin x| \leq |x| < \epsilon,$$

som viser at $\sin x$ er kontinuerlig i 0. □

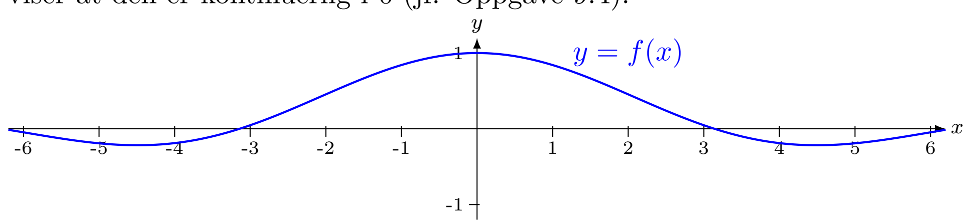
MERKNAD 9.2.7. Vi merker for senere bruk at vi i det første beviset ovenfor viste at $|\sin x| \leq |x|$ gjelder for alle $x \in \mathbb{R}$.

MERKNAD 9.2.8. En enkel modifisering av beviset viser at sinusfunksjonen $\sin : \mathbb{R} \to \mathbb{R}$ er kontinuerlig i alle punkter (jf. Oppgave 9.3(b)).

MERKNAD 9.2.9. Funksjonen $f : \mathbb{R} \to \mathbb{R}$ definert ved

$$f(x) = \begin{cases} \frac{\sin x}{x}, & x \neq 0, \\ 1, & x = 0 \end{cases}$$

er en "klassiker" i mange matematiske tekster. En videreføring av siste bevis viser at den er kontinuerlig i 0 (jf. Oppgave 9.4).



GRAFEN TIL f I MERKNAD 9.2.9

MERKNAD 9.2.10. Det kan kanskje føles litt "ubehagelig" at definisjonen på sinusfunksjonen (og de andre trigonometriske funksjonene), samt beviset for dens kontinuitet, avhenger av begrepet *buelengde*, som trenger en grenseprosess for å kunne defineres. Man kan unngå dette problemet ved å definere sinusfunksjonen (og cosinusfunksjonen) ved hjelp av rekker, som antydet i Merknad 10.9.7.

EKSEMPEL 9.2.11. Vis at funksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved $f(x) = x^2 + x$ er kontinuerlig i -2.

LØSNING (LANG, RESONNERENDE VERSJON). Vi skal, til enhver $\epsilon > 0$, finne en $\delta > 0$ slik at

$$|f(x) - f(-2)| = |(x^2 + x) - 2| < \epsilon \text{ når } |x - (-2)| < \delta,$$

med andre ord

$$|x^2 + x - 2| < \epsilon \text{ når } |x + 2| < \delta.$$

Vi starter som vanlig med uttrykket $|x^2+x-2|$ og forsøker å begrense dette ved å begrense $|x+2|$. Det er altså faktoren $|x+2|$ vi ønsker å "få ut" i dette tilfellet. Vi merker at -2 er rot i $x^2 + x - 2$, slik at $x + 2$ er en faktor ved *faktorteoremet* (Setning 4.1.9). Vi finner lett ut at $x^2+x-2 = (x+2)(x-1)$. Derfor har vi omformet uttrykket:

$$|x^2 + x - 2| = |x + 2||x - 1|,$$

og den ene faktoren, nemlig $|x+2|$, er den vi kan begrense så mye vi vil ved å velge δ liten nok, så den liker vi!

Hva så med faktoren $|x - 1|$? Kan vi også begrense den? Vel, hvis vi gjør $|x + 2|$ liten, så betyr det at x er "nær -2", og $|x - 1|$ er jo avstanden mellom x og 1. Og hvis x er "nær -2", så er den ikke så langt unna 1 heller! For eksempel, anta

$$|x + 2| < 1 \tag{213}$$

Trekantulikheten gir

$$|x - 1| = |x + 2 - 3| \leq |x + 2| + | - 3| = |x + 2| + 3 < 4.$$

Under forutsetningen (213), gjelder altså:

$$|x^2 + x - 2| = |x + 2||x - 1| < |x + 2| \cdot 4 = 4|x + 2|,$$

og vi ser at dette siste er $< \epsilon$ hvis og bare hvis

$$|x + 2| < \frac{\epsilon}{4}. \tag{214}$$

Altså har vi vist at den ønskede ulikheten $|x^2 + x - 2| < \epsilon$ holder dersom *både* (213) *og* (214) *holder*. Vi har denne gangen *to* betingelser på uttrykket $|x-a| = |x-(-2)| = |x+2|$ istedenfor bare *én*, som tidligere. Vi må derfor velge $|x + 2|$ til å være både mindre enn 1 og $\frac{\epsilon}{4}$. Dette er oppfylt når

$$|x + 2| < \min\left\{1, \frac{\epsilon}{4}\right\},$$

hvor vi minner om at uttrykket min betyr "det minste av tallene". Vi har derfor funnet en δ som gjør jobben, nemlig $\min\{1, \frac{\epsilon}{4}\}$, det minste av 1 og $\frac{\epsilon}{4}$ (dersom for eksempel $\epsilon = 5$, er $\delta = \min\{1, \frac{5}{4}\} = 1$, og dersom for eksempel $\epsilon = 1$, er $\delta = \min\{1, \frac{1}{4}\} = \frac{1}{4}$). □

Når vi nå skriver en pen, oversiktlig løsning, må vi passe på å gjøre det klart for leseren hvordan vi bruker de *to* ulikhetene (213) og (214).

Løsning ("Lærebokversjon"). Gitt $\epsilon > 0$. La $\delta = \min\{1, \frac{\epsilon}{4}\}$. Da, hvis $|x-(-2)| = |x+2| < \delta$, er for det første

$$|x+2| < \frac{\epsilon}{4}. \tag{215}$$

Dessuten er $|x+2| < 1$, som gir

$$|x-1| = |x+2-3| \leq |x+2| + 3 < 4. \tag{216}$$

Dermed vil

$$|(x^2+x)-2| = |x^2+x-2| = |x+2||x-1| \overset{(*)}{<} \frac{\epsilon}{4} \cdot 4 = \epsilon$$

(hvor vi har brukt (215) og (216) i overgangen $(*)$), som vi skulle bevise. □

Vi merker at valget å gjøre $|x+2| < 1$ i (213) egentlig var litt tilfeldig. Vi måtte bare starte med å gjøre det mindre enn *noe* og tok dette *noe* til å være 1. Vi kunne også byttet ut (213) med for eksempel $|x+2| < 2$, som ville gitt $|x+1| < 5$ istedenfor (214), som ville ført til at vi måtte velge (etter lignende utregninger) $\delta = \min\{2, \frac{\epsilon}{5}\}$. Dette svaret er like riktig som det forrige. Vi viser den pene føringen av dette svaret også:

Alternativ løsning ("Lærebokversjon"). La $\epsilon > 0$ være gitt. Sett $\delta = \min\{2, \frac{\epsilon}{5}\}$. Da, hvis $0 < |x-(-2)| = |x+2| < \delta$, er for det første

$$|x+2| < \frac{\epsilon}{5}. \tag{217}$$

Dessuten er $|x+2| < 2$, som gir

$$|x-1| = |x+2-3| \leq |x+2| + 3 < 5. \tag{218}$$

Dermed vil

$$|(x^2+x)-2| = |x^2+x-2| = |x+2||x-1| \overset{(*)}{<} \frac{\epsilon}{5} \cdot 5 = \epsilon$$

(hvor vi har brukt (217) og (218) i overgangen $(*)$), som vi skulle bevise. □

Merknad 9.2.12. Eksempel 9.2.11 fungerer like godt og ordrett om vi betrakter $f : \mathbb{C} \to \mathbb{C}$ gitt ved samme uttrykk.

Eksempel 9.2.13. Vis at funksjonen $f : \mathbb{R} \setminus \{1\} \to \mathbb{R}$ definert ved $f(x) = \frac{1}{(x-1)^2}$ er kontinuerlig i 2.

Løsning (lang, resonnerende versjon). Vi skal, gitt $\epsilon > 0$, finne en $\delta > 0$ slik at

$$|f(x) - f(2)| < \epsilon \text{ når } |x-2| < \delta,\ x \in \mathbb{R} \setminus \{1\},$$

med andre ord

$$\left|\frac{1}{(x-1)^2} - 1\right| < \epsilon \text{ når } |x-2| < \delta,\ x \neq 1.$$

Vi starter som vanlig med uttrykket $\left|\frac{1}{(x-1)^2} - 2\right|$ og omformer dette:

$$\begin{aligned}\left|\frac{1}{(x-1)^2} - 1\right| &= \left|\frac{1-(x-1)^2}{(x-1)^2}\right| = \left|\frac{1-x^2+2x-1}{(x-1)^2}\right| = \left|\frac{-x^2+2x}{(x-1)^2}\right| \\ &= \frac{|x||x-2|}{(x-1)^2} = \frac{|x|}{(x-1)^2} \cdot |x-2|.\end{aligned}$$

Faktoren $|x-2|$ kan vi begrense direkte ved hjelp av δ. Nå må vi begrense $\frac{|x|}{(x-1)^2}$. Dette betyr at vi må begrense $|x|$ *oppad*, mens vi må begrense $|x-1|$ *nedad*. Dette er altså en litt annerledes situasjon enn i forrige eksempel. Som vi har gjort tidligere, forsøker vi med å begrense $|x-2|$ som

$$|x-2| < 1. \tag{219}$$

For å begrense $|x-1|$ *nedad* bruker vi at dette er ekvivalent med

$$-1 < x-2 < 1 \Longleftrightarrow 0 < x-1 < 2$$

(addér 1 overalt). Dette sier oss at $0 < |x-1| < 2$, som ikke gir oss noen øvre begrensning på $\frac{1}{|x-1|}$. Vi klarer altså ikke å si at $\frac{|x|}{(x-1)^2}$ er "mindre enn noe". Problemet er at begrensningen i (219) denne gangen er for grov: den garanterer bare at $x \in (1,3)$, som gjør at $|x-1|$ kan komme vilkårlig nær 0. Vi må derfor begrense $|x-2|$ mer og prøver med

$$|x-2| < \frac{1}{2}. \tag{220}$$

Dette er ekvivalent med:

$$-\frac{1}{2} < x-2 < \frac{1}{2} \Longleftrightarrow \frac{1}{2} < x-1 < \frac{3}{2}$$

som gir den ønskede begrensningen nedad på $|x-1| > \frac{1}{2}$, eller ekvivalent $\frac{1}{|x-1|} < \frac{1}{\frac{1}{2}} = 2$.

Samtidig har vi at (220) er ekvivalent med

$$-\frac{1}{2} < x-2 < \frac{1}{2} \Longleftrightarrow \frac{3}{2} < x < \frac{5}{2}$$

som gir den ønskede begrensningen $|x| < \frac{5}{2}$. Når (220) holder, da vil altså

$$\frac{|x|}{(x-1)^2} = |x| \cdot \left(\frac{1}{|x-1|}\right)^2 < \frac{5}{2} \cdot 2^2 = 10.$$

Dette gir at

$$\left|\frac{1}{(x-1)^2}-1\right| = \frac{|x|}{(x-1)^2}\cdot|x-2| < 10|x-2|,$$

og dette siste er $< \epsilon$ hvis og bare hvis

$$|x-2| < \frac{\epsilon}{10}. \tag{221}$$

Altså har vi vist at den ønskede ulikheten $\left|\frac{1}{(x-1)^2}-1\right| < \epsilon$ er oppfylt dersom *både* (220) *og* (221) *holder*, og disse to er oppfylt som vanlig når

$$|x-2| < \min\left\{\frac{1}{2}, \frac{\epsilon}{10}\right\},$$

og vi har funnet at $\delta = \min\left\{\frac{1}{2}, \frac{\epsilon}{10}\right\}$ gjør jobben. □

Vi viser igjen den pene oversiktlige løsningen. For å variere, bruker vi trekantulikheten for å dedusere begrensningene på $|x|$ og $|x-1|$ gitt begrensningen på $|x-2|$. Begrensningen på $|x|$ er rett frem som tidligere, mens vi for å begrense $|x-1|$ nedad bruker konsekvensen av trekantulikheten gitt i Oppgave 3.19: den sier nemlig at $|a-b| \geq ||a|-|b||$ for alle $a, b \in \mathbb{R}$, og bruker vi den på $a = x-2$ og $b = -1$, får vi

$$|x-1| = |(x-2)-(-1)| \geq ||x-2|-|-1|| = ||x-2|-1|.$$

Løsning ("Lærebokversjon"). Gitt $\epsilon > 0$. La $\delta = \min\{\frac{1}{2}, \frac{\epsilon}{10}\}$. Da, hvis $|x-2| < \delta$, er for det første

$$|x-2| < \frac{\epsilon}{10}. \tag{222}$$

Dessuten er $|x-2| < \frac{1}{2}$, som ved trekantulikheten gir

$$|x| = |x-2+2| \leq |x-2| + 2 < \frac{5}{2}$$

og

$$|x-1| = |(x-2)-(-1)| \geq ||x-2|-1| > \left|\frac{1}{2}-1\right| = \frac{1}{2}. \tag{223}$$

Dermed vil

$$\left|\frac{1}{(x-1)^2}-1\right| = \frac{|x|}{(x-1)^2}\cdot|x-2| \overset{(*)}{<} \frac{\frac{5}{2}}{\left(\frac{1}{2}\right)^2}\cdot\frac{\epsilon}{10} = \epsilon,$$

(hvor vi har brukt (222)-(223) i overgangen $(*)$), som vi skulle bevise. □

Merknad 9.2.14. Eksempel 9.2.13, utledet som i "lærebokversjonen", fungerer like godt og ordrett om vi betrakter $f : \mathbb{C}\setminus\{1\} \to \mathbb{C}$ gitt ved samme uttrykk.

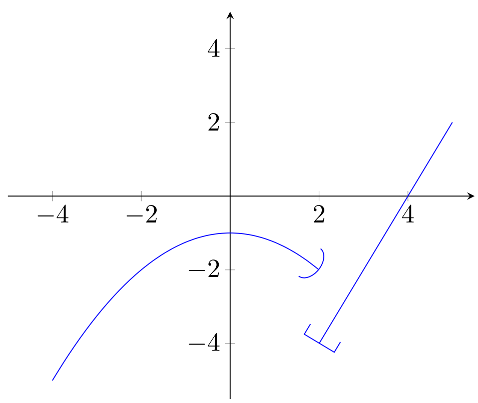
Diskontinuiteter. La oss også se på et eksempel på en diskontinuitet, av en type dere sikkert husker fra skolen.

Eksempel 9.2.15. Vis at funksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved

$$f(x) = \begin{cases} -\frac{1}{4}x^2 - 1, & x < 2, \\ 2x - 8, & x \geq 2, \end{cases}$$

er diskontinuerlig i 2.

Løsning (lang, resonnerende versjon). La oss bemerke at grafen til f ser slik ut:



Grafen til f i Eksempel 9.2.15.

Merk at $f(2) = 2 \cdot 2 - 8 = -4$. Grafisk (og intuitivt) ser vi hvorfor f er diskontinuerlig i 2: om x er veldig nær 2, med $x < 2$, så vil alltid $f(x) > -2$, slik at $|f(x) - f(2)| = |f(x) - (-4)| = |f(x) + 4| > 2$ og kan derfor ikke fås så liten vi vil. Krever vi for eksempel $|f(x) + 4| < 1$, ser vi at

$$\begin{array}{ccccccc} & & |f(x)+4| & < & 1 & \text{for} & x < 2 \\ & & \Updownarrow & & & & \\ & & \left|-\frac{1}{4}x^2 - 1 + 4\right| & < & 1 & \text{for} & x < 2 \\ & & \Updownarrow & & & & \\ & & \left|\frac{1}{4}x^2 - 3\right| & < & 1 & \text{for} & x < 2 \\ & & \Updownarrow & & & & \\ -1 & < & \frac{1}{4}x^2 - 3 & < & 1 & \text{for} & x < 2 \\ & & \Updownarrow & & & & \\ 2 & < & \frac{1}{4}x^2 & < & 4 & \text{for} & x < 2 \\ & & \Updownarrow & & & & \\ 8 & < & x^2 & < & 16 & \text{for} & x < 2 \end{array}$$

Men dette kan aldri være oppfylt for noen $x \geq 0$ (mer spesifikt for noen $x \geq -\sqrt{8}$). Vi ser altså at kriteriet for kontinuitet ikke er oppfylt for $\epsilon = 1$ og er dermed heller ikke oppfylt for noen $\epsilon \leq 1$. Et mer nøyaktig studium viser at kriteriet for kontinuitet ikke er oppfylt for $\epsilon \leq 2$. $\square$

Løsning ("Lærebokversjon"). Anta at f er kontinuerlig i 2. Da vil det finnes en $\delta > 0$ slik at $|f(x) - f(2)| = |f(x) - (-4)| = |f(x) + 4| < 1$ for alle $x \in (2-\delta, 2+\delta)$. Spesielt vil $\left|-\frac{1}{4}x^2 - 1 + 4\right| < 1$ for alle $x \in (2-\delta, 2)$. Dette betyr at

$$-1 < \frac{1}{4}x^2 - 3 < 1 \quad \text{for} \quad x \in (2-\delta, 2),$$

eller omskrevet

$$8 < x^2 < 16 \quad \text{for} \quad x \in (2-\delta, 2),$$

som er umulig uansett δ, siden $x^2 < 4$ for $0 < x < 2$. □

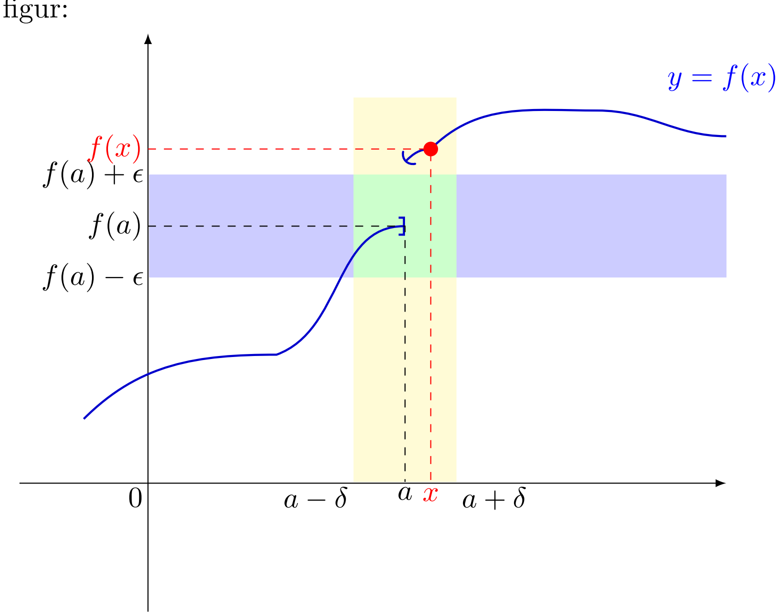
Det kan være verdt å merke seg teknikken vi brukte i forrige eksempel for å vise at funksjonen var diskontinuerlig: vi fant nemlig en $\epsilon > 0$ hvor det ikke var mulig å finne en tilhørende $\delta > 0$ slik at definisjonen av kontinuitet var oppfylt. Ser vi på definisjonen av kontinuitet, så får vi altså også ut et kriterium for diskontinuitet:

Observasjon 9.2.16: Diskontinuitet i et punkt

La $f : D(f) \to \mathbb{R}$ være en funksjon, med definisjonsmengde $D(f) \subset \mathbb{R}$, og la $a \in D(f)$. Funksjonen f er *diskontinuerlig* i a dersom det finnes et reelt tall $\epsilon > 0$ slik at uansett $\delta > 0$ finnes $x \in D(f)$ slik at

$$|x - a| < \delta \quad \text{og} \quad |f(x) - f(a)| \geq \epsilon.$$

Dette betyr at tallet ϵ er en feilmargin vi ikke kan håndtere: uansett hvor liten vi velger slingringsmonnet δ, klarer vi ikke å garantere at funksjonsverdien $f(x)$ ligger innenfor denne feilmarginen. Dette kan visualiseres i følgende figur:



$f(x)$ ligger ikke i $(f(a) - \epsilon, f(a) + \epsilon)$ for alle x i $(a - \delta, a + \delta)$, uansett hvor liten δ er.

Merknad 9.2.17. Observasjon 9.2.16 gjelder også om vi bytter ut $\mathbb{R}$ med $\mathbb{C}$.

9.3. Nødvendigheten av den formelle definisjonen

Nå er det naturlig å stanse et øyeblikk og diskutere følgende spørsmål:

(1) Den formelle ϵ–δ definisjonen av grenseverdi virker kanskje tung og unødvendig, siden vi rent intuitivt forstår hva kontinuitet er. Kunne vi ikke klart oss med en mer uformell definisjon, som for eksempel ved å si at *f er kontinuerlig i a dersom vi kan få $f(x)$ så nær vi vil $f(a)$ ved å la x være nær nok a*? Eller ved å si at *f er kontinuerlig i a dersom den har sammenhengende graf*?

(2) Gitt at vi trenger en slik formell definisjon, så virker den relativt tungvinn å bruke selv for enkle funksjoner som i Eksempel 9.2.11 i ett bestemt punkt. Siden det finnes utrolig mange funksjoner i matematikk (og i naturvitenskap generelt), må vi virkelig sjekke kontinuitet i ethvert punkt i definisjonsmengden ved hjelp av ϵ–δ definisjonen?

La oss ta svaret på (2) først, siden det er gode nyheter: Svaret er "Nei, vi må heldigvis ikke sjekke kontinuitet i ethvert punkt i definisjonsmengden til enhver funksjon ved hjelp av ϵ–δ definisjonen". Som i de fleste tilfeller i matematikk utleder vi generelle resultater basert på definisjonen som vi så kan bruke for å konkludere kontinuitet i spesielle tilfeller. Strategien går ut på å vise kontinuitet ved hjelp av den formelle ϵ–δ definisjonen for en del grunnleggende funksjoner, og så vise at dersom vi kombinerer disse funksjonene (ved hjelp av de vanlige aritmetiske operasjonene og sammensetninger), så vil de resulterende funksjonene fortsatt være kontinuerlige. Vi skal se på disse mer generelle resultatene og konsekvensene i §9.5.

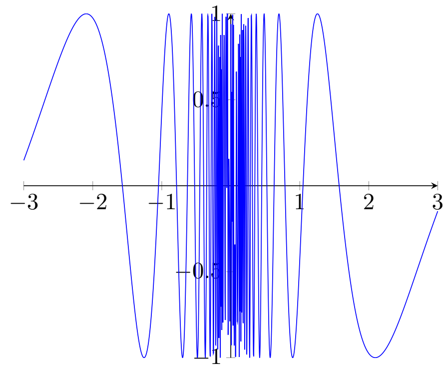
Så til spørsmål (1). Her er svaret "Nei, vi klarer oss ikke med en uformell definisjon". Grunnen til at vi kan bli villedet til å tro det er at vi fra skolen er vant til funksjoner bygd opp av pene uttrykk som polynomer, trigonometriske funksjoner, eksponensialfunksjoner (og et endelig antall operasjoner som addisjon, subtraksjon, multiplikasjon, derivasjon og rotuttrykk). Funksjoner som er definert på denne måten har meget pene egenskaper og den uformelle måten å forstå kontinuitet på som i (204)–(205) fungerer greit i slike tilfeller. Imidlertid er de fleste funksjonene som man møter på i matematikk og naturvitenskap langt mer kompliserte (se Eksemplene 9.3.2 og 9.3.3 nedenfor, for ikke å snakke om Weierstrass-funksjonen vi snakket om i §1.4), der den uformelle definisjonen nevnt ovenfor ikke er presis nok. Og det er faktisk heller ikke så enkelt å gi en skikkelig definisjon på hva det vil si at en graf er sammenhengende heller (se Definisjon 10.5.1 for en rigorøs definisjon av begrepet *sammenhengende mengde*). Vi skal i Teorem 9.6.1 gi og bevise et resultat som viser at grafen til en kontinuerlig funksjon er sammenhengende i en mulig betydning av ordet. Uansett er en vesentlig ulempe ved å basere kontinuitet på sammenhengsegenskapen at det da er vanskelig å definere kontinuitet i et enkeltstående punkt.

La oss se på noen spesielle funksjoner hvor en uformell forståelse av kontinuitet ikke vil fungere:

EKSEMPEL 9.3.1. Betrakt funksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved

$$f(x) = \begin{cases} \sin\left(\frac{1}{x}\right), & x \neq 0, \\ c, & x = 0, \end{cases}$$

hvor $c \in \mathbb{R}$ er et vilkårlig fiksert tall. Grafen (utenom $x = 0$) ser slik ut:



GRAFEN TIL FUNKSJONEN I EKSEMPEL 9.3.1 FOR $x \neq 0$.

Når x går mot 0, virker grafen å gå mot alle tall mellom -1 og 1 på y-aksen samtidig. Hvis $c \in [-1, 1]$ kommer med andre ord funksjonsverdiene $f(x)$ så nær vi vil tallet c når x er nær nok 0. Den uformelle måten å forstå kontinuitet på som i (204) og (205) virker derfor å antyde at f kan være kontinuerlig i 0 når $c \in [-1, 1]$. På den annen side, selv om $f(x)$ kommer så nær vi vil c, så vil den "svinge bort" fra c hvis x gjøres mindre (for så å igjen komme så nær vi vil c). Altså "hopper grafen" bort igjen, hvilket rent intuitivt kan tyde på en diskontinuitet. Som oppsummering: uansett $c \in [-1, 1]$, vil vi klare å sørge for at $f(x)$ er så nær vi vil c når x er liten nok, samtidig som at $f(x)$ vil "svinge bort igjen" fra c hvis x gjøres mindre. Det er uklart om f er kontinuerlig i 0 hvis vi ikke har en presis nok definisjon.

Etter at vi har definert rigorøst hva vi mener med en *sammenhengende mengde* (Definisjon 10.5.1), skal vi faktisk vise at grafen til f er sammenhengende for $c \in [-1, 1]$ (jf. Merknad 10.5.10 og Oppgave 10.45(a)).

Den formelle definisjonen kan imidlertid brukes til å avgjøre at f *er diskontinuerlig i* 0 uansett c. Man kan argumentere ved hjelp av Observasjon 9.2.16. (Beviset for dette er nok hakket vanskeligere enn eksemplene før, så det kan trygt hoppes over i første omgang og heller leses etter at man har opparbeidet seg en viss kontroll over stoffet.)

FORMELT BEVIS FOR AT f ER DISKONTINUERLIG I 0. Vi bruker Observasjon 9.2.16 med $\epsilon = 1$ og viser at uansett $\delta > 0$, så vil det finnes x slik at

$|x| < \delta$ og

$$\left|\sin\left(\frac{1}{x}\right) - c\right| \geq 1 \;\text{ og }\; 0 < |x| < \delta.$$

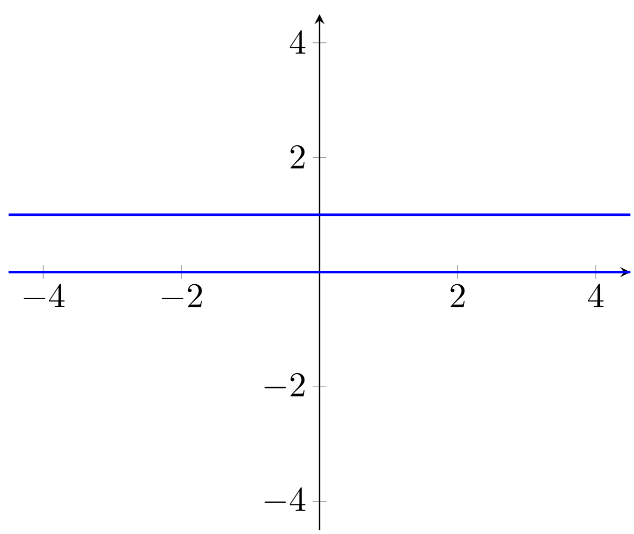
Anta først $c \leq 0$. Det vil finnes x slik at $0 < |x| < \delta$ og $\sin\left(\frac{1}{x}\right) = 1$. For å se dette, la $x_n = \frac{1}{2\pi n + \frac{\pi}{2}}$, for et hvilket som helst heltall n. Da vil $\sin\left(\frac{1}{x_n}\right) = \sin\left(2\pi n + \frac{\pi}{2}\right) = 1$. Vi kan alltid finne en stor nok $n \in \mathbb{Z}$ som er slik at $0 < x_n < \delta$. (Sjekk at det holder å velge en n slik at $n > \frac{1}{2\pi\delta} - \frac{1}{4}$.) Siden $\sin\left(\frac{1}{x_n}\right) = 1$, har vi $\left|\sin\left(\frac{1}{x_n}\right) - c\right| = |1 - c| \geq 1$ (her bruker vi at $c \leq 0$), som er det vi ønsket å vise.

Anta nå at $c \geq 0$. Det vil finnes x slik at $0 < |x| < \delta$ og $\sin\left(\frac{1}{x}\right) = -1$: la $x_n = \frac{1}{2\pi n - \frac{\pi}{2}}$ og argumentér som ovenfor. Siden $\sin\left(\frac{1}{x_n}\right) = -1$, har vi $\left|\sin\left(\frac{1}{x_n}\right) - c\right| = |-1 - c| \geq 1$ (her bruker vi at $c \geq 0$), som er det vi ønsket å vise. □

EKSEMPEL 9.3.2 (Dirichlets funksjon). *Dirichlets funksjon* $f : \mathbb{R} \to \mathbb{R}$ er definert som

$$f(x) = \begin{cases} 1, & x \text{ rasjonal;} \\ 0, & x \text{ irrasjonal,} \end{cases}$$

og er oppkalt etter den tyske matematikeren Peter Gustav Lejeune Dirichlet (1805–1859). Tegner vi grafen, vil funksjonsverdiene i de rasjonale og irrasjonale tallene ligge så tett at den (feilaktig) ser ut som to rette linjer:



GRAFEN TIL DIRICHLET-FUNKSJONEN I EKSEMPEL 9.3.2.

La oss vise at f *er diskontinuerlig overalt*. Et viktig moment i beviset er at rasjonale og irrasjonale tall ligger *tett* på tallinjen, det vil si i ethvert åpent intervall finnes både rasjonale og irrasjonale tall, jf. Setning 7.2.2.

Formelt bevis for at f er diskontinuerlig overalt. La $a \in \mathbb{R}$. Vi bruker igjen Observasjon 9.2.16 med $\epsilon = 1$ og viser at uansett $\delta > 0$, så vil det finnes x slik at $|x - a| < \delta$ og

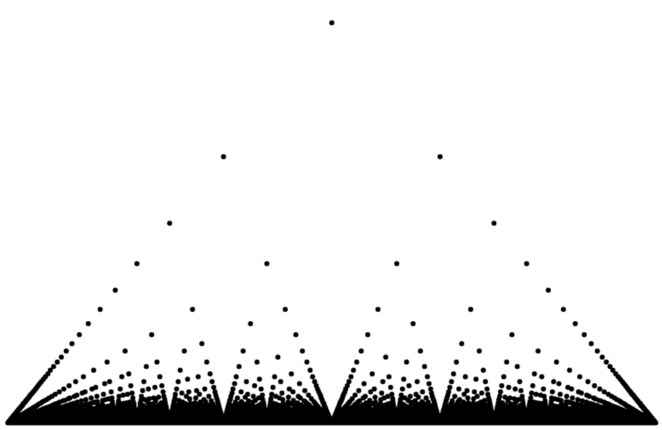
$$|f(x) - f(a)| \geq 1 \;\text{ og }\; |x - a| < \delta.$$

Uansett $\delta > 0$, vil det ved Setning 7.2.2 finnes en $x \in (a - \delta, a + \delta)$ slik at x er rasjonal hvis a er irrasjonal, og omvendt. Altså finnes $x \in (a - \delta, a + \delta)$ slik at $|f(x) - f(a)| = 1$, som er det vi skulle vise. $\square$

Eksempel 9.3.3 (Thomaes funksjon). En enda mer interessant funksjon er funksjonen $h : (0, 1) \to \mathbb{R}$ gitt ved

$$h(x) = \begin{cases} \frac{1}{q}, & \text{om } x = \frac{p}{q} \in \mathbb{Q} \text{ fullt forkortet;} \\ 0, & x \notin \mathbb{Q}. \end{cases}$$

At $x = \frac{p}{q}$ er fullt forkortet betyr at p og q ikke har felles primfaktorer, det vil si x er skrevet i *kanonisk form*, som nevnt i §3.4. Denne funksjonen kalles gjerne *Thomaes funksjon*, etter den tyske matematikeren Carl Johannes Thomae (1840–1921), eller *popcorn-funksjonen*. Her er grafen:



Grafen til Thomaes funksjon i Eksempel 9.3.3 (Kilde: Wikipedia, public domain.)

Denne funksjonen er *kontinuerlig i alle irrasjonale tall og diskontinuerlig i alle rasjonale tall i definisjonsmengden*! For å vise dette trenger man virkelig en presis definisjon på kontinuitet. La oss se på beviset.

Formelt bevis for at f er kontinuerlig presist på $\mathbb{R} \setminus \mathbb{Q}$. Anta først at $a = \frac{p}{q} \in \mathbb{Q}$, skrevet fullt forkortet. Da er $h(a) = \frac{1}{q}$. Uansett $\delta > 0$, eksisterer det en irrasjonal $x \in (a-\delta, a+\delta) \cap (0,1)$, ved *tetthetsegenskapen* (Setning 7.2.2). Da vil

$$|x-a| < \delta, \quad x \in (0,1), \quad \text{og} \quad |h(x) - h(a)| = \left|0 - \frac{1}{q}\right| = \frac{1}{q}.$$

Observasjon 9.2.16 med $\epsilon = \frac{1}{q}$ viser at h er diskontinuerlig i a.

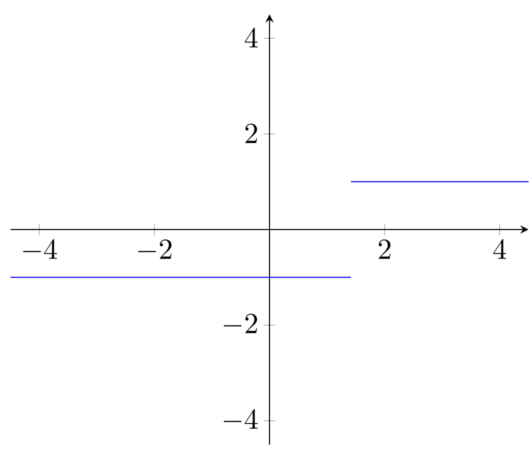
Anta så at a er irrasjonal, det vil si $a \in (\mathbb{R} \setminus \mathbb{Q}) \cap (0,1)$. Gitt $\epsilon > 0$, vil vi finne $\delta > 0$ slik at $|h(x) - h(a)| < \epsilon$ når $|x - a| < \delta$ og $x \in (0,1)$. Ved Arkimedes' prinsipp (Setning 7.2.1) finnes en $N \in \mathbb{Z}^+$ slik at $N\epsilon > 1$. I intervallet $(0,1)$ finnes bare endelig mange rasjonale tall $\frac{p}{q}$ med $q < N$. (Det finnes nemlig bare endelig mange muligheter for nevner: $1, \ldots, N-1$, og for hver nevner q, kun endelig mange muligheter for teller: $1, \ldots, q-1$.) La a_0 være tallet blant disse som ligger nærmest a og sett $\delta = |a - a_0|$, som er større enn 0, fordi $a \neq a_0$, siden a er irrasjonal og a_0 er rasjonal. I intervallet $(a-\delta, a+\delta) \cap (0,1)$ vil derfor alle rasjonale tall være på formen $\frac{p}{q}$ med $q \geq N$, og dermed oppfylle $h\left(\frac{p}{q}\right) = \frac{1}{q} \leq \frac{1}{N} < \epsilon$. Siden $h(x) = 0$ for alle irrasjonale x, vil alle $x \in (a-\delta, a+\delta) \cap (0,1)$ oppfylle $|h(x)| < \epsilon$. Derfor, hvis $|x - a| < \delta$ og $x \in (0,1)$, vil $|h(x) - h(a)| = |h(x)| < \epsilon$, og beviset er fullført. $\square$

Vi tar med et eksempel hvor funksjonen er definert på $\mathbb{Q}$, hvor svaret kanskje er overraskende:

Eksempel 9.3.4. La $f : \mathbb{Q} \to \mathbb{Q}$ være funksjonen definert ved

$$f(x) = \begin{cases} -1, & x < \sqrt{2} \\ 1, & x > \sqrt{2}. \end{cases}$$

Tegner vi grafen, vil den igjen på grunn av tetthetsegenskapen til de rasjonale tallene (Setning 7.2.2) se ut (feilaktig) som to rette linjer, og gjøre et hopp rundt $\sqrt{2}$:



Grafen til funksjonen i Eksempel 9.3.4.

Funksjonen er imidlertid kontinuerlig! Poenget er at ethvert punkt i definisjonsmengden har en omegn om seg hvor funksjonen er konstant. La oss vise dette.

La $a \in \mathbb{Q}$. Gitt $\epsilon > 0$. Hvis $a < \sqrt{2}$, la $\delta = \sqrt{2} - a > 0$. Da vil $(a - \delta, a + \delta) \subset (-\infty, \sqrt{2})$. Dermed vil alle $x \in \mathbb{Q} \cap (a - \delta, a + \delta)$ oppfylle $x < \sqrt{2}$, slik at $f(x) = -1 = f(a)$. Vi har dermed vist at vi har $|f(x) - f(a)| = 0 < \epsilon$ for alle $x \in D(f)$ slik at $|x - a| < \delta$.

Beviset for $a > \sqrt{2}$ er helt tilsvarende.

Vi avslutter seksjonen med et eksempel som viser en litt uvant side ved kontinuitet. Det viser at *diskrete funksjoner*, det vil si at funksjoner som har definisjonsmengde som er inneholdt i $\mathbb{Z}$, er automatisk kontinuerlige!

Eksempel 9.3.5. La $f : D(f) \to \mathbb{R}$ være en funksjon slik at $D(f) \subset \mathbb{Z}$. Vis at f er kontinuerlig i enhver $a \in D(f)$.

Løsning. La $a \in D(f)$. Gitt $\epsilon > 0$. La $\delta = 1$. (Enhver positiv $\delta \leq 1$ vil imidlertid fungere.) Hvis $x \in D(f)$ slik at $|x - a| < \delta$, da vil nødvendigvis $x = a$, for siden $D(f) \subset \mathbb{Z}$, så er avstanden mellom to forskjellige tall i $D(f)$ alltid minst 1. Derfor har vi at $|f(x) - f(a)| = |f(a) - f(a)| = 0 < \epsilon$, som viser at f kontinuerlig i a. □

Et lignende bevis viser at enhver funksjon er kontinuerlig i såkalte *isolerte punkter* i sin definisjonsmengde, jf. Oppgave 9.6.

Merknad 9.3.6. Eksempel 9.3.5 gjelder også om vi bytter ut $\mathbb{R}$ med $\mathbb{C}$.

9.4. Følger og kontinuitet

Vi skal nå se at kontinuitetsbegrepet henger sammen med egenskaper til konvergente følger. Neste resultat er veldig viktig og gir en nødvendig og tilstrekkelig betingelse for kontinuitet uttrykt ved hjelp av følger. Dette kriteriet blir i en god del lærebøker brukt som *definisjon* på kontinuitet, hvoretter kriteriet i Definisjon 9.1.1 blir utledet som en konsekvens.

Setning 9.4.1: Følge-kriteriet for kontinuitet

La $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$ og la $a \in D(f)$. Funksjonen f er kontinuerlig i a hvis og bare hvis $\lim_{n\to\infty} f(x_n) = f(a)$ for alle følger $\{x_n\}$ i $D(f)$ som konvergerer mot a.

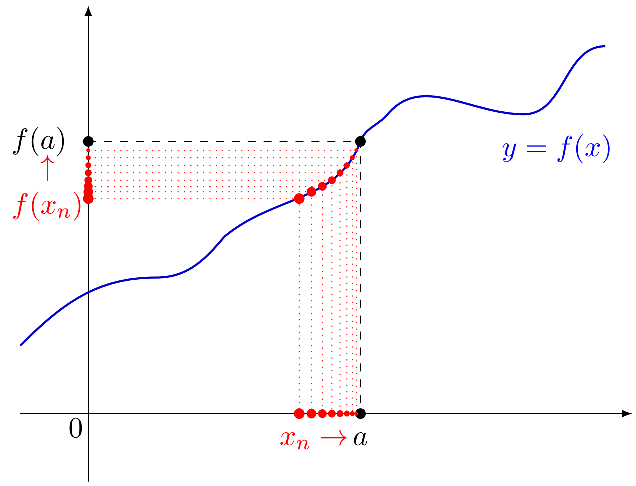
Vi husker dette resultatet kanskje best som :

$$\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n),$$

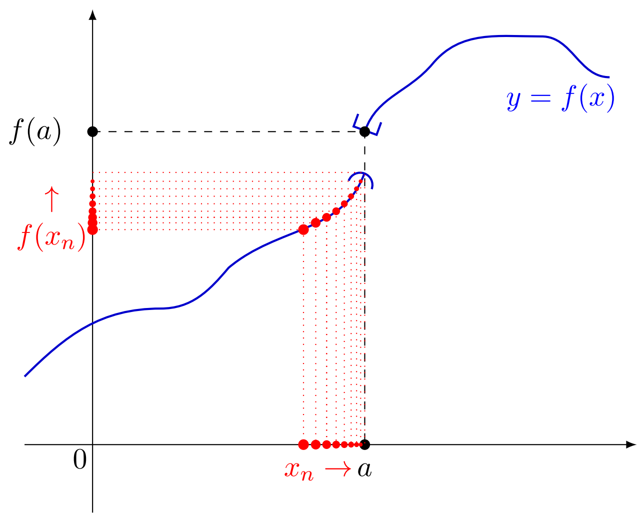
det vil si

grensen kan "trekkes innenfor" en kontinuerlig funksjon.

Følgende figurer illustrerer kriteriet:



Kontinuitet: $f(x_n)$ konvergerer mot $f(a)$ for alle x_n som konvergerer mot a.



Diskontinuitet: finnes x_n som konvergerer mot a, mens $f(x_n)$ ikke konvergerer mot $f(a)$.

Bevis for Setning 9.4.1

Vi viser først "bare hvis"-utsagnet, det vil si retningen "$\Longrightarrow$".

Anta at f er kontinuerlig i a og at $\{x_n\}$ er en følge i X som konvergerer mot a. Vi vil vise at følgen $f(x_n)$ konvergerer mot $f(a)$. Gitt $\epsilon > 0$, vil vi altså finne en $N \in \mathbb{Z}^+$ slik at $|f(x_n) - f(a)| < \epsilon$ for alle $n \geq N$. Siden f er kontinuerlig i a, finnes en $\delta > 0$ slik at $|f(x) - f(a)| < \epsilon$ for alle $x \in D(f)$ med $|x - a| < \delta$. Siden $\lim\limits_{n\to\infty} x_n = a$, finnes en $N \in \mathbb{Z}^+$ slik at $|x_n - a| < \delta$ når $n \geq N$. Men dermed er $|f(x_n) - f(a)| < \epsilon$ når $n \geq N$ og vi er ferdig.

Vi viser deretter "hvis"-utsagnet, det vil si retningen "$\Longleftarrow$", ved et kontrapositivt bevis.

Anta derfor at f ikke er kontinuerlig i a. Vi vil konstruere en følge $\{x_n\}$ i $D(f)$ slik at x_n konvergerer mot a, mens $f(x_n)$ *ikke* konvergerer mot $f(a)$. Vi bruker Observasjon 9.2.16. Siden f ikke er kontinuerlig i a, vil det finnes en $\epsilon > 0$ slik at uansett $\delta > 0$, så vil det finnes en $x \in D(f)$ slik at $|x - a| < \delta$ og $|f(x) - f(a)| \geq \epsilon$. Spesielt for $\delta = \frac{1}{n}$, med $n \in \mathbb{Z}^+$, så vil det finnes en $x_n \in D(f)$ slik at $|x_n - a| < \frac{1}{n}$ og $|f(x_n) - f(a)| \geq \epsilon$. Da har vi $|x_n - a| \to 0$ ved Skvisesetningen 7.4.13, slik at x_n konvergerer mot a (ved Oppgavene 7.18 og 7.19). Men $f(x_n)$ konvergerer ikke mot $f(a)$, siden $|f(x_n) - f(a)| \geq \epsilon$ for alle $n \in \mathbb{Z}^+$. □

Merknad 9.4.2. Setning 9.4.1 er fremdeles gyldig om vi erstatter $\mathbb{R}$ med $\mathbb{C}$. Det finnes ikke noe analogt resultat til Skvisesetningen 7.4.13 for komplekse følger, men i beviset brukte vi denne på den reelle følgen $|x_n - a|$. Resultatene fra Oppgavene 7.18 og 7.19 er fremdeles sanne i $\mathbb{C}$, jf. Oppgave 8.26, slik at beviset fungerer.

Kriteriet ovenfor er et nyttig verktøy i studiet av kontinuitet, spesielt kombinert med grensereglene for følger som vi har vist i Setning 7.4.8. For eksempel kan kriteriet brukes på alle eksemplene studert i forrige seksjon. La oss illustrere dette med følgende generalisering av Eksemplene 9.2.1 og 9.2.11.

Setning 9.4.3

La $f : \mathbb{R} \to \mathbb{R}$ være en funksjon gitt ved et polynom. Da er f kontinuerlig i alle $a \in \mathbb{R}$.

Bevis

La $a \in \mathbb{R}$ og la $\{a_n\}$ være en følge som konvergerer mot a. Da er $\lim_{n\to\infty} f(a_n) = f(a)$ ved Følgesetning 7.4.9, slik at f er kontinuerlig i a ved Setning 9.4.1. □

Merknad 9.4.4. Setning 9.4.3 gjelder fremdeles hvis vi erstatter $\mathbb{R}$ med $\mathbb{C}$, ved Merknad 9.4.2 og Følgesetning 8.5.5.

La oss nå anvende følge-kriteriet på tilfellene studert i Eksemplene 9.2.15, 9.3.1 og 9.3.3.

Eksempel 9.4.5. Vis at funksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved

$$f(x) = \begin{cases} -\frac{1}{4}x^2 - 1, & x < 2, \\ 2x - 8, & x \geq 2, \end{cases}$$

er diskontinuerlig i 2.

Løsning (ved hjelp av følge-kriteriet). Betrakt følgen $\{x_n\}$ definert ved $x_n = 2 - \frac{1}{n}$. Da vil $x_n \in (-\infty, 2)$ for alle $n \in \mathbb{Z}^+$ og $x_n \to 2$. Siden funksjonen gitt ved $x \mapsto -\frac{1}{4}x^2 - 1$ er kontinuerlig ved Setning 9.4.3, vil

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty}\left(-\frac{1}{4}x_n^2 - 1\right) = -\frac{1}{4}\cdot 2^2 - 1 = -2,$$

og siden $f(2) = -4 \neq -2$, er f diskontinuerlig i 2 ved Setning 9.4.1. □

Eksempel 9.4.6. Vis at funksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved

$$f(x) = \begin{cases} \sin\left(\frac{1}{x}\right), & x \neq 0, \\ c, & x = 0, \end{cases}$$

hvor $c \in \mathbb{R}$ er et vilkårlig fiksert tall, er diskontinuerlig i 0.

Løsning (ved hjelp av følge-kriteriet). Betrakt følgen $\{x_n\}$ definert ved $x_n = \frac{1}{\pi n + \frac{\pi}{2}}$. Da vil $x_n \to 0$. (For å være helt presis: gitt $\epsilon > 0$, da vil $|x_n| \leq \frac{1}{n}$, som vil være mindre enn ϵ såfremt $n > \frac{1}{\epsilon}$.) Samtidig er

$$f(x_n) = \sin\left(\pi n + \frac{\pi}{2}\right) = \begin{cases} 1, & \text{når } n \text{ er partall,} \\ -1, & \text{når } n \text{ er oddetall.} \end{cases}$$

Derfor konvergerer ikke $f(x_n)$, og beviset er fullført. □

Eksempel 9.4.7. Vis at Thomaes funksjon $h : (0,1) \to \mathbb{R}$ gitt ved

$$h(x) = \begin{cases} \frac{1}{q}, & \text{om } x = \frac{p}{q} \in \mathbb{Q} \text{ fullt forkortet;} \\ 0, & x \notin \mathbb{Q} \end{cases}$$

er diskontinuerlig i alle rasjonale tall i definisjonsmengden.

Løsning (ved hjelp av følge-kriteriet). Anta at $a = \frac{p}{q} \in \mathbb{Q}$. Ved Setning 7.2.2 kan vi i ethvert intervall $(a - \frac{1}{n}, a + \frac{1}{n})$, for alle $n \in \mathbb{Z}^+$, finne et irrasjonalt tall x_n. Følgen $\{x_n\}$ konvergerer mot a, mens $h(x_n) = 0$ for alle $n \in \mathbb{Z}^+$, slik at $h(x_n)$ ikke kan konvergere mot $h(a) = \frac{1}{q} \neq 0$. Altså er f diskontinuerlig i a ved Setning 9.4.1. □

9.5. Mange kontinuerlige funksjoner

Som vi nevnte i forbindelse med diskusjonen rundt spørsmål 2 helt i starten av §9.3, er det heldigvis ikke slik at vi for enhver funksjon vi studerer bruker ϵ–δ definisjonen av kontinuitet (eller følge-kriteriet) til å avgjøre kontinuitet. Vi bruker definisjonen på en del utvalgte "basisfunksjoner", og så utleder vi kontinuitet av andre funksjoner ved hjelp av noen generelle resultater. Det er nå på tide å avsløre hva disse resultatene er.

Generelle resultater om kontinuerlige funksjoner.

Setning 9.5.1: Kombinasjoner av kontinuerlige funksjoner

Anta f og g er kontinuerlige i punktet a og la $c \in \mathbb{R}$. Da er også funksjonene $f+g$, $f-g$, cf og $f \cdot g$ kontinuerlige i a. Hvis i tillegg $g(a) \neq 0$, da er også $\frac{f}{g}$ kontinuerlig i a.

Denne setningen kan vises ved hjelp av ϵ–δ definisjonen, jf. Oppgave 9.10. Man kan imidlertid også vise setningen ved å bruke følge-kriteriet (Setning 9.4.1), hvilket vi nå skal gjøre.

Bevis for Setning 9.5.1
La $\{x_n\}$ være en følge (i definisjonsmengden til de aktuelle funksjonene) som konvergerer mot a. Siden f og g antas å være kontinuerlige i a, har vi at $f(x_n) \to f(a)$ og $g(x_n) \to g(a)$. Ved Setning 7.4.8 har vi

$$\begin{aligned}
(f+g)(x_n) = f(x_n) + g(x_n) &\quad\to\quad f(a) + g(a) = (f+g)(a), \\
(f-g)(x_n) = f(x_n) - g(x_n) &\quad\to\quad f(a) - g(a) = (f-g)(a), \\
(cf)(x_n) = cf(x_n) &\quad\to\quad cf(a) = (cf)(a), \\
(f \cdot g)(x_n) = f(x_n) \cdot g(x_n) &\quad\to\quad f(a)g(a) = (f \cdot g)(a),
\end{aligned}$$

og, dersom $g(a) \neq 0$, også

$$\frac{f}{g}(x_n) = \frac{f(x_n)}{g(x_n)} \to \frac{f(a)}{g(a)} = \frac{f}{g}(a),$$

som viser setningen ved Setning 9.4.1. □

At beviset for siste setning ved hjelp av følge-kriteriet er mye enklere enn beviset for setningen via ϵ–δ-argumentasjon, skyldes selvsagt at vi har gjort alle de harde delene i beviset for Setning 7.4.8.

Setning 9.5.2: Sammensetninger av kontinuerlige funksjoner

Anta at g er kontinuerlig i punktet a og f er kontinuerlig i punktet $g(a)$. Da er den sammensatte funksjonen $f \circ g$ kontinuerlig i a.

Igjen kan dette bevises ved hjelp av ϵ–δ definisjonen (jf. Oppgave 9.10). Vi viser setningen ved hjelp av følge-kriteriet (Setning 9.4.1):

Bevis
La $\{x_n\}$ være en følge (i definisjonsmengden til $f \circ g$, som er lik definisjonsmengden til g). Siden g er kontinuerlig i a, vil $g(x_n) \to g(a)$ ved Setning 9.4.1. Følgen $\{g(x_n)\}$ er en følge i definisjonsmengden til f, og siden f er kontinuerlig i punktet $g(a)$, vil $f(g(x_n)) \to f(g(a))$ ved Setning 9.4.1. Dette siste betyr at $(f \circ g)(x_n)) \to (f \circ g)(a)$, som ved Setning 9.4.1 betyr at $f \circ g$ er kontinuerlig i a. $\square$

Merknad 9.5.3. Setningene 9.5.1 og 9.5.2 fungerer like godt for funksjoner med definisjons- og verdimengde i $\mathbb{C}$ som i $\mathbb{R}$. (Bruk Setning 8.5.4 og Merknad 9.4.2 i det komplekse tilfellet.)

Setning 9.5.4: Kontinuitet av inverse funksjoner

La $I \subset \mathbb{R}$ være et intervall og $f : I \to f(I)$ en strengt voksende (henholdsvis strengt avtagende) funksjon. Da er f bijektiv og den inverse funksjonen $f^{-1} : f(I) \to I$ er kontinuerlig og strengt voksende (hhv. strengt avtagende).

Merk at vi *ikke* krever at f er kontinuerlig! Vi vil se senere, i Følgesetning 9.6.2, at dersom f i tillegg er kontinuerlig, da er $f(I)$ igjen et intervall.

Merknad 9.5.5. Det er helt essensielt at I er et intervall for at konklusjonen skal holde. Ta for eksempel den strengt voksende (kontinuerlige!) funksjonen $f : [0,1] \cup (2,4] \to [0,2]$ definert ved

$$f(x) = \begin{cases} x, & x \in [0,1] \\ \frac{x}{2}, & x \in (2,4]. \end{cases}$$

Da er den inverse funksjonen $f^{-1} : [0,2] \to [0,1] \cup (2,4]$ gitt ved

$$f^{-1}(x) = \begin{cases} x, & x \in [0,1] \\ 2x, & x \in (1,2], \end{cases}$$

som er diskontinuerlig i $x = 1$.

Bevis for Setning 9.5.4
Vi har allerede bemerket (Merknad 7.2.8) at en strengt monoton funksjon er injektiv, slik at $f : I \to f(I)$ er bijektiv. Vi viser nå at den inverse funksjonen $f^{-1} : f(I) \to I$ er kontinuerlig når f er strengt voksende. Tilfellet f strengt avtagende behandles tilsvarende. Vi har derfor at

$$x_1 < x_2 \iff f(x_1) < f(x_2), \quad \text{for alle } x_1, x_2 \in I. \tag{224}$$

Hvis $y_1 < y_2$ i $f(I)$, la $x_1 = f^{-1}(y_1)$ og $x_2 = f^{-1}(y_2)$, slik at $f(x_1) = y_1$ og $f(x_2) = y_2$. Da sier (224) at

$$f^{-1}(y_1) < f^{-1}(y_2) \iff y_1 < y_2,$$

altså er f^{-1} også strengt voksende.

Vi vil nå vise at f^{-1} er kontinuerlig i et vilkårlig punkt $a \in f(I)$. La derfor $\epsilon > 0$ være gitt og sett $b = f^{-1}(a) \in I$. Anta i første omgang at b ikke er et endepunkt i I. Da finnes $0 < \epsilon' \leq \epsilon$ slik at

$$[b - \epsilon', b + \epsilon'] \subset I.$$

Siden f er strengt voksende, er

$$f(b - \epsilon') < a = f(b) < f(b + \epsilon').$$

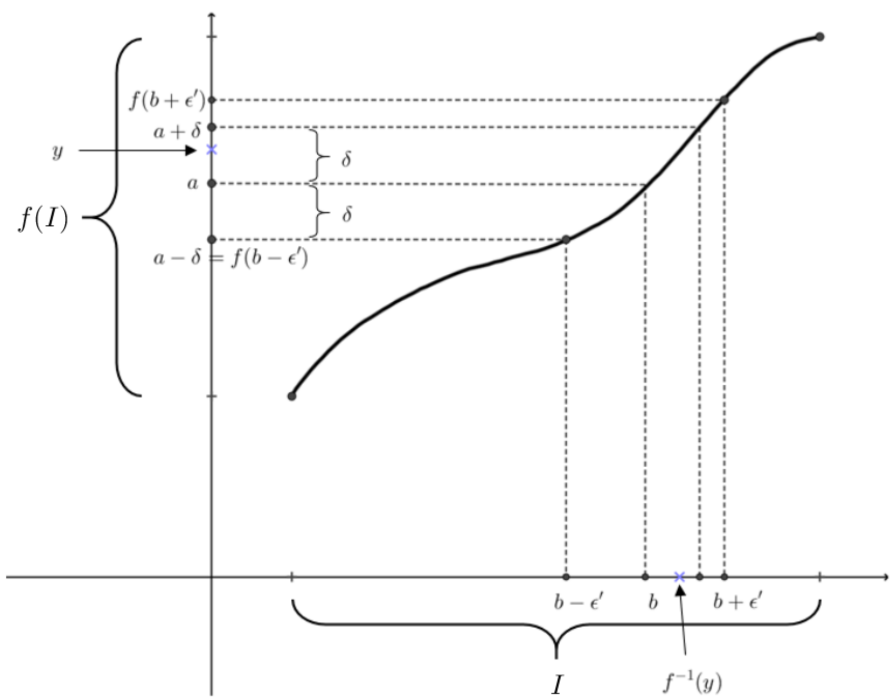
La

$$\delta = \min\{a - f(b - \epsilon'), f(b + \epsilon') - a\}.$$

Da vil alle $y \in f(I) \cap (a - \delta, a + \delta)$ oppfylle

$$f(b - \epsilon') < y < f(b + \epsilon'),$$

se figuren:



Siden f^{-1} er strengt voksende, har vi

$$b - \epsilon' = f^{-1}(f(b - \epsilon')) < f^{-1}(y) < f^{-1}(f(b + \epsilon')) = b + \epsilon',$$

altså $|f^{-1}(y) - b| < \epsilon'$. Siden $\epsilon' \leq \epsilon$, har vi

$$|f^{-1}(y) - b| < \epsilon \text{ for alle } y \in f(I) \cap (a - \delta, a + \delta).$$

Dette viser at f^{-1} er kontinuerlig i a.

Vi overlater behandlingen av tilfellet der b er et endepunkt i I til Oppgave 9.14. □

For å kunne bruke Setningene 9.5.1-9.5.4 trenger vi å vite kontinuitet for en del basisfunksjoner. La oss først merke følgende:

Definisjon 9.5.6: Kontinuitet

Vi sier at f er *kontinuerlig* i et intervall I dersom f er kontinuerlig i alle punkter $a \in I$.

Vi sier at f er en *kontinuerlig funksjon* dersom f er kontinuerlig i alle punkter a i sin definisjonsmengde $D(f)$.

Vi har allerede vist kontinuitet for en del basisfunksjoner:

- den konstante funksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved $f(x) = c$, hvor $c \in \mathbb{R}$, er kontinuerlig (Eksempel 9.2.1 med $A = 0$ og $B = c$).
- identitetsfunksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved $f(x) = x$ er kontinuerlig (Eksempel 9.2.1 med $A = 1$ og $B = 0$).
- absoluttverdifunksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved $f(x) = |x|$ er kontinuerlig (Eksempel 9.2.3).
- sinusfunksjonen $\sin : \mathbb{R} \to \mathbb{R}$ er kontinuerlig i 0 (Eksempel 9.2.6). Den er for øvrig kontinuerlig overalt ved lignende argumenter (jf. Merknad 9.2.8 og Oppgave 9.3).

Som en konsekvens av kontinuitetet for disse basisfunksjonene og Setningene 9.5.1-9.5.4 kan vi nå konkludere kontinuitet for en rekke velkjente funksjoner i de neste delseksjonene.

Polynomer, rasjonale funksjoner og rotfunksjoner. Vi minner om at en rasjonal funksjon er en funksjon gitt som $x \mapsto \frac{P(x)}{Q(x)}$, hvor P og Q er polynomer. Definisjonsmengden er alle x slik at $Q(x) \neq 0$.

Følgesetning 9.5.7

Polynomer og rasjonale funksjoner er kontinuerlige.

Bevis

Dette er en direkte konsekvens av Setning 9.5.1 og kontinuiteten til basisfunksjonene c og x ovenfor. (Vi har for øvrig vist at alle polynomer er kontinuerlige i Setning 9.4.3 ved følge-kriteriet direkte, og samme argument kan også brukes på rasjonale funksjoner.) □

MERKNAD 9.5.8. Følgesetning9.5.7 gjelder også for komplekse polynomer og rasjonale funksjoner.

Følgesetning 9.5.9

Rotfunksjoner er kontinuerlige. Mer presist, funksjonen $f : I \to \mathbb{R}$ definert ved $f(x) = x^{\frac{m}{n}} = \sqrt[n]{x^m}$, for $m, n \in \mathbb{Z}^+$, der

$$I = \begin{cases} \mathbb{R}, & \text{hvis } n \text{ er oddetall,} \\ [0, \infty), & \text{hvis } n \text{ er partall,} \end{cases}$$

er kontinuerlig i alle $a \in I$.

BEVIS
Ved Merknad 7.3.5 er funksjonen

$$\begin{aligned} g : I &\to I \\ x &\mapsto \sqrt[n]{x} \end{aligned}$$

den inverse funksjonen til funksjonen

$$\begin{aligned} I &\to I \\ x &\mapsto x^n, \end{aligned}$$

som er strengt voksende ved Merknad 7.3.5. Ved Setning 9.5.4 er funksjonen g derfor kontinuerlig. Da er $f(x) = \sqrt[n]{x^m} = g(x^m)$, slik at f er sammensetningen av g og den kontinuerlige funksjonen $x \mapsto x^m$, og er derfor kontinuerlig ved Setning 9.5.2. □

Trigonometriske funksjoner og deres inverser. Vi har sett definisjonen av den inverse sinusfunksjonen og den inverse cosinusfunksjonen i Eksempel 2.3.10. Vi henviser til Merknad 9.6.12 for definisjonen av den inverse tangensfunksjonen og til [**AE**, §P.7 og §3.5] for alle definisjonene av trigonometriske funksjoner og deres inverse funksjoner.

Følgesetning 9.5.10

Alle trigonometriske funksjoner og deres inverser er kontinuerlige.

BEVIS
Dette overlates til Oppgave 9.15. □

Eksponensial- og logaritmefunksjoner. Vi minner nå om hvordan eksponensialfunksjoner er definert. Fiksér en $a \in \mathbb{R}^+$. For enhver $n \in \mathbb{Z}^+$ har vi definert $a^{\frac{1}{n}}$ som det entydige reelle tallet slik at $\left(a^{\frac{1}{n}}\right)^n = a$ (jf. Definisjon 7.3.2). For $q = \frac{m}{n} \in \mathbb{Q}$ (med $m \in \mathbb{Z}$, $n \in \mathbb{Z}^+$) har vi definert $a^{\frac{m}{n}} = (a^m)^{\frac{1}{n}}$

(jf. Definisjon 7.3.7). Vi har utvidet dette til å gjelde for alle $x \in \mathbb{R}$ ved å definere

$$a^x = \sup\{a^t \mid t \in \mathbb{Q},\ t \leq x\}$$

(jf. Definisjon 7.3.8). Vi må selvsagt vise at de to definisjonene sammenfaller når $x \in \mathbb{Q}$, og dette ble overlatt til (den ikke så enkle) Oppgave 7.15(a). Vi har dermed en funksjon

$$\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \mapsto & a^x \end{array}$$

som vi kaller *eksponensialfunksjonen med grunntall (eller base) a*.

Vi vil vise:

Setning 9.5.11

Alle eksponensialfunksjoner er kontinuerlige.

Bevis

La $f : \mathbb{R} \to \mathbb{R}$ være en eksponensialfunksjon gitt ved $f(x) = a^x$ for en $a \in \mathbb{R}^+$.

Hvis $a = 1$, er f konstant (lik 1) ved Setning 7.3.9(vi) og det er ingenting å vise. Hvis $0 < a < 1$, er $\frac{1}{a} > 1$. Siden

$$a^x = \frac{1}{\left(\frac{1}{a^x}\right)} \overset{(*)}{=} \frac{1}{\left(\frac{1^x}{a^x}\right)} \overset{(**)}{=} \frac{1}{\left(\frac{1}{a}\right)^x},$$

hvor vi har brukt Setning 7.3.9(vi) i $(*)$ og Setning 7.3.9(v) i $(**)$, vil resultatet følge fra tilfellet $a > 1$ og Setning 9.5.1.

Vi vil derfor anta $a > 1$ i resten av beviset.

Vi viser først at det er tilstrekkelig å vise at f er kontinuerlig i 0. For $c \in \mathbb{R}$ kan vi nemlig skrive

$$f(x) = a^x = a^{c+(x-c)} \overset{(*)}{=} a^c \cdot a^{x-c},$$

hvor vi har brukt Setning 7.3.9(i) i $(*)$. Dermed er $f(x) = a^c f(g(x))$, der $g : \mathbb{R} \to \mathbb{R}$ er definert ved $g(x) = x - c$. Vi har at g er kontinuerlig i c. Dersom f er kontinuerlig i $0 = g(c)$, gir Setning 9.5.2 at $f \circ g$ er kontinuerlig i c. Da er også $f = a^c f \circ g$ kontinuerlig i c ved Setning 9.5.1.

Det gjenstår altså å vise at f er kontinuerlig i 0. Merk at $f(0) = a^0 = 1$ ved (62) eller Setning 7.3.9(vii). Gitt $\epsilon > 0$. Vi vil finne $\delta > 0$ slik at

$$|f(x) - f(0)| = |a^x - 1| < \epsilon \text{ når } |x - 0| < \delta,$$

eller med andre ord

$$1 - \epsilon < a^x < 1 + \epsilon \text{ når } -\delta < x < \delta.$$

Vi vil vise at $\delta = \frac{1}{n}$ for en $n \in \mathbb{Z}^+$ slik at $n\epsilon > a - 1 (> 0)$ vil duge. (En slik n finnes ved Arkimedes' prinsipp (Setning 7.2.1).)

Vi merker først at *Bernoullis ulikhet* (jf. Oppgave 7.13) medfører at

$$a < 1 + n\epsilon \leq (1 + \epsilon)^n,$$

slik at

$$a^{\frac{1}{n}} < 1 + \epsilon \tag{225}$$

(siden røtter bevarer orden ved Hjelpesetning 7.3.3).

Funksjonen $x \to a^x$ er strengt voksende for $a > 1$ (Merknad 7.3.11). Hvis $0 \leq x < \delta$, er derfor

$$1 = a^0 \leq a^x < a^\delta = a^{\frac{1}{n}} \overset{(225)}{<} 1 + \epsilon. \tag{226}$$

Hvis $-\delta < x \leq 0$, er av samme grunn

(227)
$$1 = a^0 \geq a^x > a^{-\delta} = \frac{1}{a^{\frac{1}{n}}} \overset{(225)}{>} \frac{1}{1+\epsilon} > \frac{1-\epsilon^2}{1+\epsilon} = \frac{(1+\epsilon)(1-\epsilon)}{1+\epsilon} = 1 - \epsilon.$$

Til sammen viser (226) og (227) at $1 - \epsilon < a^x < 1 + \epsilon$ når $|x| < \delta = \frac{1}{n}$, som viser at f er kontinuerlig i 0. □

Vi kan nå bevise eksistensen av logaritmer:

Setning 9.5.12: Eksistens og entydighet av logaritmer

La $a \in \mathbb{R}^+ \setminus \{1\}$. Til enhver $y \in \mathbb{R}^+$ finnes nøyaktig én $x \in \mathbb{R}$ slik at $a^x = y$.

Definisjon 9.5.13: Logaritme

Tallet x i setningen, som oppfyller $a^x = y$, kalles *logaritmen til y med grunntall (eller base) a* og betegnes med $\log_a y$.

Logaritmen til y med grunntall lik eulertallet e (Definisjon 7.5.7) kalles *den naturlige logaritmen til y* og betegnes med $\ln y$.

Bevis for Setning 9.5.12

Entydigheten av x følger fra Setning 7.3.9(viii), som gir at $a^{x_1} \neq a^{x_2}$ når $x_1 \neq x_2$. (Her trengs $a \neq 1$.)

For å vise eksistens, antar vi først at $a > 1$. Da vil $\left(\frac{1}{a}\right)^n \to 0$ ved Oppgave 7.21(a). For enhver $y \in \mathbb{R}^+$, finnes derfor $m \in \mathbb{Z}^+$ slik at $\left(\frac{1}{a}\right)^m \leq y$ og $\left(\frac{1}{a}\right)^m \leq \frac{1}{y}$. Dermed er $a^{-m} \leq y \leq a^m$. Sett $x_1 = -m$ og $x_2 = m$. Strategien vi bruker videre er *intervallhalveringsmetoden* som vi allerede har brukt i beviset for Setning 7.6.1 og i Oppgave 7.45.

Del intervallet $[x_1, y_1] = [-m, m]$ i to deler av like stor lengde $[-m, 0]$ og $[0, m]$. Da må enten $a^{-m} \leq y \leq a^0 = 1$ eller $1 = a^0 \leq y \leq a^m$. Velg $[x_2, y_2] = [-m, 0]$ eller $[0, m]$ slik at

$$a^{x_2} \leq y \leq a^{y_2}.$$

Merk at $[x_2, y_2]$ har lengde lik halvparten av det opprinnelige intervallet, nemlig $\frac{1}{2}(y_1 - x_1) = m$.

Vi fortsetter prosedyren med å halvere intervallet: Anta at vi har funnet $[x_n, y_n]$ slik at

$$a^{x_n} \leq y \leq a^{y_n}. \tag{228}$$

Del intervallet $[x_n, y_n]$ i to deler av like stor lengde

$$\left[x_n, \frac{x_n + y_n}{2}\right] \quad \text{og} \quad \left[\frac{x_n + y_n}{2}, y_n\right].$$

Da må enten $a^{x_n} \leq y \leq a^{\frac{x_n+y_n}{2}}$ eller $a^{\frac{x_n+y_n}{2}} \leq y \leq a^{y_n}$. Velg $[x_{n+1}, y_{n+1}]$ blant $\left[x_n, \frac{x_n+y_n}{2}\right]$ og $\left[\frac{x_n+y_n}{2}, y_n\right]$ slik at

$$a^{x_{n+1}} \leq y \leq a^{y_{n+1}}.$$

Da har vi $y_{n+1} - x_{n+1} = \frac{1}{2}(y_n - x_n)$, og dermed

$$\begin{aligned} y_{n+1} - x_{n+1} &= \frac{1}{2}(y_n - x_n) = \frac{1}{2} \cdot \frac{1}{2}(y_{n-1} - x_{n-1}) = \cdots \\ &= \left(\frac{1}{2}\right)^n (y_1 - x_1) = \frac{2m}{2^n} = \frac{m}{2^{n-1}}, \end{aligned}$$

altså

$$y_{n+1} - x_{n+1} = \frac{m}{2^{n-1}}. \tag{229}$$

Siden

$$[x_{n+1}, y_{n+1}] \subset [x_n, y_n] \subset \cdots \subset [x_2, y_2] \subset [x_1, y_1],$$

er $\{x_n\}$ voksende og oppad begrenset av $y_1 = m$. Dermed er $\{x_n\}$ konvergent ved Setning 7.5.4(i). Likeledes er $\{y_n\}$ avtagende og nedad begrenset av $x_1 = -m$, slik at $\{y_n\}$ er konvergent ved Setning 7.5.4(ii). Ved (229) er

$$\lim_{n\to\infty}(x_n - y_n) = 0,$$

og dermed, ved Setning 7.4.8(ii), er $\lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} y_n$. La x være denne grensen. Ved Setningene 9.4.1 og 9.5.11 er

$$\lim_{n\to\infty} a^{x_n} = a^{\lim_{n\to\infty} x_n} = a^x = a^{\lim_{n\to\infty} y_n} = \lim_{n\to\infty} a^{y_n}.$$

Siden $a^{x_n} \leq y \leq a^{y_n}$ for alle $n \in \mathbb{Z}^+$, følger det fra Skvisesetningen 7.4.13 at $a^x = y$. Dette fullfører beviset i tilfellet $a > 1$.

Hvis $0 < a < 1$, er $\frac{1}{a} < 1$. Ved tilfellet vi nettopp har vist finnes en $x \in \mathbb{R}$ slik at $\left(\frac{1}{a}\right)^x = y$. Siden $\left(\frac{1}{a}\right)^x = a^{-x}$ ved Setning 7.3.9(v), har vi funnet at $a^{-x} = y$, altså at $-x \in \mathbb{R}$ oppfyller den ønskede egenskapen. $\square$

Som allerede nevnt i Merknad 7.3.11 sier Setning 9.5.12 at for $a \in \mathbb{R}^+ \setminus \{1\}$ er eksponensialfunksjonen

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

bijektiv, med invers funksjon

$$\begin{aligned} f^{-1}: \mathbb{R}^+ &\to \mathbb{R} \\ y &\mapsto \log_a y. \end{aligned}$$

Siden både $f \circ f^{-1}$ og $f^{-1} \circ f$ er identitetsfunksjoner (jf. teksten rett etter Definisjon 2.3.5), har vi egenskapene

$$a^{\log_a x} = x, \quad x \in (0, \infty) = \mathbb{R}^+, \tag{230}$$

$$\log_a a^x = x, \quad x \in \mathbb{R}, \tag{231}$$

som er kjent fra skolen.

Dessuten sier Setning 7.3.9(viii) at f er strengt voksende for $a > 1$ og strengt avtagende for $0 < a < 1$. Som umiddelbar konsekvens av Setning 9.5.4 får vi derfor:

Setning 9.5.14

Alle logaritmefunksjoner $x \mapsto \log_a x$, $a \in \mathbb{R}^+ \setminus \{1\}$ er kontinuerlige. De er strengt voksende for $a > 1$ og strengt avtagende for $0 < a < 1$.

MERKNAD 9.5.15. La oss minne oss om en annerledes fremgangsmåte for å definere eksponensial- og logaritmefunksjoner som ikke går veien om Definisjon 7.3.8 og den tilhørende Oppgave 7.15(a)-(b).

Man kan definere *den naturlige logaritmefunksjonen* $\ln : (0, \infty) \to \mathbb{R}$ som en *integralfunksjon*: $\ln x = \int_1^x \frac{1}{t}\, dt$. Vi skal ikke komme nærmere innpå det her, men henviser til [**AE**, §3.3 og §5.5] for definisjonen. Det følger av *Fundamentalteoremet i Kalkulus* (se [**AE**, Thm. 5.5.5] eller [**Li**, 8.3.3]) og kontinuiteten av funksjonen $x \mapsto \frac{1}{x}$ på intervallet $(0, \infty)$ at funksjonen ln er kontinuerlig. Det følger av definisjonen at ln er strengt voksende og dermed injektiv. Den har derfor en inversfunksjon, som vi kaller *eksponensialfunksjonen* $\exp : \mathbb{R} \to (0, \infty)$ (se [**AE**, §3.3]), og Setning 9.5.4 gir at også exp er kontinuerlig.

Deretter viser man at $\exp(1) = e$, hvor e er eulertallet (Definisjon 7.5.7), som er ekvivalent med å vise at $\ln e = \int_1^e \frac{1}{t}\, dt = 1$.

Siden alle logaritme- og eksponensialfunksjoner uttrykkes enkelt ved hjelp av de to grunnleggende logaritme- og eksponensialfunksjonenene ($a^x = e^{x \ln a}$ og $\log_a x = \frac{\ln x}{\ln a}$, vi viser igjen til [**AE**, §3.3]), gir Setningene 9.5.1 og 9.5.2 at alle logaritme- og eksponensialfunksjoner er kontinuerlige.

Ulempen med denne prosedyren er at man trenger å vise at det bestemte integralet er veldefinert og trenger fundamentalteoremet i kalkulus.

Potensfunksjoner. Til slutt viser vi følgende generalisering av Følgesetning 9.5.9 til alle reelle eksponenter:

Setning 9.5.16

Alle potensfunksjoner er kontinuerlige. Mer presist, la $r \in \mathbb{R}$. Da er funksjonen

$$\begin{array}{rcl} \mathbb{R}^+ & \to & \mathbb{R} \\ x & \mapsto & x^r \end{array}$$

kontinuerlig.

Bevis

Velg en vilkårlig $a \in \mathbb{R}^+$. Da har vi

$$x^r \overset{(230)}{=} a^{\log_a(x^r)} \overset{(*)}{=} a^{r \log_a x} \overset{(**)}{=} (a^r)^{\log_a x},$$

hvor $(*)$ er Oppgave 9.18(e), mens $(**)$ er Setning 7.3.9(iii). Funksjonen $x \to x^r$ er derfor en sammensetning av funksjonene $x \mapsto \log_a x$, som er kontinuerlig ved Setning 9.5.14, og $x \mapsto (a^r)^x$, som er kontinuerlig ved Setning 9.5.11. Setningen følger da av Setning 9.5.2. $\square$

Oppsummering og eksempler. Som oppsummerende "huskeregel" av setningene og følgesetningene ovenfor har vi:

Alle kombinasjoner av polynomer, rasjonale funksjoner, trigonometriske funksjoner, inverse trigonometriske funksjoner, potens-, logaritme- og eksponensialfunksjoner ved hjelp av de aritmetiske operasjonene $+$, $-$, $\cdot$, $/$, samt absoluttverdi, røtter og sammensetninger, er kontinuerlige (på sine definisjonsmengder).

Eksempel 9.5.17. Funksjonen $f : \mathbb{R} \setminus \{0\} = (-\infty, 0) \cup (0, \infty) \to \mathbb{R}$ gitt ved $f(x) = \frac{1}{x}$ er kontinuerlig. Det er en vanlig misoppfatning at den *ikke* er det, siden den ikke er definert i 0. Men 0 er altså ikke med i definisjonsmengden til f. (Se også Eksempel 9.6.3.)

Merk at "huskeregelen" forutsetter at funksjonene har samme uttrykk i en omegn om alle punkter. For funksjoner med delt forskrift må vi argumentere spesielt, som vist i følgende eksempler:

Eksempel 9.5.18. Funksjonen f i Eksempel 9.2.15 er kontinuerlig bortsett fra i $x = 2$: vi så i eksemplet at funksjonen er diskontinuerlig i $x = 2$. På hvert av intervallene $(-\infty, 2)$ og $(2, \infty)$ er f gitt ved polynomer, slik at f er kontinuerlig der ved Følgesetning 9.5.7.

EKSEMPEL 9.5.19. Funksjonen $f : \mathbb{R} \to \mathbb{R}$ definert ved

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & x \neq 0 \\ 0 & x = 0, \end{cases} \tag{232}$$

er kontinuerlig: vi har vist at den er kontinuerlig i $x = 0$ i Eksempel 9.2.5. I hvert av intervallene $(-\infty, 0)$ og $(0, \infty)$ er $x \mapsto \sin\left(\frac{1}{x}\right)$ en sammensetning av de kontinuerlige funksjonene $x \mapsto \sin x$ og $x \mapsto 1/x$ og er dermed kontinuerlig der ved Setning 9.5.2, og funksjonen $x \mapsto x^2 \sin\left(\frac{1}{x}\right)$ er et produkt av kontinuerlige funksjoner og dermed kontinuerlig ved Setning 9.5.1. Derfor er f kontinuerlig overalt på sin definisjonsmengde.

EKSEMPEL 9.5.20 ("Sammenlapping" av kontinuerlige funksjoner). La $f : (a, b] \to \mathbb{R}$ og $g : [b, c) \to \mathbb{R}$ være kontinuerlige funksjoner slik at $f(b) = g(b)$. Vis at da er funksjonen $h : (a, c) \to \mathbb{R}$ gitt ved

$$h(x) = \begin{cases} f(x) & \text{hvis } x \in (a, b], \\ g(x) & \text{hvis } x \in [b, c) \end{cases}$$

kontinuerlig. (Merk at siden $f(b) = g(b)$, er det ikke noe tvetydig ved denne definisjonen av h.)

LØSNING. På de to åpne intervallene (a, b) og (b, c) er h gitt ved f eller g og er dermed kontinuerlig. Kontinuiteten i b vises ved et ϵ-δ-argument: gitt $\epsilon > 0$, da finnes en $\delta_1 > 0$ slik at

$$|f(x) - f(b)| < \epsilon \text{ når } |x - b| < \delta_1,\ x \in (a, b], \tag{233}$$

siden f er kontinuerlig i b, og en $\delta_2 > 0$ slik at

$$|g(x) - g(b)| < \epsilon \text{ når } |x - b| < \delta_2,\ x \in (b, c), \tag{234}$$

siden g er kontinuerlig i b. La $\delta = \min\{\delta_1, \delta_2\}$. Siden $h = f$ på $(a, b]$ og $h = g$ på $[b, c)$ (her er det viktig at $f(b) = g(c)$), kan (233)-(234) omskrives til

$$|h(x) - h(b)| < \epsilon \text{ når } |x - b| < \delta,\ x \in (a, c),$$

som viser at h er kontinuerlig i b. □

I de neste to seksjonene skal vi se på og bevise to ekstremt viktige resultater som angår kontinuerlige funksjoner, nemlig *skjæringssetningen*, og *ekstremalverdisetningen*. Bevisene til begge disse bygger på *kompletthetsprinsippet for de reelle tallene*.

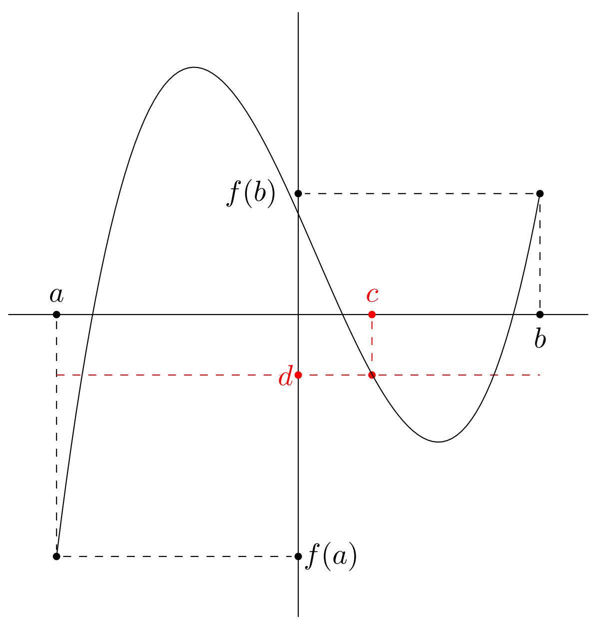
9.6. Skjæringssetningen

I denne seksjonen skal vi se på et av de fundamentale resultatene om kontinuitet, som brukes i mange sammenhenger. Det sier at en kontinuerlig funksjon ikke kan gå fra én verdi til en annen uten å anta alle mellomliggende verdier (derav navnet *mellomverdisetningen*[2]).

Teorem 9.6.1: Skjæringssetningen/mellomverdisetningen

Anta at $f : [a, b] \to \mathbb{R}$ er en kontinuerlig funksjon og d et hvilket som helst tall mellom $f(a)$ og $f(b)$. Da finnes et tall $c \in [a, b]$ slik at $f(c) = d$.

Figuren nedenunder visualiserer setningen og vi kan tenke på den som at den sier at enhver horisontal linje $y = d$ vi tegner mellom $y = f(a)$ og $y = f(b)$ vil skjære grafen i minst ett punkt, nemlig (c, d), derav navnet *skjæringssetning*:



VISUALISERING AV SKJÆRINGSSETNINGEN.

Skjæringssetningen sier ikke at punktet c er *entydig*; som antydet på figuren kan det godt finnes flere punkter som oppfyller samme betingelse.

Merk at funksjonen f godt kan være definert på et større intervall enn $[a, b]$, slik at setningen sier at en kontinuerlig funksjon definert på et intervall antar alle verdier mellom to vilkårlige funksjonsverdier på intervallet; eller sagt på en mer intuitiv måte: grafen til en kontinuerlig funksjon på et intervall er sammenhengende, det vil si "vi kan tegne den uten å løfte blyanten".

[2] *Intermediate Value Theorem* på engelsk.

Vi bruker ofte *skjæringssetningen* uten å tenke mye over det, for eksempel når vi finner verdimengden til en funksjon. Som et konkret eksempel, ta sinusfunksjonen $\mathbb{R} \to \mathbb{R}$ gitt ved $x \mapsto \sin x$. Vi vet at den antar verdier kun mellom -1 og 1 og at den antar verdiene -1 og 1. Da kan vi ved *skjæringssetningen* konkludere at verdimengden er hele intervallet $[-1, 1]$, fordi funksjonen virkelig antar alle mellomliggende verdier. (Se Eksempel 9.6.11 for et lignende eksempel.)

Litt mer generelt får vi som en direkte konsekvens av *skjæringssetningen* at en kontinuerlig funksjon avbilder intervaller på intervaller:

Følgesetning 9.6.2

Hvis $f : D(f) \to \mathbb{R}$, med $D(f) \subset \mathbb{R}$, er kontinuerlig, og $I \subset D(f)$ er et intervall, da er også $f(I)$ et intervall.

Bevis

Vi bruker karakteriseringen av intervall i Definisjon 2.1.20.

Hvis $f(I)$ er et punkt, det vil si f er konstant, er setningen oppfylt.

La $y_1, y_2 \in f(I)$ med $y_1 \neq y_2$, og d et vilkårlig reellt tall mellom y_1 og y_2. Vi ønsker å vise at $d \in f(I)$.

Siden $y_1, y_2 \in f(I)$ finnes $x_1, x_2 \in I$ slik at $f(x_1) = y_1$ og $f(x_2) = y_2$. Vi kan ved symmetri anta at $x_1 < x_2$. Ved *skjæringssetningen* anvendt på restriksjonen av f til $[x_1, x_2]$, finnes $c \in [x_1, x_2] \subset I$ slik at $f(c) = d$. Altså er $d \in f(I)$, som er det vi ønsket å vise. $\square$

Tilfellet $d = 0$ av *skjæringssetningen* ble først vist av den bohemske matematikeren, logikeren, filosofen, teologen og katolske presten Bernard Bolzano (1781–1848) i 1817 og er derfor også kjent som *Bolzanos teorem*. Den franske matematikeren og fysikeren Augustin-Louis Cauchy (1789–1857) ga et bevis i sin berømte bok *Cours d'Analyse de l'École Royale Polytechnique* fra 1821. Versjonen vi kjenner i dag tilegnes den tyske matematikeren Karl Weierstrass (1815–1897). Imidlertid har *skjæringssetningen* en tidligere opprinnelse: Den flamske matematikeren, fysikeren og ingeniøren Simon Stevin (1548–1620) beviste *skjæringssetningen* for polynomer ved å angi en algoritme for desimalutviklingen av nullpunkter, lenge før man hadde en rigorøs definisjon av de reelle tallene.

Betingelsene i skjæringssetningen. For at konklusjonen i *skjæringssetningen* skal gjelde, er det helt essensielt at funksjonen er *kontinuerlig*. Vi overlater til Oppgave 9.28 å gi et eksempel på at konklusjonen ikke trenger å gjelde for en diskontinuerlig funksjon. Det er også essensielt at funksjonen er kontinuerlig på et *intervall*, som følgende eksempel viser:

Eksempel 9.6.3. Funksjonen gitt ved $f(x) = \frac{1}{x}$ har definisjonsmengde $\mathbb{R} \setminus \{0\} = (-\infty, 0) \cup (0, \infty)$, en union av to intervaller, og f er kontinuerlig. *Skjæringssetningen* gjelder for eksempel ikke på intervallet $[-1, 1]$, siden $f(-1) = -1 < 0 < 1 = f(1)$, men det finnes ikke noe punkt c slik at $f(c) = 0$.

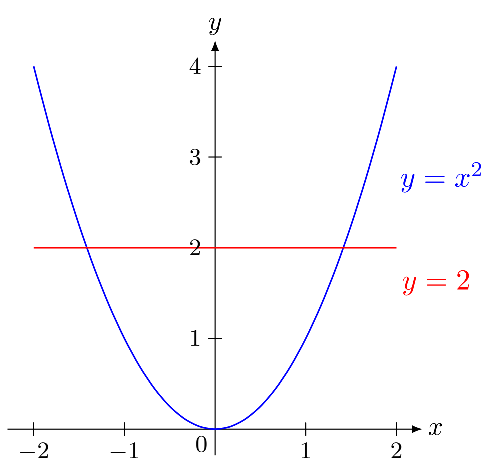
Merk også at grafen til f kun er sammenhengende over de to intervallene $(-\infty, 0)$ og $(0, \infty)$, men ikke på hele definisjonsmengden.

Det er også helt essensielt at vi jobber med intervaller i de *reelle tallene*. *Skjæringssetningen* virker for noen til å være så opplagt at de ikke synes den trenger å bevises: selvfølgelig vil en graf til en kontinuerlig funksjon som går fra den ene siden til den andre siden av den horisontale linjen $y = c$ måtte krysse linjen! Men i denne "opplagte" observasjonen ligger det faktum at vi er så vant til at de reelle tallene oppfyller kompletthetsprinsippet at vi tar det for gitt. Følgende eksempel viser at konklusjonen ikke holder hvis vi ikke hadde "oppfunnet" (eller burde vi si "oppdaget"?) de reelle tallene:

EKSEMPEL 9.6.4. Betrakt funksjonen

$$\begin{aligned} f : \mathbb{Q} &\longrightarrow \mathbb{Q} \\ x &\to x^2. \end{aligned}$$

Visualiserer vi som vanlig de rasjonale tallene i definisjonsmengden langs en tenkt horisontal x-akse og de rasjonale tallene i verdimengden langs en tenkt vertikal y-akse, kan vi tegne grafen $y = x^2$ til f, sammen med den horisontale linjen $y = 2$ i "det rasjonale planet":



GRAFEN TIL f I EKSEMPEL 9.6.4 I BLÅTT, LINJEN $y = 2$ I RØDT, BEGGE DEFINERT OVER $\mathbb{Q}$.

Poenget her er at *tetthetsegenskapen* til $\mathbb{Q}$ (Setning 7.2.2) gjør at begge aksene, samt kurvene $y = x^2$ og $y = 2$, virker som sammenhengende kurver for det blotte øyet.

Det virker opplagt at de to grafene skjærer hverandre. Vi har

$$0 = f(0) < 2 < f(2) = 4,$$

så hvis skjæringsetningen gjaldt på "det rasjonale intervallet" $[0,2] \cap \mathbb{Q}$, så ville det finnes en $c \in \mathbb{Q}$ slik at $0 < c < 2$ og $f(c) = 2$, det vil si $c^2 = 2$. Men, som vi har sett, finnes en slik c ikke, ved Setning 1.3.5.

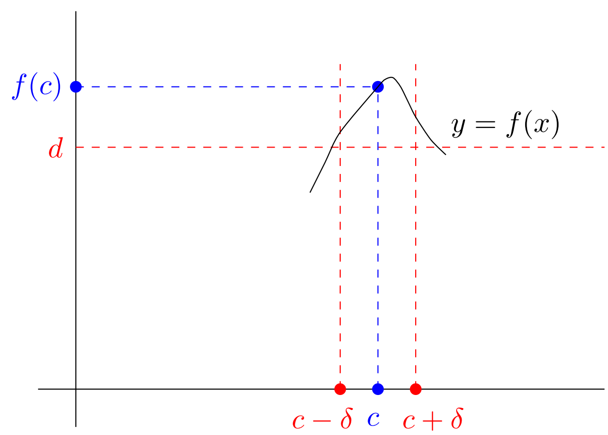
Beviset for skjæringssetningen. Før vi går løs på beviset for *skjæringssetningen* tar vi med en liten hjelpesetning som virker temmelig uskyldig, men som er helt sentral i beviset og som vi også kommer til å få bruk for senere, nemlig i beviset for *ekstremalverdisetningen* (Teorem 9.7.2). Hjelpesetningen stemmer overens med det vi intuitivt forstår med kontinuitet i et punkt, nemlig at funksjonsverdiene eller grafen til funksjonen ikke gjør plutselige hopp i punktet. Setningen sier at dersom for eksempel grafen i punktet ligger over en horisontal linje, da gjør den det også i en (liten nok) omegn om punktet:

Hjelpesetning 9.6.5

Anta at $f : D(f) \to \mathbb{R}$, med $D(f) \subset \mathbb{R}$, er kontinuerlig i $c \in D(f)$ og at $f(c) > d$ (henholdsvis $f(c) < d$) for en $d \in \mathbb{R}$. Da finnes en $\delta > 0$ slik at

$$f(x) > d \ \text{ (hhv. } f(x) < d) \ \text{ for alle } \ x \in (c-\delta, c+\delta) \cap D(f).$$

Bevis
Overlates til Oppgave 9.22. □



Illustrasjon til Hjelpesetning 9.6.5

Merk at vi i hjelpesetningen kun krever at f skal være kontinuerlig i punktet c, ikke nødvendigvis i noe annet punkt!

Vi er nå klare for å bevise *skjæringssetningen*:

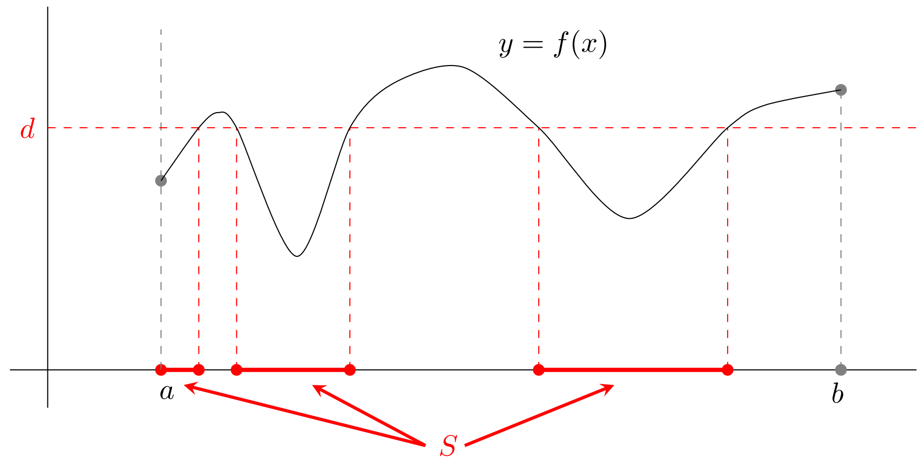
Bevis for Teorem 9.6.1 (Skjæringssetningen)

Vi betrakter tilfellet $f(a) < d < f(b)$. Beviset i det motsatte tilfellet $f(a) > d > f(b)$ er helt tilsvarende.

Betrakt mengden

$$S = \{x \in [a, b] \mid f(x) \leq d\},$$

se figur:



MENGDEN S I BEVISET FOR TEOREM 9.6.1

Vi har $a \in S$, og dermed $S \neq \emptyset$. I tillegg er b en øvre skranke for S. Dermed eksisterer den minste øvre skranken $\sup S$ ved kompletthetsegenskapen til den reelle tallene. Vi vil vise at $c = \sup S$ oppfyller den ønskede egenskapen at $f(c) = d$. Dette viser vi ved å utlede en motsigelse fra hvert av tilfellene $f(c) > d$ og $f(c) < d$. Da er eneste gjenstående mulighet (ved trikotomiegenskapen til ordenaksiomene, jf. Definisjon 3.6.9) at $f(c) = d$.

Anta at $f(c) > d$. Da er $c \neq a$, det vil si $c > a$ (mens vi ikke kan utelukke tilfellet $c = b$). Siden f er kontinuerlig i c, finnes det ved Hjelpesetning 9.6.5 en $\delta > 0$ slik at

$$f(x) > d \text{ når } x \in (c - \delta, c + \delta) \cap [a, b]. \tag{235}$$

Siden $c > a$, er $\delta' = \min\{\delta, c - a\} > 0$. Da er $(c - \delta', c] \subset (c - \delta, c + \delta) \cap [a, b]$. Da medfører (235) at

$$f(x) > d \text{ når } x \in (c - \delta', c].$$

Men dette betyr at $c - \delta'$ er en øvre skranke for S mindre enn $c = \sup S$, en motsigelse.

Anta at $f(c) < d$. Da er $c \neq b$, det vil si $c < b$ (mens vi ikke kan utelukke tilfellet $c = a$). Siden f er kontinuerlig i c, finnes det ved Hjelpesetning 9.6.5 en $\delta > 0$ slik at

$$f(x) < d \text{ når } x \in (c - \delta, c + \delta) \cap [a, b]. \tag{236}$$

Siden $c < b$, er $\delta' = \min\{\delta, b - c\} > 0$. Da er $[c, c + \delta') \subset (c - \delta, c + \delta) \cap [a, b]$. Da medfører (236) at

$$f(x) < d \text{ når } x \in [c, c + \delta').$$

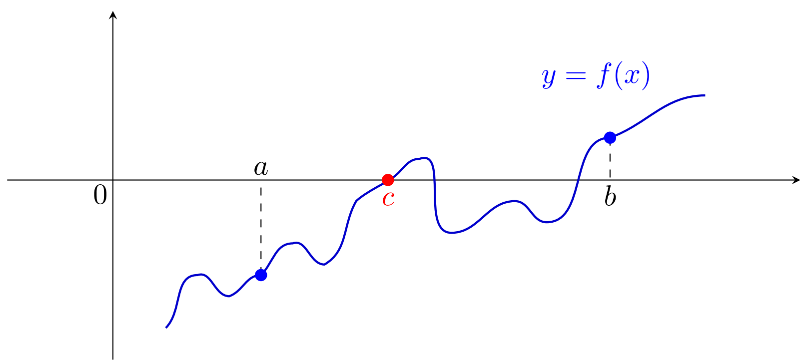
Dette motsier at c er en øvre skranke for S.

Eneste gjenværende mulighet er derfor at $f(c) = d$, og beviset er fullført. $\square$

Anvendelser av skjæringssetningen. Veldig ofte bruker vi spesialtilfellet til *skjæringssetningen* med $d = 0$ til å konkludere at en funksjon f har et nullpunkt, som vist i neste figur:

Følgesetning (av skjæringssetningen) 9.6.6

Dersom f har motsatt fortegn i a og b og f er kontinuerlig i $[a, b]$, da har f et nullpunkt i (a, b), det vil si det finnes et punkt $c \in (a, b)$ der $f(c) = 0$.

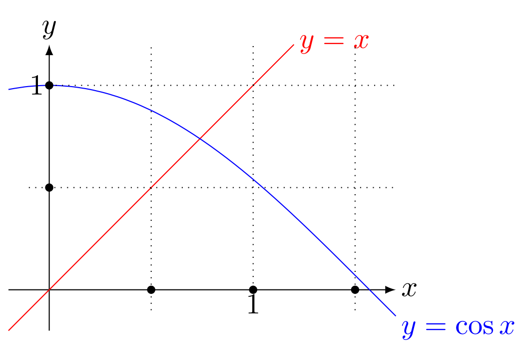


Illustrasjon til Følgesetning 9.6.6.

Mer generelt kan *skjæringssetningen* brukes til å vise at ligninger har en løsning, som i neste eksempel.

Eksempel 9.6.7. Vis at ligningen $\cos x = x$ har en løsning.

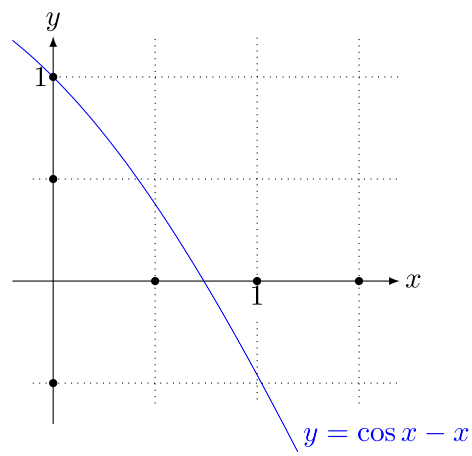
Løsning. La oss tenke litt geometrisk først: tegner vi grafene $y = \cos x$ og $y = x$, ser vi at de virker å krysse hverandre, siden grafene veksler mellom å være størst i for eksempel 0 og 1:



Grafene $y = \cos x$ i blått og $y = x$ i rødt.

Siden dette er grafer til kontinuerlige funksjoner, er det fristende å si at "de må skjære hverandre ved skjæringssetningen". Men *skjæringssetningen* sier ikke direkte dette, så vi må omformulere problemet slik at vi kan bruke den.

Ligningen kan også skrives som $\cos x - x = 0$, slik at problemet er ekvivalent med å vise at funksjonen $f(x) = \cos x - x$ har et nullpunkt. Dette er en kontinuerlig funksjon ved Setning 9.5.1, siden både $x \mapsto \cos x$ (Følgesetning 9.5.10) og $x \mapsto x$ er det. Vi har $f(0) = 1$ og $f\left(\frac{\pi}{2}\right) = -\frac{\pi}{2} < 0$. Ved *skjæringssetningen* (eller Følgesetning 9.6.6) finnes det et tall $c \in \left[0, \frac{\pi}{2}\right]$ slik at $f(c) = 0$. Altså er $\cos c - c = 0$, slik at c en løsning på ligningen $\cos x = x$. $\square$



GRAFEN $y = \cos x - x$.

Vi merker at vi ikke bare har vist at det finnes en løsning, men at vi også har klart å lokalisere løsningen til intervallet $\left[0, \frac{\pi}{2}\right]$ (faktisk til det *åpne* intervallet $\left(0, \frac{\pi}{2}\right)$, siden vi regnet ut at endepunktene ikke var nullpunkter til f). Vi kunne også lokalisere løsningen mer presist: for eksempel vet vi at $f(1) = \cos 1 - 1 < 1 - 1 = 0$, slik at vi kan si at $c \in (0, 1)$.

Vi merker oss fra forrige eksempel at den generelle strategien for å vise at en ligning

$$g(x) = h(x)$$

har en løsning er:

- "flytt over" $h(x)$ og skriv ligningen som $g(x) - h(x) = 0$;
- definér en ny funksjon f ved $f(x) = g(x) - h(x)$;
- finn to punkter a og b der f har motsatt fortegn og slik at f er kontinuerlig på $[a, b]$ (hvilket den vil være dersom både g og h er kontinuerlige).
- da gir *skjæringssetningen* at f har et nullpunkt, og dermed at ligningen $g(x) = h(x)$ har en løsning, på $[a, b]$.

Strategien ovenfor kan oppsummeres i følgende konsekvens av *skjæringssetningen*, som viser at grafene til to kontinuerlige funksjoner som krysser hverandre må ha et skjæringspunkt:

Følgesetning (av skjæringssetningen) 9.6.8

Anta at $g : [a, b] \to \mathbb{R}$ og $h : [a, b] \to \mathbb{R}$ er to kontinuerlige funksjoner slik at $g(a) < h(a)$ og $g(b) > h(b)$. Da finnes en $c \in (a, b)$ slik at $g(c) = h(c)$.

Bevis
Overlates til Oppgave 9.29. □

Det neste naturlige spørsmålet er om ligningen $g(x) = h(x)$ kan ha flere enn én løsning, for eksempel på et begrenset intervall. Dette er ekvivalent med spørsmålet om f har flere en ett nullpunkt. I tilfellet studert i forrige eksempel kan vi bruke monotoniegenskaper til å avgjøre spørsmålet:

Eksempel 9.6.9. Vis at ligningen $\cos x = x$ har høyst én løsning.

Løsning. Vi vil vise det ekvivalente utsagnet at funksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved $f(x) = \cos x - x$ har høyst ett nullpunkt. Vi deler problemet i fire deler:

- $0 \leq x \leq \frac{\pi}{2}$: her er funksjonen $x \mapsto \cos x$ strengt avtagende (tenk på hvordan den er definert geometrisk), og det samme er funksjonen $x \mapsto -x$. Funksjonen f er derfor strengt avtagende og dermed injektiv (jf. Merknad 7.2.8). Derfor kan f ha høyst ett nullpunkt.
- $x \geq \frac{\pi}{2}$: da er $f(x) = \cos x - x \leq 1 - \frac{\pi}{2} < 0$, slik at f ikke kan ha nullpunkter.
- $-\frac{\pi}{2} \leq x < 0$: da er $f(x) = \cos x - x > 0 + 0 = 0$, slik at f igjen ikke kan ha nullpunkter.
- $x < -\frac{\pi}{2}$: da er $f(x) = \cos x - x > -1 + \frac{\pi}{2} > 0$, slik at f igjen ikke kan ha nullpunkter.

□

Selv om vi gjerne er vant til fra skolen at vi kan finne løsninger på ligninger ved regning, så er ikke dette mulig for de aller fleste ligningene. Vi kan imidlertid, ved hjelp av datamaskiner, finne tilnærmingsverdier for løsninger men den grad av nøyaktighet vi måtte ønske. Vi viser til §9.8 (eller [**AE**, §4.2], [**Li**, §7.3])) for metoder for å finne tilnærmingsverdier til løsninger på ligninger, eller ekvivalent, nullpunkter til funksjoner.

Vi legger til noen flere eksempler på bruk av *skjæringssetningen*.

Vi starter med å gi en annerledes løsning på Oppgave 8.25:

Eksempel 9.6.10. Vis at et reelt polynom av odde grad alltid har en reell rot.

Løsning. Under våre antagelser kan vi skrive polynomet som

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

der $a_n, \dots, a_0 \in \mathbb{R}$, $a_n \neq 0$ og n er et oddetall. Ved Følgesetning 9.5.7 er funksjonen

$$\begin{array}{rcl} P:\mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \mapsto & P(x) \end{array}$$

kontinuerlig.

Vi kan skrive

$$P(x) = x^n \left(a_n + \frac{a_{n-1}}{x} + \cdots + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n} \right).$$

For stor $|x|$ kan vi gjøre leddene $\frac{a_{n-1}}{x}, \dots, \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n}$ så små vi vil i absoluttverdi, og $x^n a_n$ så stor vi vil i absoluttverdi. (Vi hopper over en del detaljer og viser til Oppgave 8.18). Mer spesifikt, gitt et hvilket som helst tall $M \in \mathbb{R}^+$, kan vi finne en x_1 med samme fortegn som a_n slik at $P(x_1) \geq M > 0$ og en x_2 slik at $P(x_2) \leq -M < 0$. Ved Følgesetning 9.6.6 finnes en $c \in (x_1, x_2)$ slik at $f(c) = 0$. $\square$

Løsningen ovenfor viser at gitt et hvilket som helst tall $M \in \mathbb{R}^+$, kan vi finne en x_1 med samme fortegn som a_n slik at $P(x_1) \geq M > 0$ og en x_2 slik at $P(x_2) \leq -M < 0$. Siden vi også viste at P har et nullpunkt, vil det ved *skjæringssetningen* også finnes et punkt a slik at $P(a) = M$ og et punkt b slik at $P(b) = -M$. Siden M er vilkårlig, viser dette at verdimengden til $P : \mathbb{R} \to \mathbb{R}$ er hele $\mathbb{R}$. Som nevnt i introduksjonen til seksjonen, er dette en bruk av *skjæringssetningen* vi ikke funderer så ofte over. Vi viser et eksempel til på dette:

EKSEMPEL 9.6.11. Vis at verdimengden til tangensfunksjonen restrisert til $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$,

$$\begin{array}{rcl} f:\left(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right) & \longrightarrow & \mathbb{R} \\ x & \mapsto & \tan x = \dfrac{\sin x}{\cos x}, \end{array}$$

er hele $\mathbb{R}$.

LØSNING. La $M \in \mathbb{R}$. Vi vil vise at f antar verdien M.

Vi antar for enkelhets skyld at $M \geq 0$. Tilfellet $M \leq 0$ vises på helt tilsvarende måte.

Siden $f(0) = 0$ og f er kontinuerlig ved Følgesetning 9.5.10, holder det ved *skjæringssetningen* å vise at det finnes en x slik at $f(x) \geq M$. Vi merker at $\sin x \geq \frac{1}{2}$ og $\cos x > 0$ når $x \in \left[\frac{\pi}{6}, \frac{\pi}{2}\right]$, slik at

$$f(x) = \frac{\sin x}{\cos x} \geq \frac{1}{2\cos x} \quad \text{for} \quad x \in \left[\frac{\pi}{6}, \frac{\pi}{2}\right).$$

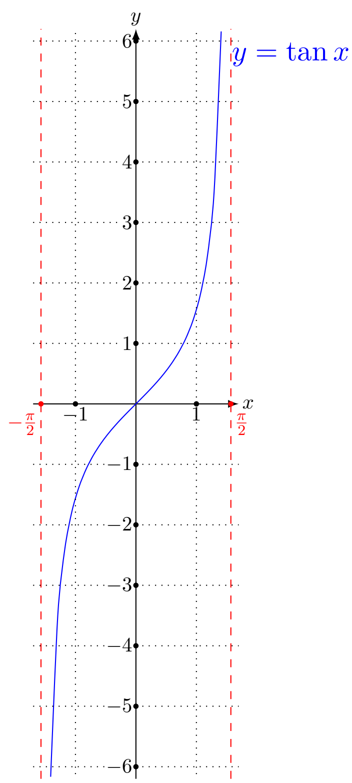
Siden $x \mapsto \cos x$ er kontinuerlig i $\frac{\pi}{2}$ (Følgesetning 9.5.10), finnes det en $\delta > 0$ slik at

$$|\cos x| = \left|\cos x - \cos\frac{\pi}{2}\right| < \frac{1}{2M} \quad \text{når} \quad x \in \left(\frac{\pi}{2} - \delta, \frac{\pi}{2} + \delta\right).$$

La $\delta' = \min\{\delta, \frac{\pi}{2} - \frac{\pi}{6}\}$. Da har vi, for $x \in \left(\frac{\pi}{2} - \delta', \frac{\pi}{2}\right)$, at:

$$f(x) = \frac{\sin x}{\cos x} \geq \frac{1}{2\cos x} > \frac{1}{2 \cdot \frac{1}{2M}} = M,$$

som er det vi ønsket å vise. □



Grafen til funksjonen f i Eksempel 9.6.11.

Merknad 9.6.12. Siden $x \mapsto \sin x$ og $x \mapsto \cos x$ er, henholdsvis, strengt voksende og strengt avtagende på $\left[0, \frac{\pi}{2}\right)$, og begge ikke-negative, er funksjonen $x \mapsto \tan x$ strengt voksende på $\left[0, \frac{\pi}{2}\right)$. Siden

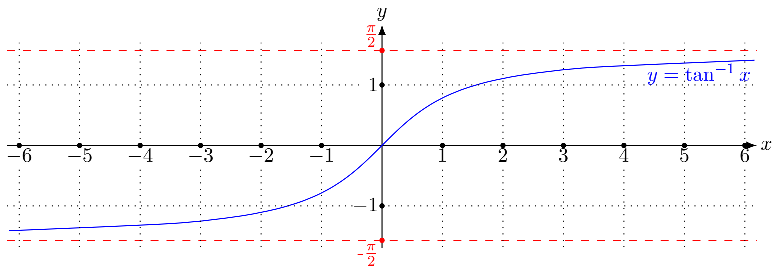
$$\tan(-x) = \frac{\sin(-x)}{\cos(-x)} = \frac{\sin(-x)}{\cos x} = -\tan x,$$

er $x \mapsto \tan x$ strengt voksende på hele $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Funksjonen f fra forrige eksempel er derfor bijektiv (jf. Merknad 7.2.8), med inversfunksjon $f^{-1} : \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ som assosierer til et tall $x \in \mathbb{R}$ det entydige tallet

$\theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ slik at $f(\theta) = \tan\theta = x$. Vi betegner gjerne θ med $\tan^{-1} x$ eller $\arctan x$, slik at den inverse funksjonen kan beskrives ved regelen

$$\begin{aligned} f^{-1} : \mathbb{R} &\longrightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \\ x &\mapsto \tan^{-1} x = \arctan x. \end{aligned}$$

Denne funksjonen kalles *den inverse tangensfunksjonen* eller også *arcus-tangens-funksjonen*. (Se også Eksempel 2.3.10.) Som bemerket i Observasjon 2.3.7 er grafen til f^{-1} speilbildet av grafen til f om linjen $y = x$.



GRAFEN TIL FUNKSJONEN $x \mapsto \tan^{-1} x$ I MERKNAD 9.6.12.

Vi har allerede bemerket at en strengt monoton funksjon er injektiv (jf. Merknad 7.2.8). Vi kan vise at det motsatte er sant for en *kontinuerlig funksjon definert på et intervall*:

Setning 9.6.13

La $f : I \to \mathbb{R}$ være en kontinuerlig funksjon definert på et intervall $I \subset \mathbb{R}$. Da er f injektiv hvis og bare hvis f er strengt monoton.

BEVIS
Det gjenstår kun å vise implikasjonen "$\Longrightarrow$", det vil si "bare hvis"-påstanden. Vi argumenterer ved selvmotsigelse. Anta f er injektiv, men er verken strengt voksende eller strengt avtagende. Da vil det finnes tre punkter $x_1 < x_2 < x_3$ i $D(f)$ slik at

(i) $f(x_1) < f(x_2)$ og $f(x_2) > f(x_3)$, eller
(ii) $f(x_1) > f(x_2)$ og $f(x_2) < f(x_3)$.

Siden f er injektiv er $f(x_1) \neq f(x_3)$, og dette medfører at begge tilfellene består av to undertilfeller, alt etter om $f(x_3) < f(x_1)$ eller omvendt:

(i-a) $f(x_1) < f(x_3) < f(x_2)$,
(i-b) $f(x_3) < f(x_1) < f(x_2)$,
(ii-a) $f(x_1) > f(x_3) > f(x_2)$,
(ii-b) $f(x_3) > f(x_1) > f(x_2)$.

I tilfellene (i-a) og (ii-a) finnes det ved *skjæringssetningen* en $c \in (x_1, x_2)$ slik at $f(c) = f(x_3)$, som motsier at f er injektiv.

I tilfellene (i-b) og (ii-b) finnes det ved *skjæringssetningen* en $c \in (x_2, x_3)$ slik at $f(c) = f(x_1)$, som igjen motsier at f er injektiv. □

Følgesetning 9.6.14: Kontinuitet av inverse funksjoner, II

La $f : I \to \mathbb{R}$ være en kontinuerlig injektiv funksjon definert på et intervall I. Da er $f(I)$ igjen et intervall og den inverse funksjonen $f^{-1} : f(I) \to I$ er kontinuerlig.

Bevis

At $f(I)$ er et intervall følger fra Følgesetning 9.6.2. Siden f er strengt monoton ved Setning 9.6.13, gir Setning 9.5.4 at f^{-1} er kontinuerlig. □

Vi avslutter seksjonen med et klassisk eksempel av mer anvendt karakter:

Eksempel 9.6.15 ("Meteorologiteoremet"). Til enhver tid finnes det to punkter langs ekvator på motsatt side av jordkloden med samme temperatur. (To punkter på motsatt side av en kule kalles gjerne *antipodale*.)

Løsning. Vi tenker oss at vi ser på jordkloden ovenfra slik at ekvator ser ut som en sirkel og parametriserer punktene på ekvator som vi til vanlig parametriserer punktene på en sirkel, nemlig med en vinkel $\theta \in [0, 2\pi)$ ved å velge et eller annet punkt langs ekvator til å tilsvare $\theta = 0$. Vi lar $\theta = 0$ og 2π parametrisere samme punkt langs ekvator og definerer funksjonen

$$T : [0, 2\pi] \longrightarrow \mathbb{R}$$

som gir temperaturen $T(\theta)$ i punktet θ langs ekvator. Spesielt er

$$T(0) = T(2\pi). \tag{237}$$

Vi gjør den rimelige antagelsen at T er en *kontinuerlig funksjon*. (Slike antagelser er typiske og helt vanlige i naturvitenskap.) Merk at to punkter $\theta_1, \theta_2 \in [0, 2\pi]$ med $\theta_1 < \theta_2$ er antipodale dersom $\theta_2 = \theta_1 + \pi$. Definér funksjonen

$$f : [0, \pi] \longrightarrow \mathbb{R}$$

ved

$$f(\theta) = T(\theta) - T(\theta + \pi).$$

Siden T er kontinuerlig og funksjonen $\theta \mapsto \theta + \pi$ er kontinuerlig (ved Følgesetning 9.5.7), er funksjonen $\theta \mapsto T(\theta + \pi)$ kontinuerlig ved Setning 9.5.2. Da er f en differanse av kontinuerlige funksjoner og dermed kontinuerlig ved Setning 9.5.1. Vi har

$$f(0) = T(0) - T(\pi)$$

og

$$f(\pi) = T(\pi) - T(2\pi) = T(\pi) - T(0) = -f(0).$$

Ved *skjæringssetningen* i varianten i Følgesetning 9.6.6 finnes en $c \in [0, \pi]$ slik at $f(c) = 0$, det vil si

$$T(c) - T(c + \pi) = 0.$$

Dette betyr at $T(c) = T(c + \pi)$. Altså er c det ønskede punktet. □

MERKNAD 9.6.16. *Meteorologiteoremet* er et spesialtilfelle av et mer generelt teorem innenfor en gren av matematikken som heter *topologi* (jf. §10.10), nemlig *Borsuk-Ulam teoremet*, oppkalt etter den polske matematikeren Karol Borsuk (1905–1982), som beviste teoremet i 1933, og den polske matematikeren Stanisław Marcin Ulam (1909–1984), som først formulerte det. I sin "populærvitenskapelige" form lyder teoremet slik:

Det finnes alltid to antipodale punkter på jordkloden med samme trykk og temperatur.

Det finnes flere slike morsomme eksempler på bruk av *skjæringssetningen*:

- **Et brød kan med ett enkelt vertikalt knivkutt alltid deles i to nøyaktig like store deler.**

  Start med kniven i posisjon til venstre for brødet. Da er hele brødet til høyre for kniven. Volumet brød til høyre for kniven er derfor likt hele brødvolumet V. Flytt kniven kontinuerlig horisontalt mot høyre. Til slutt er hele brødet til venstre for kniven, det vil si volumet brød til høyre for kniven er 0. Da må det ved *skjæringssetningen* være et punkt der volumet brød til høyre for kniven er $\frac{V}{2}$, det vil si halve brødvolumet.
- **En sandwich med ost kan med ett enkelt vertikalt knivkutt alltid deles i to deler som har nøyaktig samme andel brød og ost.**

  Velg en bestemt vinkel θ. Det forrige punktet viser at vi kan dele brødet i to nøyaktig like store deler med et vertikalt kutt i vinkel θ (sett ovenfra). Variér vinkelen θ kontinuerlig mellom 0 og π, slik at posisjonen i π er den samme som for 0 men i motsatt retning. Anta at det ved kuttet i vinkel 0 finnes *mer* enn halvparten av osten på si høyre side av kniven. Da vil det ved å kutte i vinkel π være *mindre* enn halvparten av osten på høyre side, siden kniven da ligger i motsatt retning. Ved *skjæringssetningen* vil det finnes en vinkel i $[0, \pi]$ der osten på høyre side var lik halvparten av den totale osten.
- **The Ham Sandwich Theorem: En sandwich med ost og skinke kan med ett enkelt knivkutt alltid deles i to deler som har nøyaktig samme andel brød, ost og skinke.**

  Poenget her er at vi ikke krever *vertikale knivkutt*, som i de to forrige punktene, det vil si vi tillater oss å ha helning på kniven. Da oppnår vi en frihetsgrad til. Detaljene overlates til leseren.

Nylig har *skjæringssetningen* blitt brukt til å vise at man alltid kan stabilisere et vaklende bord ved å flytte på bordet, uten å måtte legge servietter og lignende under bordbenene, jf. [**BLPR**], eller

https://www.maa.org/external_archive/devlin/devlin_02_07.html

for en mer populærvitenskapelig fremstilling.

9.7. Ekstremalverdisetningen

Vi starter med en definisjon som skal være kjent fra skolen:

Definisjon 9.7.1: Ekstremalpunkter og -verdier

La $f : D(f) \to \mathbb{R}$, med $D(f) \subset \mathbb{R}$, være en funksjon.

Et punkt $a \in D(f)$ er et *maksimumspunkt*[a] for f dersom $f(a) \geq f(x)$ for alle $x \in D(f)$. Funksjonsverdien $f(a)$ kalles da *maksimalverdien*[b] til f.

Et punkt $a \in D(f)$ er et *minimumspunkt*[c] for f dersom $f(a) \leq f(x)$ for alle $x \in D(f)$. Funksjonsverdien $f(a)$ kalles da *minimalverdien*[d] til f.

Et punkt a som er enten maksimums- eller minimumspunkt kalles et *ekstremalpunkt*[e]. Funksjonsverdien $f(a)$ kalles da en *ekstremalverdi*[f] til f.

[a] *Maximum point* på engelsk.
[b] *Maximal value* på engelsk.
[c] *Minimum point* på engelsk.
[d] *Minimal value* på engelsk.
[e] *Extremal point* på engelsk.
[f] *Extremal value* på engelsk.

Fra skolen er vi er vant til å finne *eventuelle* maksimums- og minimumspunkter til funksjoner ved å sjekke punkter der den deriverte er null eller ikke eksisterer. Men kan vi alltid være sikre på at slike ekstremalverdier finnes? Svaret er “Nei”, det finnes enkle eksempler på funksjoner som ikke har ekstremalverdier (se for eksempel Eksemplene 9.7.4-9.7.6 nedenunder). Imidlertid gir det svært viktige neste teoremet, *ekstremalverdisetningen*[3], et beroligende svar: Kontinuerlige funksjoner definert på lukkede, begrensede intervaller har alltid ekstremalpunkter:

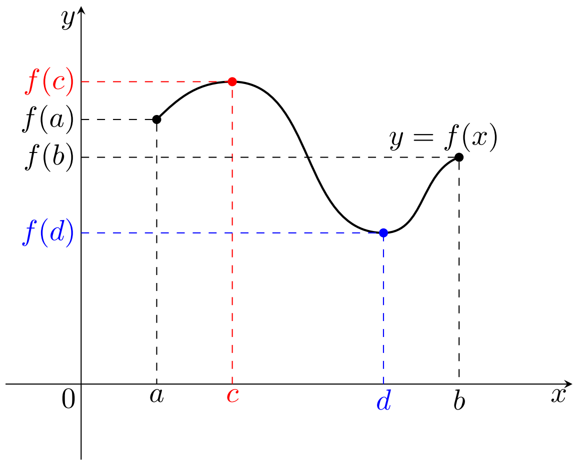
Teorem 9.7.2: Ekstremalverdisetningen

La $f : [a, b] \to \mathbb{R}$ være en kontinuerlig funksjon. Da finnes $c, d \in [a, b]$ slik at

$$f(d) \leq f(x) \leq f(c) \quad \text{for alle} \quad x \in [a, b].$$

[3] *Extreme value theorem* eller *Max-Min theorem* på engelsk.

Merk at teoremet sier både at det *finnes* to verdier slik at $f(x)$ må ligge mellom disse verdiene for alle $x \in [a, b]$ *og* at f også *antar* disse ytterverdiene. Vi merker også at *ekstremalverdisetningen* bare garanterer at det *finnes* ekstremalverdier, men på samme måte som *skjæringssetningen* sier den ingenting om hvordan vi kan bestemme disse punktene. For å eksplisitt lokalisere disse verdiene kan vi for eksempel bruke derivasjon, noe vi ikke skal komme nærmere innpå her; vi viser til [**AE**, §4.4] og [**Li**, §6.4].



ILLUSTRASJON TIL EKSTREMALVERDISETNINGEN: c OG d ER EKSTREMALPUNKTER

Ekstremalverdisetningen ble også først vist av Bolzano i 1830-årene, men arbeidet forble upublisert frem til 1930. Resultatet ble senere bevist igjen av Weierstrass.

MERKNAD 9.7.3. Setningen holder opplagt også for funksjoner som har som definisjonsmengde en *endelig union av lukkede, begrensede intervaller* (jf. [**AE**, Teorem 5 i §4.4]).

Betingelsene i ekstremalverdisetningen. Teoremet holder ikke dersom vi fjerner antagelsen om at intervallet er lukket og begrenset eller at funksjonen er kontinuerlig, som de neste eksemplene viser:

EKSEMPEL 9.7.4 (Kontinuerlig funksjon på ikke-lukket intervall). Betrakt funksjonen gitt ved $f(x) = x$ på $[0, 1)$. Da vil $f(x)$ komme så nær vi vil 1 (nedenfra) men aldri være lik 1. Verdimengden er $[0, 1)$ (ved *skjæringssetningen*), men det finnes ingen $x \in [0, 1)$ slik at $f(x) = 1$. Altså antar ikke f noen maksimalverdi (men den antar minimalverdien 0 i 0).

EKSEMPEL 9.7.5 (Kontinuerlig funksjon på ikke-begrenset intervall). Se på funksjonen gitt ved $f(x) = x$ på $[0, \infty)$. Denne funksjonen antar ingen maksimalverdi siden x kan bli så stor vi vil. Verdimengden er $V(f) = [0, \infty)$, igjen ved *skjæringssetningen*.

EKSEMPEL 9.7.6 (Diskontinuerlig funksjon). La $f : [0, 1] \to \mathbb{R}$ være gitt ved

$$f(x) = \begin{cases} x, & x \in [0, 1), \\ 0, & x = 1. \end{cases}$$

Igjen vil $f(x)$ komme så nær vi vil 1 (nedenfra) men aldri være lik 1. Verdimengden er $[0, 1)$ (ved *skjæringssetningen*)

Beviset for ekstremalverdisetningen via begrensningsteoremet. Det er vanlig å bevise *ekstremalverdisetningen* i to trinn. Første trinn går ut på å vise at en funksjon som oppfyller kriteriene i setningen er *begrenset*:

Definisjon 9.7.7: Begrenset funksjon

La X være en mengde. En funksjon $f : X \to \mathbb{R}$ kalles *begrenset*[a] dersom det finnes en $M \in \mathbb{R}$ slik at $|f(x)| \leq M$ for alle $x \in X$.

Funksjonen f kalles *oppad begrenset*[b] (henholdsvis *nedad begrenset*[c]) dersom det finnes en $M \in \mathbb{R}$ slik at $f(x) \leq M$ (hhv. $f(x) \geq M$) for alle $x \in X$.

[a] *Bounded* på engelsk.
[b] *Bounded above* på engelsk.
[c] *Bounded below* på engelsk.

Merk at definisjonene er konsistente med de tilsvarende definisjonene for følger (som jo er spesialtilfeller av funksjoner), jf. Definisjon 7.4.2. Merk også følgende:

Hjelpesetning 9.7.8

En funksjon er begrenset hvis og bare hvis den er både oppad og nedad begrenset.

BEVIS
Hvis $f : X \to \mathbb{R}$ er begrenset, finnes en $M \in \mathbb{R}$ slik at $|f(x)| \leq M$ for alle $x \in X$, og dette er ekvivalent med at $-M \leq f(x) \leq M$ for alle $x \in X$, som viser at f både er oppad og nedad begrenset.

Motsatt, hvis $f : X \to \mathbb{R}$ er både oppad og nedad begrenset, finnes $m, M \in \mathbb{R}$ slik at $m \leq f(x) \leq M$ for alle $x \in X$. Dette medfører at $|f(x)| \leq \max\{|m|, |M|\}$ for alle $x \in X$, som viser at f er begrenset. $\square$

Teorem 9.7.9: Begrensningsteoremet

La $f : [a, b] \to \mathbb{R}$ være en kontinuerlig funksjon. Da er f begrenset.

Bevis

Vi bruker Hjelpesetning 9.7.8 og viser at f både er oppad og nedad begrenset.

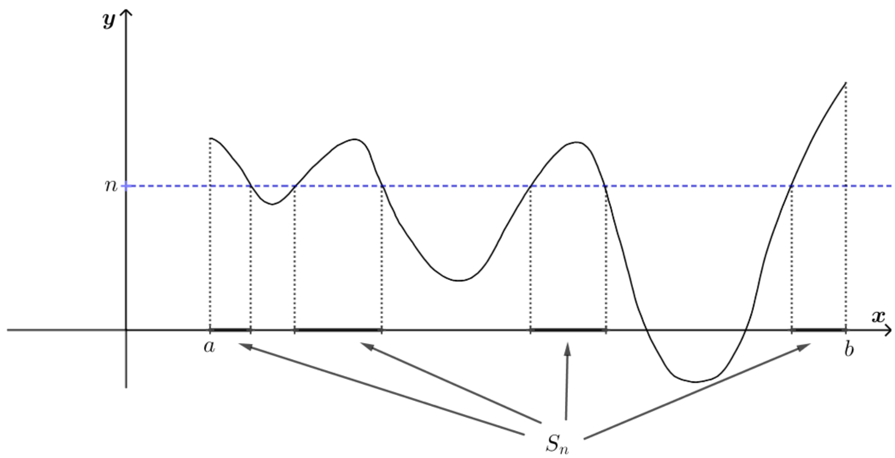
Vi viser først at f er oppad begrenset.

Idéen er å vise at vi klarer å "presse" grafen til f under en horisontal linje $y = n$ for en stor nok $n \in \mathbb{Z}^+$.

Definér derfor

$$S_n = \{x \in [a, b] \mid f(x) > n\}$$

for enhver $n \in \mathbb{Z}^+$, som vist på figuren:



Mengden S_n i beviset for begrensningsteoremet.

Vi ønsker at $S_n = \emptyset$ for en eller annen (stor nok) n, for dette vil medføre at $f(x) \leq n$ for alle $x \in [a, b]$ og vi vil være ferdig.

Vi argumenterer ved selvmotsigelse og antar derfor at $S_n \neq \emptyset$ for alle $n \in \mathbb{Z}^+$. Hver S_n er nedad begrenset, siden a er en nedre skranke, og dermed vil S_n ved største-nedre-skranke-egenskapen til $\mathbb{R}$ (jf. Teorem 6.1.14) ha en største nedre skranke inf S_n, kall den x_n, som per definisjon oppfyller $x_n \geq a$. Vi har dessuten

$$x_n < b, \tag{238}$$

for hvis $x_n = b$, vil $S_n = \{b\}$ (siden b er en nedre skranke for S_n og $S_n \neq \emptyset$), hvilket betyr at $f(x) \leq n$ for alle $x \in [a, b)$, mens $f(b) > n$, som er umulig ved Hjelpesetning 9.6.5.

Vi vil oppnå den ønskede selvmotsigelsen ved å vise begge utsagnene

$$\{f(x_n)\} \text{ konvergerer,} \tag{239}$$

$$f(x_n) \geq n \text{ for alle } n \in \mathbb{Z}^+; \tag{240}$$

dette er nemlig motstridende utsagn, siden en konvergent følge er begrenset ved Setning 7.4.10.

Vi viser først (239).

Siden $S_{n+1} \subset S_n$ for alle $n \in \mathbb{Z}^+$, er $\inf S_{n+1} = x_{n+1} \geq x_n = \inf S_n$ ved Oppgave 6.3. Følgen $\{x_n\}$ er derfor voksende. Siden også $x_n \leq b$ for alle $n \in \mathbb{Z}^+$, er følgen også oppad begrenset. Den vil derfor konvergere ved Setning 7.5.4(i), si mot L. Siden $a \leq x_n \leq b$ for alle $n \in \mathbb{Z}^+$ og grenser bevarer orden (Setning 7.4.14), er også $a \leq L \leq b$. Dermed er f kontinuerlig i L, slik at $f(x_n) \to f(L)$ ved Setning 9.4.1. Dette viser (239).

Vi viser dernest (240) ved selvmotsigelse. Anta at vi har $f(x_n) < n$ for en $n \in \mathbb{Z}^+$. Siden f er kontinuerlig i x_n, finnes det ved Hjelpesetning 9.6.5 en $\delta > 0$ slik at

$$f(x) < n \quad \text{for alle } x \in (x_n - \delta, x_n + \delta) \cap [a, b]. \tag{241}$$

La $\delta' = \min\{\delta, b - x_n\}$, som oppfyller $\delta' > 0$ ved (238). Da er

$$[x_n, x_n + \delta') \subset (x_n - \delta, x_n + \delta) \cap [a, b],$$

slik at (241) medfører

$$f(x) < n \quad \text{for alle } x \in [x_n, x_n + \delta') \subset [a, b],$$

som viser at $x_n + \delta'$ er en nedre skranke for S_n, hvilket motsier at x_n er den største nedre skranken. Dette er den ønskede selvmotsigelsen som viser (240).

Vi har dermed vist at både (239) og (240) er sanne, som fullfører beviset for at f er oppad begrenset.

For å vise at f også er nedad begrenset, betrakter vi funksjonen $-f : [a, b] \to \mathbb{R}$, som er kontinuerlig ved Setning 9.5.1. Ved det vi nettopp har vist, er denne funksjonen oppad begrenset, hvilket vil si at det finnes en $M \in \mathbb{R}$ slik at $-f(x) \leq M$ for alle $x \in [a, b]$. Men da er $f(x) \geq -M$ for alle $x \in [a, b]$, som viser at f er nedad begrenset.

Dermed har vi vist at f både er oppad og nedad begrenset, og beviset er fullført. $\square$

Før vi går løs på beviset for *ekstremalverdisetningen* er det verdt å merke seg at funksjoner godt kan være begrenset uten å ha ekstremalpunkter, jf. Eksemplene 9.7.4 og 9.7.6.

Bevis for Teorem 9.7.2 (Ekstremalverdisetningen)

Først viser vi eksistensen av *maksimumspunktet* c, det vil si, punktet c som oppfyller $f(c) \geq f(x)$ for alle $x \in [a, b]$.

Ved *begrensningsteoremet* 9.7.9 vet vi at verdimengden $V(f)$ er begrenset, spesielt oppad begrenset. Ved kompletthetsprinsippet til $\mathbb{R}$ har $V(f)$ en minste øvre skranke $M = \sup V(f)$. Vi vil vise at det finnes en $c \in [a, b]$ slik at $f(c) = M$. Da har vi at $f(x) \leq f(c) = M$ for alle $x \in [a, b]$ og vi har funnet det ønskede punktet c.

Vi argumenterer ved selvmotsigelse og antar at det *ikke* finnes et slikt punkt c. Det betyr at $f(x) < M$ for alle $x \in [a, b]$. Ved Setning 9.5.1 er funksjonen $g : [a, b] \to \mathbb{R}$ definert ved

$$g(x) = \frac{1}{M - f(x)}$$

kontinuerlig. Siden $M - f(x) > 0$, er $g(x) > 0$ for alle $x \in [a, b]$. Ved *begrensningsteoremet* 9.7.9 er funksjonen begrenset, spesielt vil det finnes en $K \in \mathbb{R}^+$ slik at

$$g(x) = \frac{1}{M - f(x)} \leq K \text{ for alle } x \in [a, b].$$

Siden $M - f(x) > 0$ kan vi multiplisere med $M - f(x)$ og bevare ulikheten, som gir at $1 \leq K(M - f(x))$, som igjen kan omskrives til $f(x) \leq M - \frac{1}{K}$ for alle $x \in [a, b]$. Men dette gir at $M - \frac{1}{K}$ er en øvre skranke for $V(f)$ mindre enn den minste øvre skranken M, som er den ønskede motsigelsen.

Dette fullfører beviset for eksistensen av maksimumspunktet c.

For å vise eksistensen av minimumspunktet d, betrakter vi funksjonen $-f : [a, b] \to \mathbb{R}$, som er kontinuerlig ved Setning 9.5.1. Ved det vi nettopp har vist, finnes et maksimumspunkt for denne funksjonen, det vil si et punkt d slik at $-f(x) \leq -f(d)$ for alle $x \in [a, b]$. Men da er $f(x) \geq f(d)$ for alle $x \in [a, b]$, og d er det ønskede minimumspunktet for f.

Dermed har vi vist at f både har et maksimumspunkt og et minimumspunkt, og beviset er fullført. □

Vi avslutter med en teoretisk, meget interessant konsekvens av *ekstremalverdisetningen* og *skjæringssetningen*:

Følgesetning 9.7.10

Verdimengden $V(f)$ til en kontinuerlig funksjon $f : [a, b] \to \mathbb{R}$ er et lukket, endelig intervall; nærmere bestemt

$$V(f) = [f(d), f(c)]$$

der c og d er som i *ekstremalverdisetningen* (9.7.2). (Dette inkluderer muligheten $f(c) = f(d)$, der $V(f) = \{f(c)\}$.)

BEVIS

Ved *ekstremalverdisetningen* (9.7.2) er $V(f) \subset [f(d), f(c)]$. Ved *skjæringssetningen* (9.6) må f anta alle verdier mellom $f(d)$ og $f(c)$, slik at $V(f) = [f(d), f(c)]$. □

9.8. Fikspunktiterasjon*

Å løse ligninger er en viktig del av skolematematikken. Dette skyldes den enkle grunnen at mange problemer i et vidt spekter av fagfelt kan oversettes til en ligning, og ved å løse ligningen kan man løse problemet.

På skolen har vi lært å løse lineære og kvadratiske ligninger, samt noen trigonometriske ligninger og ligninger som involverer logaritmer og eksponensialfunksjoner. Å løse disse ligningene går som regel ut på å bruke en bestemt prosedyre, og de fleste av oss har vært vant til å kunne løse omtrent alle ligninger vi fikk presentert på skolen. Dette har kanskje gitt oss inntrykket av at det alltid finnes metoder for å finne løsninger av ligninger.

Sannheten er imidlertid at de aller fleste ligningene *ikke* kan løses ved hjelp av "tradisjonelle" metoder. Polynomielle ligninger, det vil si ligninger på formen $P(x) = 0$ der P er et (komplekst eller reellt) polynom, er et godt eksempel på dette. Vi kjenner løsningsformlene for slike ligninger av grad 1 og 2, der vi med graden på ligningen mener graden til polynomet P. Som nevnt i innledningen til Kapittel 4, finnes det løsningsformler for slike ligninger av grad 3 og 4, hvor vi med *løsningsformel* mener et uttrykk som involverer koeffisientene i ligningen og endelig mange av de aritmetiske operasjonene addisjon, subtraksjon, multiplikasjon, divisjon, samt røtter. Men *Abel–Ruffini teoremet*, som vi også nevnte i innledningen til Kapittel 4, slår fast at slike formler ikke finnes for et generelt polynom av grad > 4. Det finnes riktignok polynomielle ligninger av vilkårlig høy grad som er løsbare ved rotutdragning, som for eksempel ligninger av typen

$$z^n - 1 = 0,$$

jf. Setning 8.3.1. Men disse er alle av en *spesiell* type, det vil si at koeffisientene må oppfylle visse betingelser.

Når så de aller fleste ligningene ikke kan løses ved hjelp av enkle løsningsformler, hvordan kan vi da gå frem for å finne løsninger? Den generelle idéen er å gi opp å finne eksakte løsninger og istedenfor finne tilnærmede løsninger med den graden av nøyaktighet vi måtte ønske. Det gjør vi ved å

- finne en følge $\{x_n\}$ som konvergerer mot løsningen r;
- finne en øvre skranke for *feilen i tilnærmingen* $|x_n - r|$ uttrykt ved n.

Vi skal se på en slik prosedyre i denne seksjonen.

Iterasjoner. Uttrykket *iterasjon*[4] kan brukes generelt for alle gjentagende prosesser. I matematikk refererer uttrykket vanligvis til en algoritme av følgende type, hvor X er en mengde:

$$x_0 \in X \text{ gitt}, \quad x_{n+1} = f(x_n) \text{ for alle } n \in \mathbb{N}, \tag{242}$$

hvor $f : X \to X$ er en funksjon. Punktet x_0 kalles gjerne *startpunktet for iterasjonen*, og funksjonen f gjør oss altså i stand til å generere en følge $\{x_n\}$ i mengden X, som gjerne kalles *iterasjonsfølgen generert av f og x_0*. Hvis vi bruker notasjonen f^n for sammensetningen av funksjonen f med seg selv n ganger, det vil si

$$f^n : X \to X, \quad f^n = \underbrace{f \circ \cdots \circ f}_{n \text{ ganger}},$$

[4] *Iteration* på engelsk.

har vi altså

$$\begin{aligned} x_1 &= f(x_0), \\ x_2 &= f(x_1) = f \circ f(x_1) = f^2(x_0), \\ x_3 &= f(x_2) = f \circ f \circ f(x_1) = f^3(x_0), \\ \cdots & \cdots \\ x_n &= f^n(x_0), \end{aligned}$$

slik at iterasjonen også kan skrives som

$$x_0 \in X \text{ gitt}, \quad x_n = f^n(x_0) \text{ for alle } n \in \mathbb{Z}^+. \tag{243}$$

Eksempel 9.8.1. Herons metode fra Eksempel 6.3.6 (og Oppgave 7.39) for å finne en tilnærmingsverdi for $\sqrt{K}$ er basert på følgen $\{x_n\}$ i $\mathbb{R}$ gitt ved

$$x_0 = K \text{ og } x_{n+1} = \frac{1}{2}\left(x_n + \frac{K}{x_n}\right),$$

hvor vi har skalert indeksene i (104). Denne prosedyren er en iterasjon med startpunkt $x_0 = K$ og funksjonen $f : \mathbb{R} \to \mathbb{R}$ definert ved

$$f(x) = \frac{1}{2}\left(x + \frac{K}{x}\right).$$

Vi har vist i Oppgave 7.39 at $\{x_n\}$ konvergerer mot $\sqrt{K}$. Det er interessant å merke seg at

$$f(\sqrt{K}) = \frac{1}{2}\left(\sqrt{K} + \frac{K}{\sqrt{K}}\right) = \frac{1}{2}\left(2\sqrt{K}\right) = \sqrt{K}.$$

Et slikt punkt, som "blir bevart" av funksjonen, kalles et *fikspunkt* til f, se Definisjon 9.8.3 som vi skal komme til om et øyeblikk.

Eksempel 9.8.2. La $c \in \mathbb{C}$ og betrakt iterasjonen $f : \mathbb{C} \to \mathbb{C}$ gitt ved

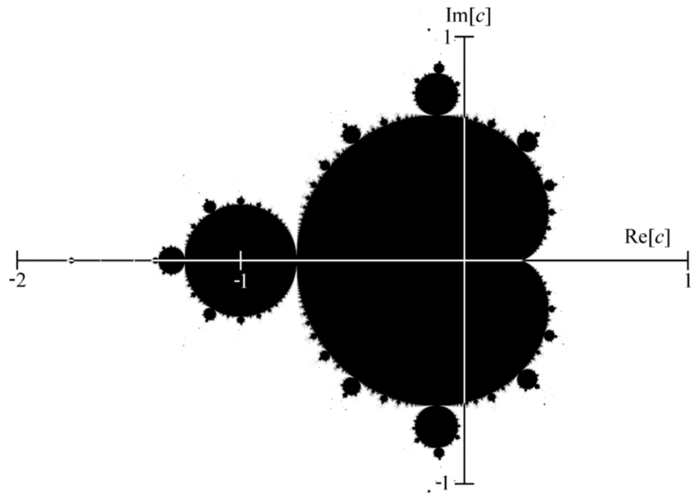
$$x_0 = c \text{ og } x_{n+1} = x_n^2 + c.$$

Mengden

$$\mathbb{M} = \{c \in \mathbb{C} \mid \{x_n\} \text{ er begrenset}\} \subset \mathbb{C}$$

kalles *Mandelbrot-mengden*, som vi allerede nevnte i §1.4. Lar vi for eksempel $c = 1$, er $\{x_n\} = \{0, 1, 2, 5, 26, \ldots\}$ og er ubegrenset, slik at $1 \notin \mathbb{M}$. For $c = -1$ er $\{x_n\} = \{0, -1, 0, -1, 0, -1, \ldots\}$, slik at $-1 \in \mathbb{M}$.

Tegner vi mengden $\mathbb{M}$ i svart i det komplekse planet, ser den slik ut:



MANDELBROT-MENGDEN I DET KOMPLEKSE PLANET (I SVART) (KILDE: WIKIPEDIA, PUBLIC DOMAIN)

Hvis man zoomer inn langs randen av Mandelbrot-mengden, vil man finne små kopier av mengden selv, i det uendelige. Dette er et eksempel på det man i matematikk kaller *fraktaler*.

Studiet av Mandelbrot-mengden startet med et arbeid av den franske matematikeren Adrien Douady (1935–2006) og den amerikanske matematikeren John H. Hubbard (1945–) fra 1985, som ga mengden navnet til ære for den polskfødte fransk-amerikanske matematikeren Benoit Mandelbrot (1924–2010), grunnet hans mange viktige arbeider innenfor fraktalgeometri.

Mandelbrot-mengden har blitt kjent utenfor matematiske kretser ikke fordi den er spesielt nyttig, men fordi den ser vakker ut og kan beskrives relativt enkelt.

Fikspunkter.

Definisjon 9.8.3: Fikspunkt

La X være en mengde og $f : X \to X$ en funksjon. Et *fikspunkt*[a] for f er et punkt $r \in X$ som oppfyller $f(r) = r$.

[a] *Fixed point* på engelsk.

La nå $X \subset \mathbb{C}$. (Dette inkluderer selvsagt spesialtilfellet $X \subset \mathbb{R}$.) Neste resultat viser at det vi observerte i Eksempel 9.8.1 er ingen tilfeldighet: dersom $f : X \to X$ er *kontinuerlig* og den genererte følgen $\{x_n\}$ konvergerer, da vil den nødvendigvis konvergere mot et *fikspunkt* til f.

Setning 9.8.4: Konvergens mot fikspunkt

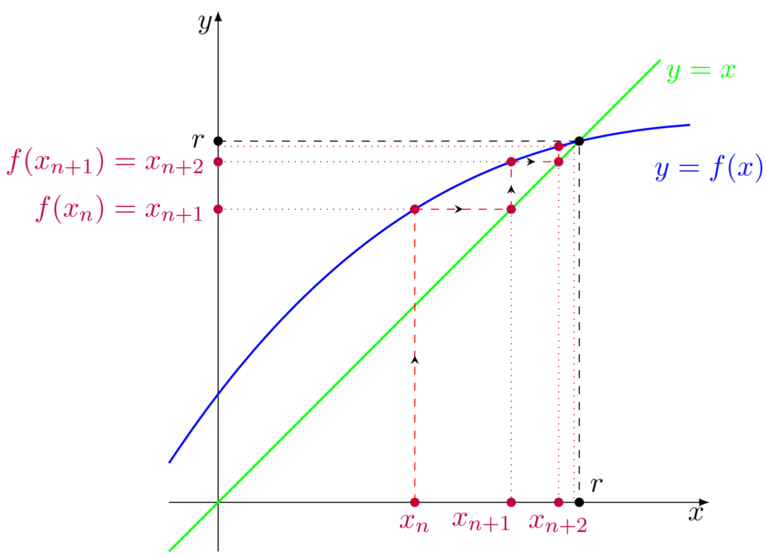
La $f : X \to X$ med $X \subset \mathbb{C}$ være en kontinuerlig funksjon og $x_0 \in X$. Dersom følgen $\{x_n\}$ definert ved (242) konvergerer mot $r \in X$, da er r et fikspunkt til f.

Bevis
Vi har $\lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = r$. Ved å ta grenser på begge sider av likheten $x_{n+1} = f(x_n)$, får vi

$$r = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} f(x_n) \overset{(*)}{=} f(\lim_{n\to\infty} x_n) = f(r),$$

hvor likheten i $(*)$ skyldes at f er kontinuerlig, ved følge-kriteriet (Setning 9.4.1 og Merknad 9.4.2). Dette viser at r er et fikspunkt til f. $\square$

Denne setningen er årsaken til at iterasjonen definert ved (242) også kalles en *fikspunktiterasjon*. Når $X \subset \mathbb{R}$ kan vi visualisere iterasjonen grafisk: Et fikspunkt til f er en løsning på ligningen $f(x) = x$, som er x-verdien til skjæringspunktet mellom kurvene $y = f(x)$ (grafen til f) og $y = x$ i planet. Figuren nedenfor viser grafisk lokaliseringen til punktene x_n og x_{n+1} i iterasjonen gitt ved (242):



Graf til f i blått, $y = x$ i grønt, og fikspunktet er r

Merknad 9.8.5. Det er helt essensielt at funksjonen f i setningen er kontinuerlig: den diskontinuerlige funksjonen $f : \mathbb{C} \to \mathbb{C}$ gitt ved

$$f(x) = \begin{cases} \frac{x}{2}, & x \neq 0 \\ 1, & x = 0 \end{cases}$$

genererer følgen $\{\frac{x_0}{2^n}\}$ hvis $x_0 \neq 0$ og $\{\frac{1}{2^{n-1}}\}$ hvis $x_0 = 0$, som uansett konvergerer mot 0. Men 0 er ikke et fikspunkt til f.

Har vi en ligning (i $\mathbb{R}$ eller $\mathbb{C}$) på formen $f(x) = x$, er altså fikspunktiterasjon en lovende prosedyre for å finne en tilnærmingsverdi for løsningen, fordi *om* den genererte følgen konvergerer, da vil den konvergere mot en løsning. Spørsmålet er da om følgen konvergerer i det hele tatt. La oss se på et enkelt eksempel:

EKSEMPEL 9.8.6. Betrakt ligningen

$$x^2 - 2x - 3 = 0,$$

hvor vi lett kan regne ut løsningene $x = -1$ og $x = 3$. La oss se hvordan fikspunktiterasjon fungerer for å finne den positive løsningen/roten $r = 3$. Vi må da omskrive ligningen slik at den får formen $x = f(x)$. Dette kan gjøres på mange måter; la oss betrakte to av dem:

(I) $x = \sqrt{2x+3}$ (som fungerer for den positive løsningen); her er $f : [0, \infty) \to [0, \infty)$ gitt ved $f(x) = \sqrt{2x+3}$.

(II) $x = \frac{x^2-3}{2}$; her er $f : \mathbb{R} \to \mathbb{R}$ gitt ved $f(x) = \frac{x^2-3}{2}$.

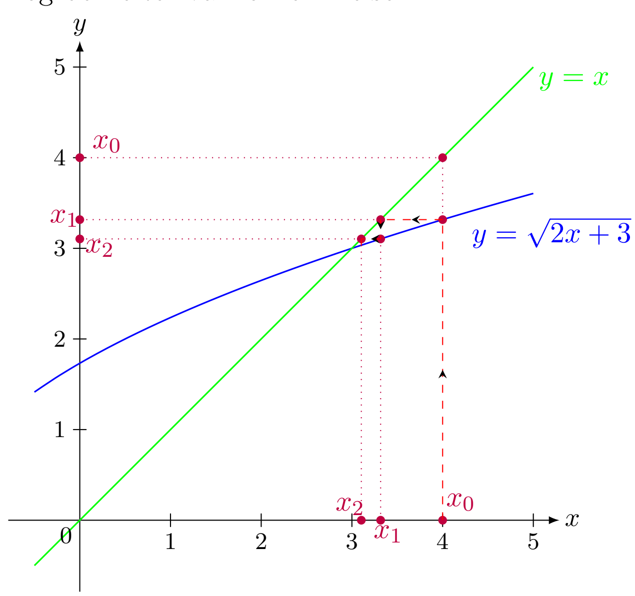
La oss studere tilfelle (I) først. Iterasjonsformelen blir

$$x_{n+1} = f(x_n) = \sqrt{2x_n + 3}.$$

Med startpunkt $x_0 = 4$, blir de første leddene i følgen:

$$\begin{aligned} x_1 &= \sqrt{11} = 3.3166 \\ x_2 &= 3.1037 \\ x_3 &= 3.0344 \\ x_4 &= 3.0114 \\ x_5 &= 3.0038 \end{aligned}$$

Denne følgen virker å konvergere mot løsningen $x = 3$. Følgende figur viser hva som skjer og bekrefter vår fornemmelse:



FIKSPUNKTITERASJON PÅ $f(x) = \sqrt{2x+3}$ MED $x_0 = 4$. GRAF TIL f I BLÅTT, LINJEN $y = x$ I GRØNT

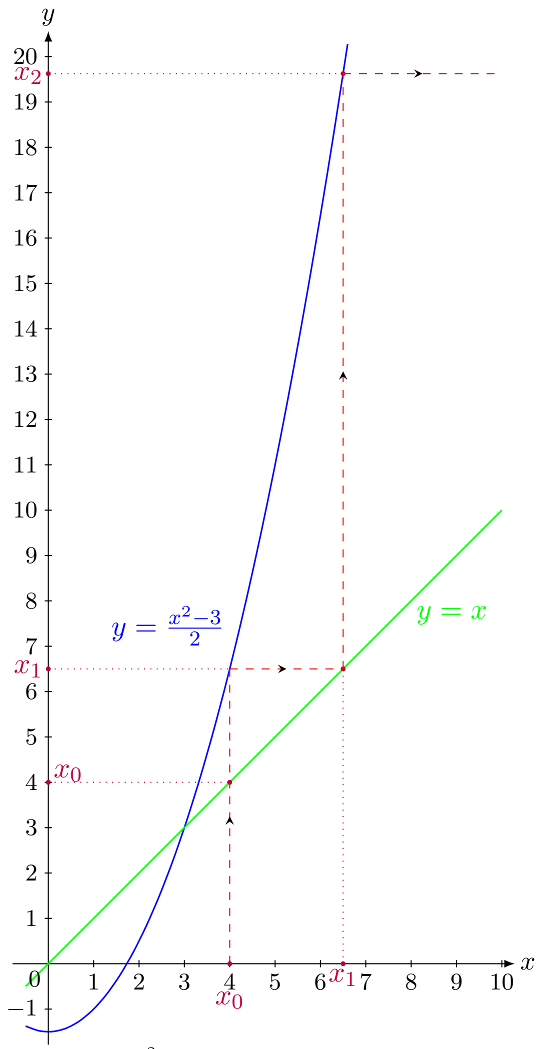
La oss dernest studere tilfelle (II). Iterasjonsformelen blir

$$x_{n+1} = f(x_n) = \frac{x_n^2 - 3}{2}.$$

Med startpunkt $x_0 = 4$ som i forrige tilfelle, blir de første leddene i følgen:

$$\begin{aligned} x_1 &= \tfrac{4^2-3}{2} = 6.500 \\ x_2 &= 19.625 \\ x_3 &= 191.07 \end{aligned}$$

Denne følgen virker å divergere. Det kan vi også se i følgende figur:



Fikspunktiterasjon på $f(x) = \frac{x^2-3}{2}$ med $x_0 = 4$. Graf til f i blått, linjen $y = x$ i grønt.

Vi skal videre i denne seksjonen utvikle en del teori som gjør oss i stand til å forklare (og forutsi!) hvorfor vi har konvergens i tilfelle (I) og divergens i tilfelle (II). Så skal vi ta opp igjen dette eksemplet i Eksempel 9.8.15.

Kontraksjoner og fikspunktteoremet. Vi så i forrige eksempel at en følge generert av fikspunktiterasjon både kan konvergere og divergere. Det er derfor viktig å finne kriterier som garanterer konvergens. En annen sak er at en funksjon kan ha flere fikspunkter, på samme måte som en ligning kan ha flere løsninger, slik at det også er viktig å finne kriterier som garanterer entydighet av fikspunkter. Før vi gir og beviser et resultat som gir slike kriterier, trenger vi en definisjon:

Definisjon 9.8.7: Kontraksjon

En funksjon $f : X \to X$ med $X \subset \mathbb{C}$ kalles en *kontraksjon*[a] dersom det finnes en $K \in \mathbb{R}$ slik at $K < 1$ og

$$|f(a) - f(b)| \leq K|a - b| \text{ for alle } a, b \in X. \tag{244}$$

Konstanten K (som nødvendigvis er ≥ 0) kalles en *Lipschitz-konstant* for f (etter den tyske matematikeren Rudolf Lipschitz (1832–1903)) og betingelsen (244) kalles gjerne en *Lipschitz-betingelse* for f.

[a] *Contraction* på engelsk.

MERKNAD 9.8.8. Ulikheten i definisjonen er *ikke* ekvivalent med ulikheten $|f(a) - f(b)| < |a - b|$. Det finnes nemlig funksjoner som oppfyller denne siste ulikheten, men $|f(a) - f(b)|$ kommer vilkårlig nær $|a - b|$ for $a, b \in X$, slik at det ikke finnes noen konstant $K < 1$ slik at (244) er oppfylt. Et konkret eksempel på en slik funksjon gis i Merknad 9.8.12 nedenunder.

MERKNAD 9.8.9. Det er lett å se at en kontraksjon er kontinuerlig. Det følger av det mer generelle resultatet i Oppgave 9.59.

MERKNAD 9.8.10. Dersom $X \subset \mathbb{R}$ er et intervall, kan man bruke derivasjon til å avgjøre om $f : X \to X$ er en kontraksjon, jf. Oppgave 9.41.

Teorem 9.8.11: Kontraksjonsteorem/Banach-fikspunktteorem

La $f : X \to X$ med $X \subset \mathbb{C}$ være en kontraksjon. Da har f et entydig fikspunt $r \in X$. Dessuten vil følgen $\{x_n\}$ definert ved $x_{n+1} = f(x_n)$ konvergere mot r for et vilkårlig startpunkt $x_0 \in X$.

BEVIS
Per antagelse finnes et reellt tall $K < 1$ slik at $|f(a) - f(b)| \leq K|a-b|$ for alle $a, b \in X$. Entydigheten av fikspunktet er en umiddelbar konsekvens

av dette: hvis $f(r_1) = r_1$ og $f(r_2) = r_2$, vil

$$|r_1 - r_2| = |f(r_1) - f(r_2)| \leq K|r_1 - r_2|,$$

som kun er mulig dersom $|r_1 - r_2| = 0$, det vil si $r_1 = r_2$.

For å vise resten av setningen, viser vi at iterasjonsfølgen konvergerer, siden den da automatisk vil konvergere mot et fikspunkt ved Setning 9.8.4. For å vise at $\{x_n\}$ konvergerer, viser vi at den er Cauchy, som medfører konvergens ved Setning 8.5.12.

For $n \geq 1$ har vi

$$|x_{n+1} - x_n| = |f(x_n) - f(x_{n-1})| \leq K|x_n - x_{n-1}|.$$

Ved induksjon får vi dermed

$$|x_{n+1} - x_n| \leq K^n |x_1 - x_0|. \tag{245}$$

For $m > n$ har vi derfor

$$\begin{aligned}
|x_m - x_n| &= |x_m - x_{m-1} + x_{m-1} - \cdots - x_{n+1} + x_{n+1} - x_n| \\
&\leq |x_m - x_{m-1}| + |x_{m-1} - x_{m-2}| \cdots + |x_{n+2} - x_{n+1}| + |x_{n+1} - x_n| \\
&\leq \left(K^{m-1} + K^{m-2} + \cdots + K^{n+1} + K^n\right) |x_1 - x_0| \\
&= K^n \left(K^{m-n-1} + K^{m-n-2} + \cdots + K^2 + K^1 + 1\right) |x_1 - x_0| \\
&\overset{(*)}{=} K^n \cdot \frac{1 - K^{m-n}}{1 - K} |x_1 - x_0| \leq \frac{K^n}{1 - K} |x_1 - x_0|,
\end{aligned}$$

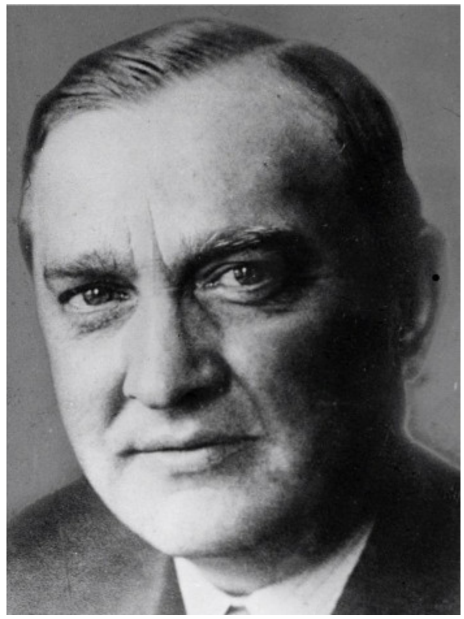
hvor vi har brukt resultatet fra Oppgave 5.10 i likheten $(*)$. Siden vi har at $0 \leq K < 1$, er $\lim_{n\to\infty} K^n = 0$ (ved Oppgave 7.21), og følgelig er også $\lim_{n\to\infty} \frac{K^n}{1-K}|x_1 - x_0| = 0$ (ved Setning 7.4.8(iii)). Gitt $\epsilon > 0$, finnes derfor en $N \in \mathbb{Z}^+$ slik at $\frac{K^n}{1-K}|x_1 - x_0| < \epsilon$ for alle $n \geq N$. Dermed er også $|x_m - x_n| < \epsilon$ for alle $m \geq n \geq N$, og vi har vist at $\{x_n\}$ er Cauchy. □

MERKNAD 9.8.12. I teoremet kan vi ikke erstatte betingelsen at f er en kontraksjon med den svakere betingelsen

$$|f(a) - f(b)| < |a - b| \text{ for alle } a, b \in X, \tag{246}$$

jf. Merknad 9.8.8. Betrakt for eksempel funksjonen $f : [1, \infty) \to [1, \infty)$ gitt ved $f(x) = 1 + \frac{1}{x}$. Denne oppfyller betingelsen (246), men har ingen fikspunkter. (Detaljene overlates til Oppgave 9.38.) Det følger spesielt at f ikke er en kontraksjon og at betingelsen (246) virkelig er en svakere betingelse enn betingelsen i Definisjon 9.8.7.

MERKNAD 9.8.13. Teoremet bærer sitt navn etter den polske matematikeren Stefan Banach (1892–1945), som beviste en mer generell versjon av teoremet i 1922. (For én mulig generalisering, se Oppgave 10.12.) Banach regnes som en av de mest innflytelsesrike matematikerne på 1900-tallet og som grunnleggeren av den moderne *funksjonalanalysen*.



Stefan Banach. (Kilde: Wikipedia, Creative Commons Attribution-Share Alike license)

Følgende viser at dersom en funksjon oppfyller den motsatte ulikheten av den i definisjonen for kontraksjon, da vil fikspunktiterasjonen alltid divergere, med mindre vi er så heldige at vi starter den i et fikspunkt.

Observasjon 9.8.14

La $X \subset \mathbb{C}$ og $f : X \to X$ være en funksjon. La $F \subset X$ være mengden av alle fikspunktene til f (som kan være tom).

Hvis $|f(a) - f(b)| \geq |a - b|$ for alle $a, b \in X$, da vil følgen generert av fikspunktiterasjon for et vilkårlig startpunkt $x_0 \in X \setminus F$ divergere.

Bevis

For $n \in \mathbb{Z}^+$ har vi

$$|x_{n+1} - x_n| = |f(x_n) - f(x_{n-1})| \geq |x_n - x_{n-1}|,$$

hvilket medfører induktivt at $|x_{n+1} - x_n| \geq |x_1 - x_0|$ for alle $n \in \mathbb{Z}^+$. En slik følge kan ikke konvergere med mindre $x_1 = x_0$. Siden $x_{n+1} = f(x_n)$, medfører dette ved induksjon at $x_n = x_0$ for alle $n \in \mathbb{Z}^+$, og følgen konvergerer mot x_0. $\square$

La oss ta en nøyere gjennomgang av Eksempel 9.8.6.

Eksempel 9.8.15. Vi betrakter igjen de to tilfellene (I) og (II) i Eksempel 9.8.6, hvor fikspunktet er 3.

(I) Her er $f : [0,\infty) \to [0,\infty)$, gitt ved $f(x) = \sqrt{2x+3}$. Siden

$$\begin{aligned}
|f(a) - f(b)| &= \left|\sqrt{2a+3} - \sqrt{2b+3}\right| \\
&= \left|\frac{(\sqrt{2a+3} - \sqrt{2b+3})(\sqrt{2a+3} + \sqrt{2b+3})}{\sqrt{2a+3} + \sqrt{2b+3}}\right| \\
&= \left|\frac{|2a+3| - |2b+3|}{\sqrt{2a+3} + \sqrt{2b+3}}\right| \leq \frac{2|a-b|}{2\sqrt{3}} \\
&= \frac{|a-b|}{\sqrt{3}},
\end{aligned}$$

for alle $a, b \geq 0$, er f en kontraksjon med Lipschitz-konstant $K = \frac{1}{\sqrt{3}}$. Følgen $\{x_n\}$ generert av fikspunktiterasjon vil ved Fikspunkteoremet 9.8.11 konvergere mot fikspunktet 3, uansett starverdi $x_0 \in [0,\infty)$.

(II) Her er $f : \mathbb{R} \to \mathbb{R}$ gitt ved $f(x) = \frac{x^2-3}{2}$. Vi merker at

$$f(x) \geq x \iff \frac{x^2-3}{2} \geq x \iff x^2 - 2x - 3 \geq 0 \iff (x-3)(x+1) \geq 0,$$

slik at $f(x) \geq x$ for alle $x \geq 3$. Det betyr at vi har en funksjon $f|_{[3,\infty)} : [3,\infty) \to [3,\infty)$. For alle $a, b \in [3,\infty)$ har vi

$$\begin{aligned}
|f(a) - f(b)| &= \left|\frac{a^2-3}{2} - \frac{b^2-3}{2}\right| = \frac{1}{2}|a^2 - b^2| \\
&= \frac{1}{2}|a-b||a+b| \geq \frac{1}{2}|a-b| \cdot (3+3) = 3|a-b|.
\end{aligned}$$

Derfor vil følgen $\{x_n\}$ generert av fikspunktiterasjon divergere ved Observasjon 9.8.14 for ethvert startpunkt $x_0 \in (3,\infty)$, spesielt for $x_0 = 4$.

I de situasjonene hvor fikspunktiterasjon gir oss en følge som konvergerer mot fikspunktet r er vi også interessert i å begrense *feilen i approksimasjonen*

$$r \approx x_n$$

som defineres som absoluttverdien til differansen mellom den virkelige verdien r og tilnærmingsverdien x_n, altså $|x_n - r|$. Følgende begrensninger er en konsekvens av beviset for fikspunktteoremet:

Setning 9.8.16

Under de samme betingelsene og med samme notasjon som i Teorem 9.8.11 gjelder, for alle $n \in \mathbb{Z}^+$:

(i) $|x_n - r| \leq \frac{K^n}{1-K}|x_1 - x_0|$;

(ii) $|x_n - r| \leq \frac{K}{1-K}|x_n - x_{n-1}|$;

(iii) $|x_n - r| \leq K|x_{n-1} - r|$,

hvor K er en Lipschitz-konstant for f.

Bevis
Siden f er en kontraksjon med Lipschitz-konstant K har vi

$$|x_n - r| = |f(x_{n-1}) - f(r)| \leq K|x_{n-1} - r|, \tag{247}$$

som gir oss (iii). Videre har vi

$$|x_{n-1} - r| = |x_{n-1} - x_n + x_n - r| \leq |x_{n-1} - x_n| + |x_n - r|,$$

som sammen med (iii) gir oss

$$|x_n - r| \leq K|x_{n-1} - x_n| + K|x_n - r|.$$

Dette er ekvivalent med (ii).

Induksjon og (iii) gir oss

$$|x_n - r| \leq K^n|x_0 - r|. \tag{248}$$

Siden

$$|x_0 - r| = |x_0 - x_1 + x_1 - r| \leq |x_0 - x_1| + |x_1 - r| \overset{(iii)}{\leq} |x_1 - x_0| + K|x_0 - r|,$$

får vi $|x_0 - r| \leq \frac{1}{1-K}|x_1 - x_0|$, som innsatt i (248) gir oss (i). (Alternativt kan vi utlede (i) ved å la $m \to \infty$ i ulikheten $|x_m - x_n| \leq \frac{K^n}{1-K}|x_1 - x_0|$ fra beviset på Teorem 9.8.11.) $\square$

Begrensningene i (i)-(iii) i Setning 9.8.16 er alle nyttige på hver sin måte:

- (i) gjør oss i stand til å gi en begrensning på alle feilene $|x_n - r|$ uansett n etter kun å ha regnet ut x_0 og x_1. Denne begrensningen er derfor nyttig for å finne n slik at feilen i tilnærmingen $r \approx x_n$ er mindre enn den tålegrensen vi setter oss (jf. Oppgave 9.39).
- (ii) gjør oss i stand til å begrense feilen $|x_n - r|$ ved hjelp av de to siste ledden i følgen. Denne begrensningen er nyttig når vi merker at følgen $\{x_n\}$ har begynt å stabilisere seg i et visst antall desimaler og vi ønsker å stoppe prosessen.
- (iii) sier at feilen $|x_n - r|$ i hvert trinn blir mindre med faktoren K (siden $K < 1$ er dette en forbedring i hvert trinn av iterasjonen). Denne er kanskje ikke så nyttig i praksis, men gir oss et mål på hvor god (rask) prosedyren er. Vi skal se senere, i Definisjon 9.8.20, at (iii) sier at *konvergensen er lineær*.

I forbindelse med feil i approksimasjoner er følgende språkbruk vanlig:

Definisjon 9.8.17: Nøyaktighet

Vi sier at approksimasjonen $r \approx x$ *gjelder med* n *desimalers nøyaktighet* dersom

$$r \in [x - 5 \cdot 10^{-n-1}, x + 5 \cdot 10^{-n-1}).$$

(Litt mer intuitivt betyr dette at avrundingen skjer i n-te desimal etter komma.)

Vi merker at betingelsen i definisjonen er oppfylt hvis $|r - x| < 5 \cdot 10^{-n-1}$.

EKSEMPEL 9.8.18. At $r \approx 4.79856$ med 5 desimalers nøyaktighet betyr altså at $r \in [4.798555, 4.798564)$. Intervallet består nettopp av alle tall som avrundes til 4.79856 i henhold til de vanlige avrundingsreglene.

La oss se på et omfattende eksempel.

EKSEMPEL 9.8.19. La oss betrakte ligningen

$$x^3 - 7x + 2 = 0 \tag{249}$$

i $\mathbb{R}$. Algebraens fundamentalteorem 8.4.1 sier oss at ligningen kan ha høyst tre reelle løsninger. Funksjonen $x \mapsto x^3 - 7x + 2$ er kontinuerlig, og ved å sette inn noen funksjonsverdier, ser vi at funksjonen skifter fortegn i hvert av intervallene $(-3, -2)$, $(0, 1)$ og $(2, 3)$. Ved *skjæringssetningen* (i versjonen fra Følgesetning 9.6.6) har derfor ligningen (249) en løsning (og kun én løsning) i hvert av intervallene $(-3, -2)$, $(0, 1)$ og $(2, 3)$.

La oss omforme ligningen slik at løsningene blir fikspunkter til en funksjon. Det finnes flere muligheter, som for eksempel

$$x = \frac{x^3 + 2}{7}, \tag{250}$$

Det betyr at løsningene på (249) oppfyller $x = f(x)$ med $f : \mathbb{R} \to \mathbb{R}$ definert ved $f(x) = \frac{x^3+2}{7}$. For å sjekke om fikspunktiterasjon kan fungere for å finne tilnærmede verdier på løsningene undersøker vi $|f(a) - f(b)|$ på hvert av intervallene ovenfor. Vi har

$$\begin{aligned}
|f(a) - f(b)| &= \left| \frac{a^3 + 2}{7} - \frac{b^3 + 2}{7} \right| = \frac{1}{7} \left| a^3 - b^3 \right| \\
&\overset{(*)}{=} \frac{1}{7} \left| (a - b)^3 + 3a^2 b - 3ab^2 \right| = \frac{1}{7} \left| (a - b)^3 + 3ab(a - b) \right| \\
&= \frac{1}{7} |a - b| \Big((a - b)^2 + 3ab \Big) = \frac{1}{7} |a - b| \left| a^2 + b^2 + ab \right|,
\end{aligned}$$

hvor vi har brukt *binomialteoremet* 3.8.7 i overgangen $(*)$. Vi får dermed

$$|f(a) - f(b)| \leq \frac{3}{7} |a - b| \quad \text{når} \quad a, b \in [0, 1], \tag{251}$$

$$|f(a) - f(b)| \geq \frac{12}{7} |a - b| \quad \text{når} \quad a, b \in [-3, -2] \text{ eller } [2, 3]. \tag{252}$$

Dette indikerer, ved *fikspunktteoremet* 9.8.11 og Observasjon 9.8.14 at følgen generert av fikspunktiterasjonen vil konvergere mot løsningen i $(0, 1)$, mens følgen vil divergere rundt de to andre løsningene. La oss derfor først konsentrere oss om løsningen i $(0, 1)$.

Funksjonen f er (strengt) voksende. Minimal- og maksimalverdien til f på intervallet $[0, 1]$ er derfor $f(0) = \frac{2}{7}$ og $f(1) = \frac{3}{7}$ henholdsvis. Altså ligger alle funksjonsverdiene i intervallet $[\frac{2}{7}, \frac{3}{7}] \subset [0,1]$. Vi kan dermed betrakte restriksjonen til f på intervallet $[0,1]$ (jf. Definisjon 2.2.10):

$$f|_{[0,1]} : [0,1] \longrightarrow [0,1]$$

(gitt ved samme uttrykk). Ved (251) er $f|_{[0,1]}$ en kontraksjon med Lipschitz-konstant $K = \frac{3}{7}$. Da vet vi ved *fikspunkteoremet* 9.8.11 at følgen generert av fikspunktiterasjonen vil konvergere mot fikspunktet til $f|_{[0,1]}$ i $(0,1)$, det vil si løsningen på ligningen (249), uansett starverdi $x_0 \in [0,1]$.

Velger vi startverdi $x_0 = 0.5$ og lar en datamaskin utføre iterasjonen med en nøyaktighet på 10 desimaler, får vi:

$$\begin{aligned}
x_1 &= \frac{17}{56} = 0.3035714286 \\
x_2 &= 0.2897108301 \\
x_3 &= 0.2891880165 \\
x_4 &= 0.2891692442 \\
x_5 &= 0.2891685715 \\
x_6 &= 0.2891685473
\end{aligned}$$

Det er fristende å stoppe opp å si at vi tror $r \approx 0.289169$ med 6 desimalers nøyaktighet, eller til og med $r \approx 0.2891685$ med 7 desimalers nøyaktighet. Men kan vi være sikre på at dette er riktig?

Det finnes flere måter å finne ut av dette. Vi kan bruke begrensningene på feilen $|x_n - r|$ gitt i Setning 9.8.16(i)-(ii) med $K = \frac{3}{7}$, som lyder henholdsvis:

$$|x_n - r| \leq \left(\frac{3}{7}\right)^n \cdot \frac{11}{2^5} \tag{253}$$

$$|x_n - r| \leq \frac{3}{4}\,|x_n - x_{n-1}|\,. \tag{254}$$

(etter litt regning). Begrensningen (253) med $n = 6$ gir oss

$$|x_6 - r| \leq 0.0021300\ldots < 5 \cdot 10^{-3}$$

og ut fra dette kan vi dessverre kun konkludere med at $r \approx x_6$ med 2 desimalers nøyaktighet, det vil si $r \approx 0.29$ med 2 desimalers nøyaktighet. Dette er ikke så rart, siden (253) kun er et grovt estimat for feilen. Ønsker vi å kunne konkludere for eksempel 6 desimalers nøyaktighet med denne skranken, må vi sørge for at

$$\left(\frac{3}{7}\right)^n \cdot \frac{11}{2^5} < 5 \cdot 10^{-7},$$

og litt regning (jf. Oppgave 9.39) viser at dette er oppfylt for $n \geq 16$. Vi kan altså la datamaskinen holde på frem til $n = 16$ for å være sikre på å oppnå 6 desimalers nøyaktighet.

Når vi undersøker hva begrensningen (254) gir oss, må vi huske på at kalkulatoren har regnet ut x_5 og x_6 med 10 desimalers nøyaktighet, hvilket betyr at

$$\begin{aligned} x_5 &\in [0.28916857145, 0.28916857154)], \quad \text{og} \\ x_6 &\in [0.28916854725, 0.28916854734). \end{aligned}$$

Vi kan dermed kun si at

$$|x_6 - x_5| \leq 0.28916857154 - 0.28916854725 = 0.000000024$$

og dermed at

$$|x_6 - r| \leq \frac{3}{4} \cdot 0.000000024 = 0.000000018 = 1.8 \cdot 10^{-8} < 5 \cdot 10^{-8}$$

ved (254). Dette sier oss at $r \approx x_6$ med 7 desimalers nøyaktighet, det vil si

(255) $$r \approx 0.2891685 \text{ med } 7 \text{ desimalers nøyaktighet.}$$

Det finnes også en annen metode for å sjekke at siste approksimasjon gjelder med det antallet desimalers nøyaktighet vi hevder. At (255) gjelder er ved Definisjon 9.8.17 ekvivalent med at

(256) $$r \in [0.28916845, 0.28916854).$$

For å sjekke om dette er sant kan vi bruke *skjæringssetningen* (i versjonen fra Følgesetning 9.6.6) nok en gang: vi regner ut, ved hjelp av en kalkulator, at verdien til $x^3 - 7x + 2$ er positiv i 0.28916845 og negativ i 0.28916854, hvilket betyr at det finnes en løsning i intervallet gitt i (256), som vi hevdet.

Det finnes altså flere måter å gå frem på for å oppnå en ønsket grad av nøyaktighet i tilnærmingen.

La oss nå konsentrere oss om de to andre løsningene på ligningen (249), de som ligger i intervallene $(-3, -2)$ og $(2, 3)$. Ved (252) og Observasjon 9.8.14 vet vi at fikspunkiterasjon ikke vil fungere på f. Men vi må ikke fortvile. Siden f er strengt voksende på begge intervallene, er den injektiv og har følgelig en invers funksjon. Ligningen $x = f(x)$ er ekvivalent med $f^{-1}(x) = x$, der f^{-1} er den inverse funksjonen til f over hvert av intervallene $(-3, -2)$ og $(2, 3)$. Verdimengden til f på intervallene $[-3, -2]$ og $[2, 3]$ er henholdsvis $[f(-3), f(-2)] = [-\frac{25}{7}, -\frac{6}{7}]$ og $[f(2), f(3)] = [\frac{10}{7}, \frac{29}{7}]$ ved Følgesetning 9.7.10. Dette er da lik definisjonsmengden til den tilsvarende inverse funksjonen. Nå er imidlertid begge disse intervallene større enn intervallene $[-3, -2]$ og $[2, 3]$ hvor vi vet løsningen ligger, så vi kan like godt begrense oppmerksomheten til de sistnevnte intervallene. Vi er altså interessert i funksjonene

(257) $$f^{-1}|_{[-3,-2]} : [-3, -2] \to [-3, -2] \text{ og } f^{-1}|_{[2,3]} : [2, 3] \to [2, 3].$$

Lar vi $b' = f(b)$ og $a' = f(a)$, slik at $b = f^{-1}(b')$ og $a = f^{-1}(a')$, er ulikheten (252) ekvivalent med

$$|a' - b'| \geq \frac{12}{7} \left| f^{-1}(a') - f^{-1}(b') \right|,$$

det vil si

$$\left|f^{-1}(a') - f^{-1}(b')\right| \le \frac{7}{12}|a' - b'|, \tag{258}$$

for alle a', b' i $[-3, -2]$ eller $[2, 3]$. Dermed er begge disse funksjonene i (257) kontraksjoner med Lipschitz-konstant $\frac{7}{12}$. Vi vet dermed ved *fikspunktteoremet* 9.8.11 at følgen generert av fikspunktiterasjonen vil konvergere mot fikspunktet til f^{-1} i intervallene, det vil si løsningen på ligningen (249), uansett starverdi x_0 i intervallene.

I dette tilfellet er det svært enkelt å finne et eksplisitt uttrykk for f^{-1} ved å løse ligningen $y = f(x)$ med hensyn på x:

$$y = \frac{x^3 + 2}{7} \iff \sqrt[3]{7y - 2} = x.$$

Altså er $f^{-1}(y) = \sqrt[3]{7y - 2}$ et uttrykk for den inverse funksjonen, eller

$$f^{-1}(x) = \sqrt[3]{7x - 2}. \tag{259}$$

(For å dobbeltsjekke, kan vi regne ut $f(f^{-1}(x)) = x$ og $f^{-1}(f(x)) = x$, som viser at vi har funnet riktig uttrykk for f^{-1}.) Løsningene på den opprinnelige ligningen oppfyller altså

$$x = \sqrt[3]{7x - 2}. \tag{260}$$

Det er lett å se at også denne ligningen er ekvivalent med (249).

Vi overlater til Oppgave 9.40 å utføre iterasjonen og drøftingen.

Vi legger merke til at selv om vi ikke hadde resonnert med den inverse funksjonen til f, så ville vi lett ha klart å finne at (260) er ekvivalent med (249), og vi kunne ha resonnert derfra, uten å betrakte inverse funksjoner.

Konvergenshastighet. Vi avslutter seksjonen med noen bemerkninger om hvor fort en følge konvergerer. Vi trenger først en definisjon:

Definisjon 9.8.20: Konvergenshastighet

La $\{x_n\}$ være en følge av komplekse tall som konvergerer mot x.

Vi sier at $\{x_n\}$ *konvergerer lineært*[a], eller at *konvergensen er lineær*, dersom det finnes en $N \in \mathbb{Z}^+$ slik at $|x_n - x| \le K|x_{n-1} - x|$ for et reellt tall $K < 1$ for alle $n \ge N$.

Vi sier at $\{x_n\}$ *konvergerer kvadratisk*[b], eller at *konvergensen er kvadratisk*, dersom det finnes en $N \in \mathbb{Z}^+$ slik at $|x_n - x| \le M|x_{n-1} - x|^2$ for et reellt tall M for alle $n \ge N$.

[a] *Converges linearly* på engelsk.
[b] *Converges quadratically* på engelsk.

Det finnes mange andre former for konvergens, for eksempel *kubisk*.

Setning 9.8.16(iii) sier at konvergensen i fikspunktiterasjon er lineær. Nærmere bestemt vil feilen i approksimasjonen $r \approx x_n$ bli mindre med en faktor $0 < K < 1$ for hvert trinn. Dette er imidlertid ikke en spesielt rask form for konvergens. Kvadratisk konvergens er en mye raskere form for konvergens, og det finnes approksimasjonsmetoder som garanterer denne formen for konvergens, som for eksempel *Newton-Raphson metoden* (jf. [**AE**, §4.2] eller [**Li**, §7.3]) dere lærer i MAT111. Flere numeriske metoder lærer dere i kursene *MAT160–Regnealgoritmer I* og *MAT260–Regnealgoritmer II*.

9.9. Uniform kontinuitet*

La oss betrakte en kontinuerlig funksjon $f : D(f) \to \mathbb{R}$, med $D(f) \subset \mathbb{R}$, og gå tilbake til definisjonen av kontinuitet i et enkeltstående punkt $a \in D(f)$ (Definisjon 9.1.1). Tallet δ i definisjonen kan avhenge av ϵ, som bemerket i Merknad 9.2.2, men også av a. At δen vi finner når vi skal vise at en funksjon er kontinuerlig i et punkt a i definisjonsmengden ofte avhenger av a, ser vi i følgende eksempel:

EKSEMPEL 9.9.1. (i) Vis at funksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved $f(x) = x^2$ er kontinuerlig ved å finne, til enhver $\epsilon \in \mathbb{R}^+$ og enhver $a \in \mathbb{R}$, en $\delta \in \mathbb{R}^+$ som oppfyller betingelsen (208) i Definisjon 9.1.1.

(ii) Vis at det (for fiksert $\epsilon \in \mathbb{R}^+$) ikke finnes en δ som oppfyller (208) for alle $a \in \mathbb{R}$ samtidig.

LØSNING. (i) Vi må vise at f er kontinuerlig i et vilkårlig punkt $a \in \mathbb{R}$. Vi skal, til enhver $\epsilon > 0$, finne en $\delta > 0$ slik at

$$|f(x) - f(a)| = \left|x^2 - a^2\right| < \epsilon \text{ når } |x - a| < \delta.$$

Vi har

$$\left|x^2 - a^2\right| = |x - a||x + a|,$$

og vi vil begrense $|x + a|$. Hvis for eksempel $|x - a| < 1$, da er

$$|x + a| = |x - a + 2a| \leq |x - a| + 2|a| < 1 + 2|a|,$$

og dermed er

$$\left|x^2 - a^2\right| = |x - a||x + a| < |x - a| \left(1 + 2|a|\right),$$

og dette siste uttrykket er $< \epsilon$ hvis (og bare hvis) $|x - a| < \frac{\epsilon}{1+2|a|}$. Vi ser dermed at $\delta = \min\left\{1, \frac{\epsilon}{1+2|a|}\right\}$ oppfyller betingelsen (208) i Definisjon 9.1.1.

(ii) Fiksér $\epsilon \in \mathbb{R}^+$ og anta at det finnes en $\delta \in \mathbb{R}^+$ som oppfyller (208) for alle $a \in \mathbb{R}$ samtidig. Dette betyr at, *for alle* $x, a \in \mathbb{R}$, gjelder

$$|f(x) - f(a)| = \left|x^2 - a^2\right| = |x - a||x + a| < \epsilon \text{ når } |x - a| < \delta. \tag{261}$$

Vi ser hva problemet er: når a og x vokser seg store og $|x - a|$ holdes fast, vil $|x - a||x + a|$ bli så stor vi bare vil. Hvis for eksempel $x = a + \frac{\delta}{2}$, slik at $|x - a| = \frac{\delta}{2} < \delta$, og $|x + a| \geq \frac{2\epsilon}{\delta}$, vil

$$|x - a||x + a| \geq \frac{\delta}{2} \cdot \frac{2\epsilon}{\delta} = \epsilon,$$

som motsier (261). □

Noen ganger kan vi imidlertid til enhver $\epsilon \in \mathbb{R}^+$ finne en $\delta \in \mathbb{R}^+$ som gjelder for alle a i definisjonsmengden samtidig. Vi har allerede sett noen eksempler på dette:

- Lineære funksjoner $f : \mathbb{R} \to \mathbb{R}$ gitt ved $f(x) = Ax+B$, for $A, B \in \mathbb{R}$: I Eksempel 9.2.1 så vi at $\delta = \frac{\epsilon}{|A|}$, hvis $A \neq 0$, fungerer for alle punkter i definisjonsmengden, mens en vilkårlig δ fungerer hvis $A = 0$.
- Absoluttverdifunksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved $f(x) = |x|$: I Eksempel 9.2.3 så vi at $\delta = \epsilon$ fungerer for alle punkter i definisjonsmengden.
- Kvadratrotfunksjonen $f : [0, \infty) \to \mathbb{R}$ gitt ved $f(x) = \sqrt{x}$: I Oppgave 9.2 skal du ha funnet ut at $\delta = \epsilon^2$ fungerer for alle punkter i definisjonsmengden.
- Sinusfunksjonen $f : [0, \infty) \to \mathbb{R}$ gitt ved $f(x) = \sin x$: I Oppgave 9.3 skal du ha funnet ut at $\delta = \epsilon$ fungerer for alle punkter i definisjonsmengden.

Det samme gjelder funksjonen $x \mapsto x^2$ fra Eksempel 9.9.1 dersom vi restriserer definisjonsmengden:

Eksempel 9.9.2. Betrakt funksjonen $f : I \to \mathbb{R}$ gitt ved $f(x) = x^2$ på et vilkårlig begrenset intervall I. Vis at det til enhver $\epsilon \in \mathbb{R}^+$ finnes en $\delta \in \mathbb{R}^+$ som oppfyller betingelsen (208) i Definisjon 9.1.1 for alle $a \in I$ samtidig.

Løsning. At I er begrenset betyr at I er på formen (c, d), $[c, d)$, $(c, d]$ eller $[c, d]$. For alle $x, a \in I$ er $|x + a| \leq |x| + |a| \leq 2\max\{|c|, |d|\}$, og dermed er

$$|f(x) - f(a)| = \left|x^2 - a^2\right| = |x - a||x + a| \leq |x - a| \cdot 2\max\{|c|, |d|\},$$

og dette siste uttrykket er $< \epsilon$ hvis (og bare hvis) $|x - a| < \frac{\epsilon}{2\max\{|c|,|d|\}}$. Vi ser dermed at $\delta = \frac{\epsilon}{2\max\{|c|,|d|\}}$ oppfyller betingelsen (208) i Definisjon 9.1.1 for alle $a \in I$ samtidig. □

Eksemplene ovenfor illustrerer forskjellen mellom begrepene *kontinuerlig* og *uniformt kontinuerlig* på en mengde A: en funksjon er uniformt kontinuerlig på A dersom den er kontinuerlig i ethvert punkt i A slik at det til enhver $\epsilon \in \mathbb{R}^+$ finnes en δ *slik at betingelsen* (208) *i Definisjon 9.1.1 er oppfylt for alle $a \in A$ samtidig*, det vil si slik at

$$|f(x) - f(a)| < \epsilon \text{ når } |x - a| < \delta \text{ og } x, a \in A.$$

For å vise symmetrien mellom x og a bedre bytter vi dem gjerne ut med x og y (eller med x_1 og x_2), slik at den formelle definisjonen blir som følger:

Definisjon 9.9.3: Uniform kontinuitet

La $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$ og la $A \subset D(f)$. Funksjonen f er *uniformt kontinuerlig*[a] på A dersom det til ethvert reelt tall $\epsilon > 0$ finnes et reelt tall $\delta > 0$ slik at

$$|f(x) - f(y)| < \epsilon \text{ når } |x - y| < \delta \text{ og } x, y \in A.$$

Vi sier at f er *uniformt kontinuerlig* hvis f er uniformt kontinuerlig på hele $D(f)$.

[a] *Uniformly continuous* på engelsk.

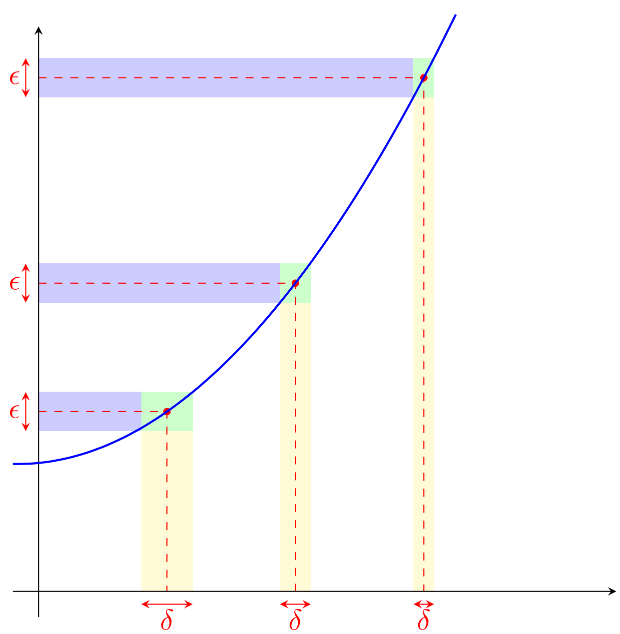
Definisjonen dukket først opp i et arbeid av den tyske matematikeren Eduard Heine (1821–1881) i 1870 og har senere spilt en sentral rolle i matematisk analyse. Ikke minst dukker begrepet opp i det viktige resultatet om at en kontinuerlig funksjon på et lukket, begrenset intervall er integrerbar, hvor *Heine-Cantor Teoremet* 9.9.13 nedenunder spiller en sentral rolle (som dere vil se i MAT112, jf. [**AE**, Teorem 5 i App. IV]). Begrepet var imidlertid kjent hos Bolzano, men ble publisert mye senere, i 1930.

Fra definisjonen merker vi spesielt at:

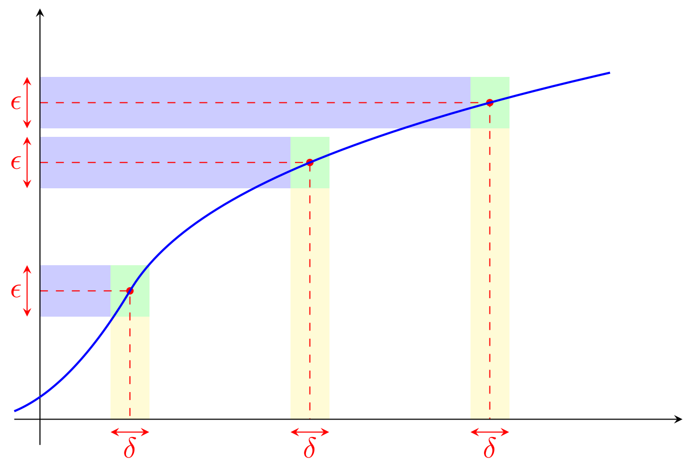
$$f \text{ uniformt kontinuerlig på } A \Longrightarrow f \text{ kontinuerlig på } A. \tag{262}$$

Den motsatte implikasjonen gjelder ikke: for eksempel er $x \mapsto x^2$ kontinuerlig på $\mathbb{R}$, men ikke uniformt kontinuerlig på $\mathbb{R}$ ved Eksempel 9.9.1(ii).

Følgende figurer illustrerer forskjellen på kontinuitet og uniform kontinuitet:



Ikke uniformt kontinuerlig: samme ϵ krever mindre og mindre δ når vi beveger oss mot høyre.



UNIFORMT KONTINUERLIG: SAMME δ FUNGERER OVERALT FOR FIKSERT ϵ.

MERKNAD 9.9.4. Begrepet *kontinuitet* kalles noen ganger for *punktvis kontinuitet*[5], siden det i motsetning til *uniform kontinuitet* er definert i enkeltstående punkter.

MERKNAD 9.9.5. Det er opplagt fra definisjonen at dersom f er uniformt kontinuerlig på A, så er f fremdeles uniformt kontinuerlig på enhver delmengde $B \subset A$.

Vi har allerede vist ovenfor at en del funksjoner er uniformt kontinuerlige. Vi oppsummerer disse eksemplene:

EKSEMPEL 9.9.6. Lineære funksjoner, absoluttverdifunksjonen, kvadratrotfunksjonen og sinusfunksjonen er uniformt kontinuerlige på sine definisjonsmengder (jf. teksten rett etter Eksempel 9.9.1).

Det samme er funksjonen $x \mapsto x^2$ på ethvert begrenset intervall (faktisk på enhver begrenset delmengde av $\mathbb{R}$) ved Eksempel 9.9.2.

Imidlertid er funksjonen $x \mapsto x^2$ ikke uniformt kontinuerlig på hele $\mathbb{R}$ ved Eksempel 9.9.1.

Før vi ser på et ytterligere eksempel, er det nyttig å se på hva det betyr at en funksjon *ikke* er uniformt kontinuerlig. Skriver vi ned negasjonen av betingelsen i Definisjon 9.9.3, så får vi ut et kriterium for at en funksjon ikke er uniformt kontinuerlig som er parallell til kriteriet for diskontinuitet fra Observasjon 9.2.16:

[5] *Pointwise continuity* på engelsk.

Observasjon 9.9.7

La $f : D(f) \to \mathbb{R}$ være en funksjon, med definisjonsmengde $D(f) \subset \mathbb{R}$. Funksjonen f er ikke uniformt kontinuerlig i $A \subset D(f)$ hvis og bare hvis det finnes et reelt tall $\epsilon > 0$ slik at for alle $\delta > 0$ finnes $x, y \in A$ som oppfyller

$$|x - y| < \delta \;\text{ og }\; |f(x) - f(y)| \geq \epsilon. \tag{263}$$

EKSEMPEL 9.9.8. Betrakt funksjonen $f : (0, \infty) \to \mathbb{R}$ gitt ved $f(x) = \frac{1}{x}$.

(i) Vis at f ikke er uniformt kontinuerlig på $(0, 1)$.

(ii) Vis at f er uniformt kontinuerlig på ethvert intervall $[a, \infty)$, for $a > 0$.

LØSNING. I begge tilfeller trenger vi å regne ut

$$|f(x) - f(y)| = \left|\frac{1}{x} - \frac{1}{y}\right| = \frac{|x - y|}{xy}. \tag{264}$$

Dette uttrykket vokser seg større og større når x og y går mot null, som er kjerneobservasjonen i dette eksemplet.

(i) Vi bruker Observasjon 9.9.7. Vi viser faktisk at kriteriet i observasjonen holder for alle ϵ. Gitt $\epsilon \in \mathbb{R}^+$ og $\delta \in \mathbb{R}^+$. For $y = \frac{x}{2}$ er $|x - y| = \frac{x}{2}$, som er $< \delta$ hvis og bare hvis $x < 2\delta$. Dessuten er

$$|f(x) - f(y)| = \frac{|x - y|}{xy} = \frac{1}{2y} = \frac{1}{x},$$

og dette er $\geq \epsilon$ hvis og bare hvis $x \leq \frac{1}{\epsilon}$. For å oppsummere: for $x < \min\left\{2\delta, \frac{1}{\epsilon}\right\}$ og $y = \frac{x}{2}$, er (263) oppfylt, som viser at f ikke er uniformt kontinuerlig på $(0, 1)$.

(ii) For $x, y \in [a, \infty)$ viser (264) at

$$|f(x) - f(y)| = \frac{|x - y|}{xy} \leq \frac{|x - y|}{a^2},$$

og dermed er $|f(x) - f(y)| < \epsilon$ såfremt $|x - y| < \delta = a^2\epsilon$. Dette viser at f er uniformt kontinuerlig på $[a, \infty)$. $\square$

MERKNAD 9.9.9. Definisjon 9.9.3 har sin opplagte generalisering til funksjoner med definisjonsmengde og verdimengde i $\mathbb{C}$, og mer generelt til metriske rom (Definisjon 7.2.4). Det samme gjelder de fleste resultatene i denne seksjonen. Vi overlater detaljene til oppgaver, jf. Oppgavene 10.39, 10.40 og 10.56.

Resultater og eksempler. Vi vil i denne delseksjonen samle en del resultater om uniform kontinuitet. Vi starter med å bemerke at man ut i fra definisjonen kan utlede resultater som er (nesten helt) parallelle til Setningene 9.5.1 og 9.5.2 om kontinuerlige funksjoner:

Setning 9.9.10: Kombinasjoner av unif. kont. funksjoner

Anta at $f : X \to \mathbb{R}$ og $g : X \to \mathbb{R}$, med $X \subset \mathbb{R}$, er uniformt kontinuerlige og la $c \in \mathbb{R}$. Da er også funksjonene $f + g$, $f - g$, cf uniformt kontinuerlige på X.

Hvis f og g i tillegg er begrenset, er også funksjonen fg uniformt kontinuerlig på X.

Bevis
Overlates til Oppgave 9.46. □

Merknad 9.9.11. Betingelsen om at f og g er begrenset er nødvendig for å konkludere at fg er uniformt kontinuerlig: vi har sett at funksjonen $x \mapsto x$ er uniformt kontinuerlig på $\mathbb{R}$, mens $x \mapsto x^2$ er det ikke (Eksempel 9.9.6).

Setning 9.9.12: Sammensetninger av unif. kont. funksjoner

En sammensetning av uniformt kontinuerlige funksjoner (med definisjonsmengder og verdimengder i $\mathbb{R}$) er uniformt kontinuerlig.

Bevis
Overlates til Oppgave 9.47. □

Vi har merket at funksjoner som ikke er uniformt kontinuerlige, som for eksempel $x \mapsto x^2$ og $x \mapsto \frac{1}{x}$, kan være uniformt kontinuerlige på delmengder av definisjonsmengden. For eksempel er ved Eksemplene 9.9.6 og 9.9.8 (samt Merknad 9.9.5) begge de sistnevnte funksjonene uniformt kontinuerlige på lukkede, begrensede intervaller. Dette er et generelt fenomen, som følgende svært viktige resultat viser:

Teorem 9.9.13: Heine-Cantor (for funksjoner definert på $\mathbb{R}$)

La $f : [a, b] \to \mathbb{R}$ være en kontinuerlig funksjon. Da er f også uniformt kontinuerlig.

Merknad 9.9.14. Den generelle versjonen av teoremet gjelder på vilkårlige metriske rom, med den naturlige generaliseringen av begrepet “lukket, begrenset intervall” til “kompakt metrisk rom” (jf. Definisjon 10.6.3). Se Oppgave 10.56.

Dette er kanskje det aller mest kjente resultatet om uniform kontinuitet og spiller, som nevnt ovenfor, en sentral rolle i beviset for at en kontinuerlig funksjon på et lukket, begrenset intervall er integrerbar (jf. [**AE**, Teorem 5 i App. IV]). Heine publiserte et bevis for teoremet i 1872, men hevdet ingen originalitet. Beviset var nemlig likt et bevis presentert av Dirichlet under en forelesning om integrerbarhet i 1854. Teoremet hadde imidlertid allerede tidligere blitt formulert av Bolzano, sammen med en bevisidé.

Det finnes et relativt kort og elegant bevis for dette teoremet som bygger på stoffet om delfølger og Bolzano-Weiestrass Teoremet fra §7.7. Vi overlater dette til Oppgave 9.58. Vi gir et bevis som er uavhengig av dette stoffet, men som er lengre. Det bygger på følgende observasjon som er interessant i seg selv og som viser at problemet med å "finne δ, gitt ϵ" kan deles opp i å finne δ'er på mindre delmengder:

Hjelpesetning 9.9.15

La $\epsilon \in \mathbb{R}^+$ og $f : D(f) \to \mathbb{R}$ med $D(f) \subset \mathbb{R}$. Anta at det finnes intervaller $I_1, I_2 \subset D(f)$ og $\delta_1, \delta_2 \in \mathbb{R}^+$ slik at

$$I_1 \cap I_2 \neq \emptyset \text{ og } f \text{ er kontinuerlig i minst ett punkt i } I_1 \cap I_2 \tag{265}$$

og

$$|f(x) - f(y)| < \epsilon \text{ når } |x-y| < \delta_i \text{ og } x, y \in I_i, \; i = 1, 2. \tag{266}$$

Da finnes en $\delta \in \mathbb{R}^+$ slik at

$$|f(x) - f(y)| < \epsilon \text{ når } |x-y| < \delta \text{ og } x, y \in I_1 \cup I_2. \tag{267}$$

Før vi beviser hjelpesetningen, merker vi at vi får følgende interessante konsekvens:

Følgesetning 9.9.16

La $f : D(f) \to \mathbb{R}$ være en funksjon med $D(f) \subset \mathbb{R}$. Hvis f er uniformt kontinuerlig på intervallene I_1 og I_2 og $I_1 \cap I_2 \neq \emptyset$, da er f uniformt kontinuerlig på $I_1 \cup I_2$.

Bevis
Overlates til Oppgave 9.56. □

Bevis for Hjelpesetning 9.9.15
Per antagelse finnes en $c \in I_1 \cap I_2$ slik at f er kontinuerlig i c. Dermed finnes en $\delta_3 \in \mathbb{R}^+$ slik at

$$|f(x) - f(c)| < \frac{\epsilon}{2} \text{ når } |x-c| < \delta_3 \text{ og } x \in I_1 \cup I_2. \tag{268}$$

La nå $\delta = \min\{\delta_1, \delta_2, \delta_3\}$ og anta at $|x-y| < \delta$. Hvis $x, y \in I_1$ eller $x, y \in I_2$, er (267) oppfylt på grunn av (266), siden $\delta \leq \delta_i$, $i = 1, 2$. Det resterende tilfellet er $x \in I_1 \setminus I_2$ og $y \in I_2 \setminus I_1$ (eller motsatt). Da er

$x < c < y$ (eller motsatt), slik at $|x-c|, |y-c| < \delta \leq \delta_3$. Ved (268) er

$$\begin{aligned} |f(x)-f(y)| &= |f(x)-f(c)+f(c)-f(y)| \\ &\leq |f(x)-f(c)|+|f(c)-f(y)| \\ &< \frac{\epsilon}{2}+\frac{\epsilon}{2}=\epsilon, \end{aligned}$$

slik at (267) igjen er oppfylt. □

Bevis for Teorem 9.9.13
Gitt $\epsilon \in \mathbb{R}^+$. La oss kalle f for *ϵ-god på intervallet I* dersom det finnes en $\delta \in \mathbb{R}^+$ slik at

$$|f(x)-f(y)| < \epsilon \text{ når } |x-y| < \delta,\ x, y \in I.$$

Vår oppgave er altså å vise at f er ϵ-god på hele $[a, b]$.

Betrakt mengden

$$S(\epsilon) = \{x \in [a, b] \mid f \text{ er } \epsilon\text{-god på } [a, x]\}.$$

Siden f er kontinuerlig i a, er f selvsagt ϵ-god på intervallet $[a, a] = \{a\}$, slik at $a \in S(\epsilon)$, og dermed er $S(\epsilon) \neq 0$. Dessuten er $S(\epsilon)$ oppad begrenset av b. Dermed eksisterer $s = \sup S(\epsilon)$, ved kompletthetsegenskapen til reelle tall. Vi har selvsagt $s \leq b$.

Vi viser først at tilfellet $s < b$ leder til en motsigelse.

Anta at $s < b$. Siden f er kontinuerlig i s, finnes $\delta' \in \mathbb{R}^+$ slik at $|f(x)-f(s)| < \frac{\epsilon}{2}$ når $|x-s| < \delta'$ og $x \in [a, b]$. Velg $\delta'' < \delta'$ liten nok slik at $0 < \delta'' < b-s$. Hvis $x, y \in [\max\{a, s-\delta''\}, s+\delta'']$, vil derfor $|x-s|, |y-s| \leq \delta'' < \delta'$, slik at

$$\begin{aligned} |f(x)-f(y)| &= |f(x)-f(s)+f(s)-f(y)| \\ &\leq |f(x)-f(s)|+|f(s)-f(y)| \\ &< \frac{\epsilon}{2}+\frac{\epsilon}{2}=\epsilon. \end{aligned}$$

Dette viser at f er ϵ-god på intervallet $[\max\{a, s-\delta''\}, s+\delta'']$. Hvis $\max\{a, s-\delta''\} = a$, betyr dette at f er ϵ-god på $[a, s+\delta'']$, som betyr at $s+\delta'' \in S(\epsilon)$, hvilket motsier at s er en øvre skranke for $S(\epsilon)$. Hvis $\max\{a, s-\delta''\} = s-\delta'' > a$, betyr dette at f er ϵ-god på intervallet $[s-\delta'', s+\delta'']$. Siden $a < s-\delta'' < s$, og s er den minste øvre skranken til $S(\epsilon)$, er $s-\delta'' \in S(\epsilon)$, som betyr at f er ϵ-god på intervallet $[a, s-\delta'']$. Ved Hjelpesetning 9.9.15, er f ϵ-god på $[a, s-\delta''] \cup [s-\delta'', s+\delta''] = [a, s+\delta'']$. Dette betyr igjen at $s+\delta'' \in S(\epsilon)$, som motsier at s er en øvre skranke for $S(\epsilon)$.

Eneste mulighet er derfor at $s = b$. Siden f er kontinuerlig i b, finnes en $\delta_b > 0$ slik at $|f(x)-f(b)| < \epsilon$ når $x \in (\max\{a, b-\delta_b\}, b]$. Velg $\delta < \delta_b$ liten nok slik at $0 < \delta < b-a$. Da er f ϵ-god på $[b-\delta, b]$. Siden $b-\delta < b = \sup S(\epsilon)$, er f også ϵ-god på $[a, b-\delta]$. Ved Hjelpesetning 9.9.15, er f ϵ-god på $[a, b-\delta] \cup [b-\delta, b] = [a, b]$, som er det vi skulle bevise. □

MERKNAD 9.9.17. For et ytterligere bevis for Teorem 9.9.13, se [**AE**, Teorem 4 i App. IV] (som er pensum i MAT112).

Det neste resultatet gir en tilstrekkelig betingelse for at en funksjon er uniformt kontinuerlig, som er et hendig kriterium. Vi trenger først en definisjon

Definisjon 9.9.18: Lipschitz-funksjon

La $X \subset \mathbb{C}$. En funksjon $f : X \to \mathbb{C}$ kalles en *Lipschitz-funksjon* (etter den tyske matematikeren Rudolf Lipschitz (1832–1903)) dersom det finnes en $K \in \mathbb{R}$ slik at

$$|f(x) - f(y)| \leq K|x - y| \text{ for alle } x, y \in X. \tag{269}$$

for alle $x, y \in X$.

Betingelsen (269) kalles en *Lipschitz-betingelse* med *Lipschitz-konstant* K.

MERKNAD 9.9.19. Ved denne definisjonen kan vi altså si at en *kontraksjon* (Definisjon 9.8.7) er en Lipschitz-funksjon med Lipschitz-konstant $K < 1$.

Setning 9.9.20

En Lipschitz-funksjon $f : X \to \mathbb{R}$, hvor $X \subset \mathbb{R}$, er uniformt kontinuerlig.

BEVIS
Overlates til Oppgave 9.59. □

MERKNAD 9.9.21. Det motsatte av siste setning gjelder ikke: f kan være uniformt kontinuerlig uten å være en Lipschitz-funksjon. For eksempel er kvadratrotfunksjonen $x \mapsto \sqrt{x}$ uniformt kontinuerlig på $[0, \infty)$ ved Eksempel 9.9.6, men ikke Lipschitz, siden vi for $x < y$ har

$$|\sqrt{x} - \sqrt{y}| = \left| \frac{(\sqrt{x} - \sqrt{y})(\sqrt{x} + \sqrt{y})}{\sqrt{x} - \sqrt{y}} \right| = \frac{|x - y|}{\sqrt{x} + \sqrt{y}}.$$

Dersom dette er $\leq K|x - y|$, for en konstant K, for alle $x, y \in [0, \infty)$ med $x \leq y$, vil $\frac{1}{K} \leq \sqrt{x} + \sqrt{y} \leq 2\sqrt{y}$ for alle $y > 0$, en umulighet.

MERKNAD 9.9.22. Som konsekvens av siste setning kan man utlede det interessante resultatet at en funksjon $f : I \to \mathbb{R}$, hvor I er et intervall, med *begrenset derivert* er uniformt kontinuerlig (jf. Oppgave 9.60).

EKSEMPEL 9.9.23. Vis at $\ln : (0, \infty) \to \mathbb{R}$ er uniformt kontinuerlig på ethvert intervall $[a, \infty)$, for $a > 0$.

BEVIS. For $0 < a \le y \le x$ har vi

$$\frac{x}{y} = \frac{x-y}{y} + 1 \le \frac{x-y}{a} + 1. \tag{270}$$

Dette gir oss

$$\ln x - \ln y \overset{(*)}{=} \ln \frac{x}{y} \overset{(**)}{\le} \ln\left(\frac{x-y}{a} + 1\right) \overset{(***)}{\le} \frac{x-y}{a},$$

hvor vi har brukt identiteten fra Oppg.9.18(d) i overgangen (*), at ln er strengt voksende (Setning 9.5.14) og (270) i overgangen (**), og ulikheten i Oppgave 9.21(b) i overgangen (***). Dette viser at ln er en *Lipschitz-funksjon* med *Lipschitz-konstant* $\frac{1}{a}$ på $[a, \infty)$, og er derfor uniformt kontinuerlig der ved Setning 9.9.20. □

EKSEMPEL 9.9.24. Vis at $f : (0, \infty) \to \mathbb{R}$ gitt ved $f(x) = x \sin\left(\frac{1}{x}\right)$ er uniformt kontinuerlig på $(0, \infty)$.

LØSNING. Ved Følgesetning 9.9.16 er det nok å vise at f er uniformt kontinuerlig på intervallene $(0, 1]$ og $[1, \infty)$.

Ved Oppgave 9.1(e) er funksjonen $\overline{f} : \mathbb{R} \to \mathbb{R}$ definert ved

$$\overline{f}(x) = \begin{cases} x \sin\left(\frac{1}{x}\right), & x \neq 0 \\ 0 & x = 0, \end{cases}$$

kontinuerlig i 0 (beviset er svært likt beviset utført i Eksempel 9.5.19). Utenom 0 er funksjonen en kombinasjon av kontinuerlige funksjoner, og er derfor kontinuerlig. Spesielt er $\overline{f}$ kontinuerlig på $[0, 1]$ og dermed uniformt kontinuerlig der ved Teorem 9.9.13. Den er da følgelig også uniformt kontinuerlig på $(0, 1]$ ved Merknad 9.9.5. Siden $\overline{f}|_{(0,1]} = f$, er f uniformt kontinuerlig på $(0, 1]$.

For å vise at f er uniformt kontinuerlig på intervallet $[1, \infty)$ holder det ved Setning 9.9.20 å vise at f er en Lipschitz-funksjon på $[1, \infty)$. (For å gjøre dette kan man enkelt derivere funksjonen og vise at den deriverte er begrenset og konkludere ved hjelp av Oppgave 9.60, jf. Merknad 9.9.22.

Siden derivasjon ikke omtales i denne boken, vil vi bruke en alternativ, mer direkte, prosedyre.)

Kjerneobservasjonen er identiteten

$$\sin\alpha - \sin\beta = 2\cos\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2}.$$

Denne gir oss at

$$\begin{aligned}\left|\sin\left(\frac{1}{x}\right) - \sin\left(\frac{1}{y}\right)\right| &= \left|2\cos\frac{\frac{1}{x}+\frac{1}{y}}{2}\sin\frac{\frac{1}{x}-\frac{1}{y}}{2}\right| \leq 2\left|\sin\frac{\frac{1}{x}-\frac{1}{y}}{2}\right| \\ &= 2\left|\sin\frac{y-x}{2xy}\right| \overset{(*)}{\leq} 2\left|\frac{y-x}{2xy}\right| = \frac{|y-x|}{|x||y|},\end{aligned}$$

hvor vi har brukt at $\sin|\alpha| \leq |\alpha|$ (jf. Merknad 9.2.7) i overgangen $(*)$. Den oppnådde ulikheten kan vi nå bruke til å utlede:

$$\begin{aligned}&\left|x\sin\left(\frac{1}{x}\right) - y\sin\left(\frac{1}{y}\right)\right| \\ =\ &\left|x\sin\left(\frac{1}{x}\right) - y\sin\left(\frac{1}{x}\right) + y\sin\left(\frac{1}{x}\right) - y\sin\left(\frac{1}{y}\right)\right| \\ \leq\ &\left|x\sin\left(\frac{1}{x}\right) - y\sin\left(\frac{1}{x}\right)\right| + \left|y\sin\left(\frac{1}{x}\right) - y\sin\left(\frac{1}{y}\right)\right| \\ =\ &|x-y|\left|\sin\left(\frac{1}{x}\right)\right| + |y|\left|\sin\left(\frac{1}{x}\right) - \sin\left(\frac{1}{y}\right)\right| \\ \leq\ &|x-y|\cdot 1 + |y|\cdot\frac{|y-x|}{|x||y|} \\ =\ &|x-y| + \frac{|y-x|}{|x|} \\ \overset{x\geq 1}{\leq}\ &|x-y| + |y-x| = 2|x-y|.\end{aligned}$$

Dette viser at f er en Lipschitz-funksjon på $[1,\infty)$. ☐

Det neste resultat gir en *nødvendig betingelse* for at en funksjon skal være uniformt kontinuerlig på et intervall. Spesielt interessant er resultatet hvis vi minner oss om at en kontinuerlig funksjon på et *lukket, begrenset* intervall er begrenset der ved *Begrensningsteoremet* 9.7.9. Fjerner vi betingelsen om at intervallet skal være *lukket*, gjelder ikke *Begrensningsteoremet*: tenk for eksempel på den kontinuerlige funksjonen $x \mapsto \frac{1}{x}$ på det begrensede intervallet $(0,1]$. En *uniformt* kontinuerlig funksjon er imidlertid alltid begrenset på et begrenset intervall, selv om intervallet ikke skulle være lukket:

Setning 9.9.25

La $f : I \to \mathbb{R}$ være en uniformt kontinuerlig funksjon definert på et begrenset intervall $I \subset \mathbb{R}$. Da er f begrenset på I.

BEVIS. At I er begrenset, betyr at I er på formen $[a,b]$, $(a,b]$, $[a,b)$ eller (a,b), med $a,b \in \mathbb{R}$.

Velg en hvilken som helst $\epsilon > 0$, f.eks. $\epsilon = 1$. Siden f er uniformt kontinuerlig, finnes en $\delta > 0$ slik at

$$|f(x) - f(y)| < \epsilon = 1 \text{ når } x, y \in I \text{ og } |x - y| < \delta.$$

Del I opp i N like store delintervaller $I_1, \ldots, I_N$, der N er så stor at bredden på intervallene er

$$\frac{b-a}{N} < \delta.$$

(Det holder med N slik at $N > \frac{b-a}{\delta}$.) Velg en $z_i \in I_i$ for alle i (for eksempel kan vi la $z_i = a + (i - \frac{1}{2})\frac{b-a}{N}$, midtpunktet i I_i). For hver i og $x \in I_i$, er $|x - z_i| < \delta$, og vi har derfor

$$|f(x)| = |f(x) - f(z_i) + f(z_i)| \leq |f(x) - f(z_i)| + |f(z_i)| \leq 1 + |f(z_i)|.$$

Da må, for $x \in I$,

$$|f(x)| \leq 1 + \max_{1 \leq i \leq N}\{|f(z_i)|\}.$$

Sett $M := \max_{1 \leq i \leq N}\{|f(z_i)|\}$. Da er $|f(x)| \leq M + 1$. □

MERKNAD 9.9.26. Det motsatte av setningen holder ikke, selv for kontinuerlige funksjoner: f kan være kontinuerlig og begrenset på et begrenset intervall I uten å være uniformt kontinuerlig der. Et eksempel på dette er funksjonen $x \mapsto \sin\left(\frac{1}{x}\right)$ på $(0,1)$, se Eksempel 9.9.31 nedenunder.

EKSEMPEL 9.9.27. Vis at $\ln : (0,1) \to \mathbb{R}$ ikke er uniformt kontinuerlig.

LØSNING (ALTERNATIV I). Funksjonen er ubegrenset (nedad): gitt en $M \in \mathbb{R}$, velg $x < e^M$; da er $\ln x < \ln e^M = M$, siden ln er strengt voksende (Setning 9.5.14). Dermed er ln ikke uniformt kontinuerlig ved Setning 9.9.25. □

LØSNING (ALTERNATIV II). Vi bruker kriteriet i Observasjon 9.9.7 med $\epsilon = 1$: Gitt $\delta \in \mathbb{R}^+$. La $x = \delta$ og $y = \frac{\delta}{3}$. Da er $|x - y| = \frac{\delta}{3} < \delta$ og

$$|\ln x - \ln y| = \left|\ln \frac{x}{y}\right| = \ln 3 > \ln e = 1,$$

siden $e < 3$ og ln er strengt voksende (Setning 9.5.14). □

Vi tar også med følgende resultat som er parallelt til *følge-kriteriet for kontinuitet* (Setning 9.4.1). Dette kriteriet blir i en god del lærebøker brukt som *definisjon* på uniform kontinuitet, hvoretter kriteriet i Definisjon 9.9.3 blir utledet som en konsekvens.

Setning 9.9.28: Følge-kriteriet for uniform kontinuitet

La $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$ og la $a \in D(f)$. Funksjonen f er uniformt kontinuerlig på en delmengde $A \subset D(f)$ hvis og bare hvis $\lim_{n\to\infty} (f(x_n) - f(y_n)) = 0$ for alle følger $\{x_n\}, \{y_n\}$ i A slik at $\lim_{n\to\infty} (x_n - y_n) = 0$.

Bevis
Dette ligner på beviset for Setning 9.4.1 og overlates til Oppgave 9.61.
□

Merknad 9.9.29. Man kan bruke kriteriet fra Setning 9.9.28 på følgende måte for å vise at en funksjon f ikke er uniformt kontinuerlig på en mengde A. Start med en følge $\{\delta_n\}$ slik at $\delta_n \to 0$. Finn deretter en følge $\{x_n\}$ i A slik at $x_n + \delta_n \in A$ for alle $n \in \mathbb{Z}^+$ og

$$\lim_{n\to\infty} (f(x_n) - f(x_n + \delta_n)) \neq 0.$$

Ved å sette $y_n = x_n + \delta_n$ har vi funnet følger $\{x_n\}$ og $\{y_n\}$ i A som oppfyller

$$\lim_{n\to\infty} (x_n - y_n) = 0 \quad \text{og} \quad \lim_{n\to\infty} (f(x_n) - f(y_n)) \neq 0.$$

Setning 9.9.28 viser da at f *ikke* er uniformt kontinuerlig på A.

Eksempel 9.9.30. Vis at funksjonen $f : (0, \infty) \to \mathbb{R}$ gitt ved $f(x) = \frac{1}{x}$ ikke er uniformt kontinuerlig på $(0, 1)$.

Løsning. Vi argumenterer som foreslått i Merknad 9.9.29 og starter med en følge δ_n slik at $\delta_n > 0$ for alle $n \in \mathbb{Z}^+$ og $\delta_n \to 0$. For en vilkårlig følge $\{x_n\}$ med $x_n > 0$ for alle $n \in \mathbb{Z}^+$ har vi at

$$|f(x_n) - f(x_n + \delta_n)| = \left| \frac{1}{x_n} - \frac{1}{x_n + \delta_n} \right| = \frac{\delta_n}{x_n(x_n + \delta_n)}.$$

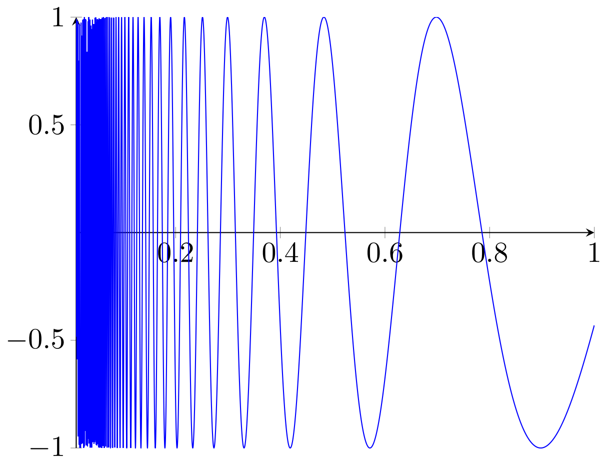
Nå ser vi at hvis vi velger $x_n = \delta_n$, da vil

$$|f(x_n) - f(x_n + \delta_n)| = \frac{1}{2x_n} \to \infty.$$

Vi må nå bare sørge for at $x_n, x_n + \delta_n \in (0, 1)$. Dette er for eksempel oppfylt for $x_n = \delta_n = \frac{1}{n+2}$. □

Eksempel 9.9.31. Avgjør om funksjonen $f : (0, \infty) \to \mathbb{R}$ gitt ved $f(x) = \sin\left(\frac{1}{x}\right)$ er uniformt kontinuerlig på $(0, 1)$.

Løsning. Det er instruktivt å ta en titt på grafen til f (som er "halvparten" av grafen vi har sett på side 341):



GRAFEN TIL f I EKSEMPEL 9.9.31.

Funksjoner som har grafer som "svinger mer og mer inn mot et punkt", som i dette eksempelet, er typisk *ikke* uniformt kontinuerlige, siden punktene som gir toppene og bunnene på grafen gir følger som nærmer seg hverandre mer og mer, men hvor de tilhørende funksjonsverdiene ikke nærmer seg hverandre. Ved å finne punktene x_n og y_n som gir toppene og bunnene på grafen vil vi finne følger $\{x_n\}$ og $\{y_n\}$ slik at $\lim_{n\to\infty}(x_n - y_n) = 0$, mens $\lim_{n\to\infty}(f(x_n) - f(y_n)) \neq 0$. Funksjonen vil da ikke være uniformt kontinuerlig ved Setning 9.9.28.

I dette konkrete tilfelle merker vi at

$$\sin\left(\frac{1}{x_n}\right) = 1 \iff \frac{1}{x_n} = \frac{\pi}{2} + 2\pi n \iff x_n = \frac{2}{\pi(4n+1)}$$

og

$$\sin\left(\frac{1}{y_n}\right) = -1 \iff \frac{1}{y_n} = \frac{3\pi}{2} + 2\pi n \iff y_n = \frac{2}{\pi(4n+3)}.$$

Vi har

$$\left|x_n - y_n\right| = \left|\frac{2}{\pi(4n+1)} - \frac{2}{\pi(4n+3)}\right| = \frac{4}{\pi(4n+1)(4n+3)} \to 0$$

og

$$\left|\sin\left(\frac{1}{x_n}\right) - \sin\left(\frac{1}{y_n}\right)\right| = |1-(-1)| = 2,$$

slik at $\lim_{n\to\infty}(f(x_n) - f(y_n)) \neq 0$. Dermed er f ikke uniformt kontinuerlig på $(0,1)$ ved Setning 9.9.28. □

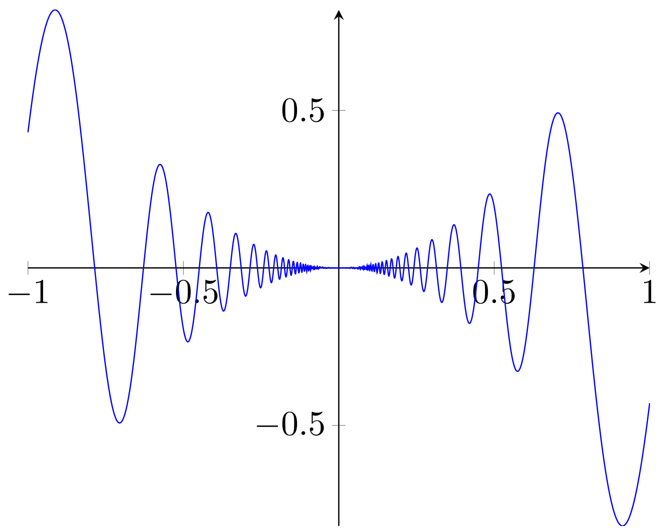
9.10. Grenseverdier*

Vi avslutter kapitlet med et begrep som har mye til felles med kontinuitet, nemlig grenser. Seksjonen inneholder ikke noe vesentlig nytt stoff i forhold til MAT111 (se [**AE**, §1.2, 1.3, 1.5] og [**Li**, §5.4]), bortsett fra definisjonen av *opphopningspunkt/grensepunkt* (Definisjon 9.10.1) og de forskjellige følgekriteriene for grenser (Setningene 9.10.9, 9.10.14, 9.10.19 og 9.10.36).

La $f : X \to \mathbb{R}$ være en funksjon, med definisjonsmengde $X \subset \mathbb{R}$. At f er kontinuerlig i $a \in X$ betyr at funksjonsverdiene $f(x)$ kommer så nær vi vil $f(a)$ så lenge x er nær nok a. Det er også naturlig å spørre seg om en svakere betingelse er oppfylt: uavhengig av om $x \in X$ eller ikke, *finnes* i det hele tatt et tall $L \in \mathbb{R}$, ikke nødvendigvis lik $f(a)$, som har egenskapen at $f(x)$ kommer så nær vi vil L så lenge x er nær nok a?

Et godt eksempel er funksjonen

$$\begin{array}{rcl} (-\infty, 0) \cup (0, \infty) & \longrightarrow & \mathbb{R} \\ x & \mapsto & x^2 \sin\left(\dfrac{1}{x}\right). \end{array}$$



GRAFEN TIL $x \mapsto x^2 \sin\left(\frac{1}{x}\right),\ x \neq 0$

Funksjonen er ikke definert i 0, men funksjonsverdiene kommer så nær vi vil 0 når x går mot null. Vi kan derfor utvide funksjonen til å være definert på hele $\mathbb{R}$ ved å definere den til å være 0 i 0, og få funksjonen

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & x \neq 0 \\ 0, & x = 0, \end{cases}$$

som er kontinuerlig i 0, som vi så i Eksempel 9.2.5. Vi kaller en slik utvidelse for en *kontinuerlig utvidelse*[6] av den opprinnelige funksjonen.

Vi spør altså: Enten f er definert i a eller ikke, kan vi (om)definere funksjonsverdien til f i a til å være et tall L slik at den nye funksjonen

$$\begin{array}{rcl} X \cup \{a\} & \longrightarrow & \mathbb{R} \\ x & \mapsto & \begin{cases} f(x), & x \in X \setminus \{a\}, \\ L, & x = a \end{cases} \end{array} \tag{271}$$

[6] *Continuous extension* på engelsk.

blir kontinuerlig? (Hvis $a \in X$, da er selvsagt $X \cup \{a\} = X$.) Hvis svaret er ja, sier vi at L er *grensen* eller *grenseverdien* til f når x nærmer seg a.

Vi merker først at spørsmålet ikke er interessant hvis a er et *isolert punkt* i $X \cup \{a\}$ (jf. Oppgave 9.6): dette betyr at det finnes en δ-omegn $(a-\delta, a+\delta)$ om a slik at $(X \cup \{a\}) \cap (a-\delta, a+\delta) = \{a\}$, som er ekvivalent med å si at

$$\text{det finnes en } \delta > 0 \text{ slik at } X \cap (a-\delta, a+\delta)^* = \emptyset, \tag{272}$$

hvor $(a-\delta, a+\delta)^* = (a-\delta, a+\delta) \setminus \{a\}$ er en *punktert omegn* om a, jf. Definisjon 2.1.23. I dette tilfellet vil funksjonen i (271) være kontinuerlig uansett L ved Oppgave 9.6. Vi betrakter derfor kun punkter a som *ikke* er isolerte punkter i $X \cup \{a\}$, det vil si slik at (272) *ikke* er oppfylt. Slike punkter har fått et eget navn:

Definisjon 9.10.1: Opphopningspunkt/grensepunkt

La $X \subset \mathbb{R}$. Et punkt $a \in \mathbb{R}$ er et *opphopningspunkt*[a] eller *grensepunkt*[b] til X dersom vi for alle $\delta > 0$ har at

$$X \cap (a-\delta, a+\delta)^* \neq \emptyset.$$

[a] *Accumulation point* på engelsk.

[b] *Limit point* på engelsk.

Sagt med ord betyr dette at *enhver omegn om a inneholder minst ett punkt i X utenom a.*

Følgende resultat klargjør navnebruken:

Setning 9.10.2: Karakterisering av opphopningspunkter

La $X \subset \mathbb{R}$. Et punkt $a \in \mathbb{R}$ er et opphopningspunkt til X hvis og bare hvis det finnes en følge $\{a_n\}$ slik at $a_n \in X$, $a_n \neq a$ for alle n og $a_n \to a$.

Bevis

Overlates til Oppgave 9.64. □

Eksempel 9.10.3. La $X = (0,1) \cup \{2\}$. Da er alle $x \in [0,1]$ opphopningspunkter til X, siden vi alltid kan finne en følge i $(0,1) \setminus \{x\} \subset X$ som konvergerer mot x. Samtidig kan ingen punkter i $\mathbb{R} \setminus [0,1]$ være opphopningspunkter til X; spesielt er ikke $\{2\}$ et opphopningspunkt, for det finnes ingen følge i $X \setminus \{2\} = (0,1)$ som konvergerer mot 2. Altså er mengden av alle opphopningspunkter til X lik intervallet $[0,1]$.

Vi kan skrive ned definisjonen på at funksjonen (271) er kontinuerlig i a ved å bruke Definisjon 9.1.1: Til enhver $\epsilon \in \mathbb{R}^+$ finnes en $\delta \in \mathbb{R}^+$ slik at

$$|f(x) - L| < \epsilon \text{ når } |x-a| < \delta \text{ og } x \in X \setminus \{a\}.$$

At $x \in X \setminus \{a\}$ er ekvivalent med $x \neq a$, eller med $0 < |x-a|$. Vi kan derfor skrive den formelle definisjonen på grenseverdi slik:

Definisjon 9.10.4: Grense

La $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$ og la $a \in \mathbb{R}$ være et opphopningspunkt til $D(f)$. Vi sier at grensen[a] til f er L når x nærmer seg a (uttrykt som "$\lim_{x\to a} f(x) = L$" eller "$f(x) \to L$ når $x \to a$"), hvis det til ethvert reelt tall $\epsilon > 0$ finnes et reelt tall $\delta > 0$ slik at

(273) $$|f(x) - L| < \epsilon \text{ når } 0 < |x-a| < \delta \text{ og } x \in D(f).$$

[a] *Limit* på engelsk.

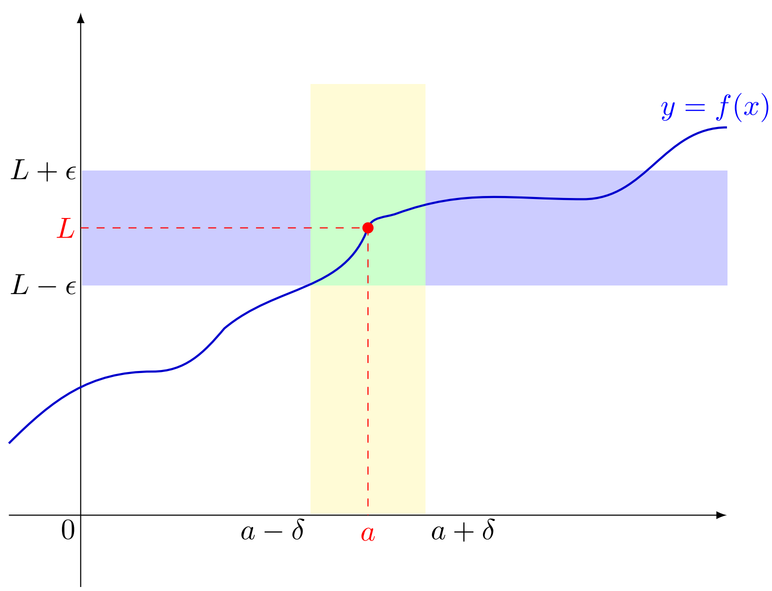
Merk at betingelsen $0 < |x-a| < \delta$ også kan skrives som

$$x \in (a-\delta, a+\delta) \setminus \{a\} = (a-\delta, a+\delta)^*,$$

slik at den betyr at x ligger i en punktert δ-omegn om a. Betingelsen $|f(x) - L| < \epsilon$ kan skrives som $f(x) \in (L-\epsilon, L+\epsilon)$, som er en ϵ-omegn om L. Hele betingelsen (273) kan derfor også uttrykkes som

(274) $$f(x) \in (L-\epsilon, L+\epsilon) \text{ når } x \in (a-\delta, a+\delta)^* \cap D(f)$$

og visualiseres i følgende figur:



$f(x)$ LIGGER I INTERVALLET $(L-\epsilon, L+\epsilon)$ SÅFREMT x LIGGER I DEN PUNKTERTE OMEGNEN $(a-\delta, a+\delta)^*$.

Intuitivt betyr dette at vi kan få $|f(x) - L|$ så liten vi vil bare $|x-a|$ er liten nok og $x \neq a$. Om funksjonen f er definert i a eller ikke, og hvilken verdi den eventuelt har i a, spiller ingen rolle.

Fra diskusjonen ovenfor er det klart at følgende holder for funksjoner som er definert i a:

Setning 9.10.5

La $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$ og la $a \in D(f)$ være et opphopningspunkt til $D(f)$.

Da er f kontinuerlig i a hvis og bare hvis $\lim_{x \to a} f(x) = f(a)$.

Merk imidlertid at kontinuitet også er definert i punkter som ikke er opphopningspunkter til $D(f)$ (og at f automatisk er kontinuerlig i slike punkter ved Oppgave 9.6).

Anvendelsen av definisjonen for grenser er i praksis helt lik anvendelsen av den formelle definisjonen for kontinuitet, bare at vi bytter ut $f(a)$ med L. Merk også at det overhodet ikke har noe å si om f er definert i a eller hva $f(a)$ er, som følgende eksempel viser:

Eksempel 9.10.6. La $f : \mathbb{R} \setminus \{3\} \to \mathbb{R}$ være definert ved

$$f(x) = \frac{2x^2 - 7x + 3}{x - 3}.$$

Vis at $\lim_{x \to 3} f(x) = 5$.

Løsning. Vi merker at 3 er et opphopningspunkt til definisjonsmengden $\mathbb{R} \setminus \{3\}$.

Siden $2x^2 - 7x + 3 = (2x - 1)(x - 3)$, kan vi skrive

$$f(x) = 2x - 1, \; x \neq 3.$$

Dette er samme funksjon som i Eksempel 9.1.4, bortsett fra at vi har tatt bort tallet 3 fra definisjonsmengden. Samme løsning som i Eksempel 9.1.4 viser at $\lim_{x \to 3} f(x) = 5$: Gitt $\epsilon > 0$. La $\delta = \frac{\epsilon}{2}$. Da, hvis $0 < |x - 3| < \delta = \frac{\epsilon}{2}$, er

$$|f(x) - f(3)| = |2x - 1 - 5| = |2x - 6| = 2|x - 3| < 2 \cdot \frac{\epsilon}{2} = \epsilon.$$

Betingelsen at $0 < |x - 3|$ garanterer at $x \neq 3$, slik at vi kan sette inn uttrykket $2x - 1$ for $f(x)$.

Vi kunne også brukt Setning 9.10.5 og sagt at siden funksjonen $x \mapsto 2x - 1$ er kontinuerlig, er $\lim_{x \to 3} (2x - 1) = 5$. Siden f er identisk lik denne funksjonen når $x \neq 3$, og definisjonen på grenseverdi forutsetter $x \neq 3$, kan vi konkludere at $\lim_{x \to 3} f(x) = \lim_{x \to 3} (2x - 1) = 2 \cdot 3 - 1 = 5$. □

Eksempel 9.10.7. Finn $\lim_{x \to \frac{\pi}{2}} f(x)$ for funksjonen $f : (0, \frac{\pi}{2}) \to \mathbb{R}$ gitt ved $f(x) = \sin x$.

Løsning. Vi merker at $\frac{\pi}{2}$ er et opphopningspunkt til definisjonsmengden $(0, \frac{\pi}{2})$.

Siden funksjonen $x \mapsto \sin x$ er kontinuerlig, har vi $\lim\limits_{x \to \frac{\pi}{2}} \sin x = \sin \dfrac{\pi}{2} = 1$ ved Setning 9.10.5. Siden f er lik denne funksjonen på sin innskrenkede definisjonsmengde, har vi $\lim\limits_{x \to \frac{\pi}{2}} f(x) = \lim\limits_{x \to \frac{\pi}{2}} \sin x = 1$. □

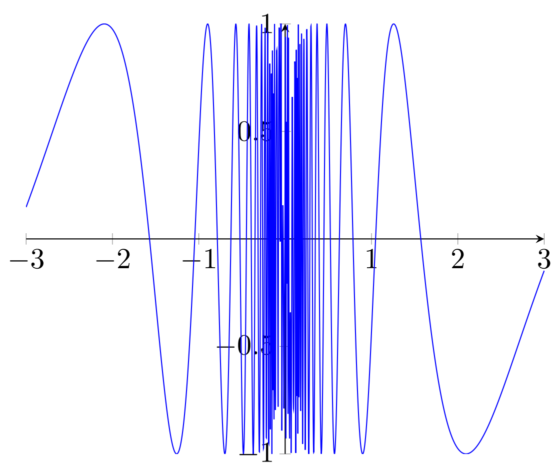
Noen ganger spesifiserer vi ikke definisjonsmengden til en funksjon f og skriver bare $\lim\limits_{x \to a} f(x)$. Når vi gjør det er det underforstått at vi gir f den størst mulige definisjonsmengden, det vil si, lik alle punkter hvor uttrykket er definert, og at a er et opphopningspunkt til denne. Dette er samme konvensjon som vi nevnte i §2.2.

Eksempel 9.10.8. Avgjør om grensen $\lim\limits_{x \to 0} \sin\left(\dfrac{1}{x}\right)$ eksisterer.

Løsning. Som nevnt ovenfor, er det underforstått at vi gir funksjonen $x \mapsto \sin\left(\frac{1}{x}\right)$ definisjonsmengden $\mathbb{R} \setminus \{0\} = (-\infty, 0) \cup (0, \infty)$. Grensen $\lim\limits_{x \to 0} \sin\left(\dfrac{1}{x}\right)$ eksisterer ikke. For hvis den eksisterte, si $\lim\limits_{x \to 0} \sin\left(\dfrac{1}{x}\right) = c$, ville funksjonen

$$f(x) = \begin{cases} \sin\left(\frac{1}{x}\right), & x \neq 0, \\ c, & x = 0, \end{cases}$$

vært kontinuerlig ved Setning 9.10.5, men vi viste i Eksemplene 9.3.1 og 9.4.6 at dette ikke er tilfellet, uansett $c \in \mathbb{R}$. □



Grafen til $x \mapsto \sin\left(\frac{1}{x}\right)$ $x \neq 0$.

Som vi så i eksemplene ovenfor, kan vi i mange tilfeller bruke Setning 9.10.5 for å finne grenseverdier; det gjelder dersom funksjonen vi betrakter er kontinuerlig i a, eller kan utvides til en kontinuerlig funksjon i a eller gjøres om til en kontinuerlig funksjon i a ved å endre funksjonsverdien i a.

Ikke overraskende har vi et følge-kriterium for grenseverdi, på lik linje med følge-kriteriet for kontinuitet (Setning 9.4.1). Dette brukes som definisjon på grenseverdi i mange lærebøker.

Setning 9.10.9: Følge-kriteriet for grense

La $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$ og la $a \in \mathbb{R}$ være et opphopningspunkt til $D(f)$.

Da er $\lim_{x \to a} f(x) = L$ hvis og bare hvis $\lim_{n \to \infty} f(a_n) = L$ for alle følger $\{a_n\}$ i $D(f) \setminus \{a\}$ som konvergerer mot a.

Bevis

Beviset er veldig likt beviset for Setning 9.4.1 og overlates til Oppgave 9.68. Betingelsen om at a er et opphopningspunkt til $D(f)$ garanterer ved Setning 9.10.2 at det finnes minst én følge i $D(f) \setminus \{a\}$ som konvergerer mot a. □

Siden grenser av følger er entydige (Setning 7.4.6), får vi som umiddelbar konsekvens at også grenser for funksjoner er entydige, dersom de eksisterer:

Følgesetning 9.10.10: Entydighet av grenser

Grenser er entydige, det vil si at hvis $\lim_{x \to a} f(x) = L$ og $\lim_{x \to a} f(x) = M$, da er $L = M$.

Eksempel 9.10.11. Betrakt funksjonen

$$\begin{aligned} f : (-\infty, 0) \cup (0, \infty) &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{1}{x}. \end{aligned}$$

Da er 0 et opphopningspunkt til $(-\infty, 0) \cup (0, \infty)$. Betrakt følgen $\{\frac{1}{n}\}$. Da er $\lim_{n \to \infty} \frac{1}{n} = 0$ ved Hjelpesetning 7.4.7, mens følgen $\{f(\frac{1}{n})\} = n$ divergerer siden den er ubegrenset (Setning 7.4.10). Ved Setning 9.10.9 eksisterer ikke $\lim_{x \to 0} f(x)$.

Ensidige grenser. Når vi snakker om ensidige grenser, betyr det i praksis at vi kun studerer funksjonens oppførsel på den ene siden av et punkt a. Med andre ord betrakter vi funksjonen $f : D(f) \to \mathbb{R}$ sin oppførsel når vi restriserer den til $D(f) \cap (-\infty, a)$ eller $D(f) \cap (a, \infty)$ (jf. Definisjon 2.2.10). Ellers er tankegangen og definisjonen de samme som i forrige delseksjon; den eneste forskjellen er at vi bytter ut hele den punkterte omegnen $(a - \delta, a + \delta)^* = (a - \delta, a + \delta) \setminus \{a\} = (a - \delta, a) \cup (a, a + \delta)$ i (274) med kun delen på den ene siden, $(a - \delta, a)$ eller $(a, a + \delta)$.

Definisjonen for ensidige grenser blir da som følger:

Definisjon 9.10.12: Ensidige grenser

La $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$.

La $a \in \mathbb{R}$ være et opphopningspunkt til $D(f) \cap (a, \infty)$. Vi sier at *grensen til f er L når punktet x nærmer seg punktet a ovenfra* (uttrykt som $\lim_{x \to a^+} f(x) = L$), hvis det til ethvert reelt tall $\epsilon > 0$ finnes et reelt tall $\delta > 0$ slik at

$$|f(x) - L| < \epsilon \text{ når } a < x < a + \delta. \tag{275}$$

La $a \in \mathbb{R}$ være et opphopningspunkt til $D(f) \cap (-\infty, a)$. Vi sier at *grensen til f er L når punktet x nærmer seg punktet a nedenfra* (uttrykt som $\lim_{x \to a^-} f(x) = L$), hvis det til ethvert reelt tall $\epsilon > 0$ finnes et reelt tall $\delta > 0$ slik at

$$|f(x) - L| < \epsilon \text{ når } a - \delta < x < a. \tag{276}$$

Merknad 9.10.13. Når funksjonen i utgangspunktet kun er definert på den ene siden av a, er det egentlig det samme om vi skriver $\lim_{x \to a} f(x)$ eller $\lim_{x \to a^\pm} f(x)$. Det gjelder for eksempel funksjonen $f : [0, \infty) \to \mathbb{R}$ gitt ved $f(x) = \sqrt{x}$. Her kan vi skrive både $\lim_{x \to 0} f(x) = 0$ og $\lim_{x \to 0+} f(x) = 0$. Det er egentlig unødvendig å bruke siste typen notasjon når vi allerede har spesifisert definisjonsmengden til å være $(0, \infty)$, men mange vil foretrekke å skrive $\lim_{x \to 0^+} \sqrt{x}$ hvis man ikke eksplisitt har spesifisert definisjonsmengden på forhånd.

Bruken av disse definisjonene for å vise ensidige grenser er egentlig helt lik bruken for å vise tosidige grenser.

Vi har igjen et følge-kriterium:

Setning 9.10.14: Følge-kriteriet for ensidig grense

La $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$ og la $a \in \mathbb{R}$ være et opphopningspunkt til $D(f) \cap (a, \infty)$ (henholdsvis $D(f) \cap (-\infty, a)$).

Da er $\lim_{x \to a^+} f(x) = L$ (hhv. $\lim_{x \to a^-} f(x) = L$) hvis og bare hvis $\lim_{n \to \infty} f(a_n) = L$ for alle følger $\{a_n\}$ i $D(f) \cap (a, \infty)$ (hhv. $D(f) \cap (-\infty, a)$) som konvergerer mot a.

Bevis

Beviset er igjen veldig likt beviset for Setning 9.4.1 og overlates til Oppgave 9.68. $\square$

Som konsekvens gjelder Setning 9.10.10 om entydighet av grenser også for ensidige grenser.

Vi merker oss dessuten at ved å sammenligne definisjonene 9.10.4 og 9.10.12, så får vi:

Setning 9.10.15

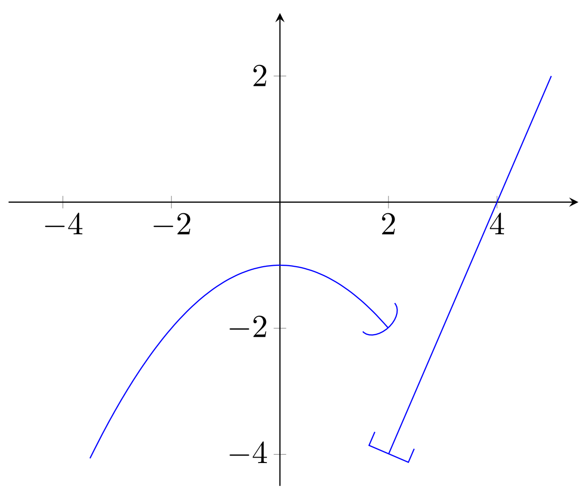
$$\lim_{x\to a} f(x) = L \iff \lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x) = L$$

Dette resultatet gir oss ofte en enkel måte å bevise at en (tosidig) grense ikke eksisterer:

EKSEMPEL 9.10.16. La oss betrakte funksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved

$$f(x) = \begin{cases} -\frac{1}{4}x^2 - 1, & x < 2 \\ 2x - 8, & x \geq 2, \end{cases}$$

allerede betraktet i Eksemplene 9.2.15, 9.4.5 og 9.5.18, med graf:



Ved å bruke at begge funksjonene $x \mapsto -\frac{1}{4}x^2 - 1$ og $x \mapsto 2x - 8$ er kontinuerlige (siden de er gitt ved polynomer), finner vi, ved å bruke Setning 9.10.5:

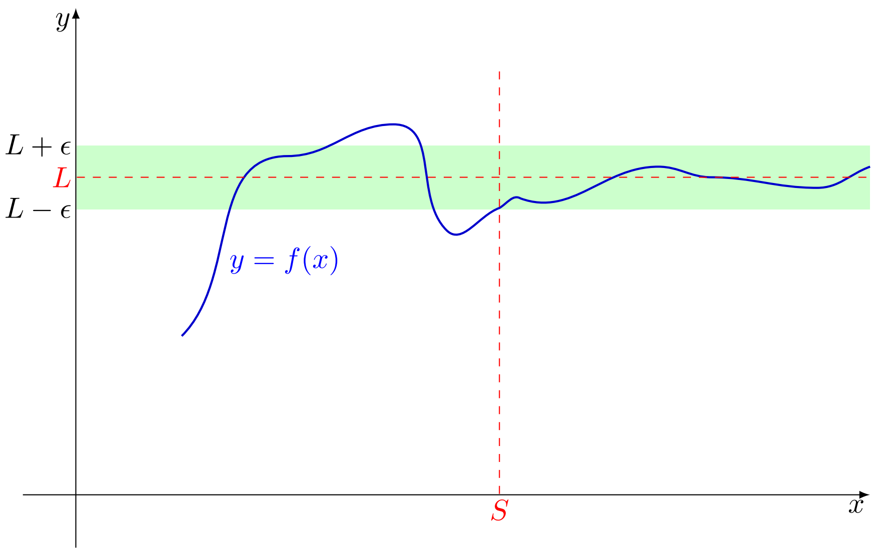
$$\lim_{x\to 2^-} f(x) = \lim_{x\to 2^-} \left(-\frac{1}{4}x^2 - 1\right) = \lim_{x\to 2} \left(-\frac{1}{4}x^2 - 1\right) = -\frac{1}{4} \cdot 2^2 - 1 = -2$$

og

$$\lim_{x\to 2^+} f(x) = \lim_{x\to 2^+} (2x - 8) = \lim_{x\to 2} (2x - 8) = 2 \cdot 2 - 8 = -4.$$

Imidlertid eksisterer ikke grensen $\lim_{x\to 2} f(x)$ ved Setning 9.10.15, siden vi har at $\lim_{x\to 2^+} f(x) \neq \lim_{x\to 2^-} f(x)$.

Grenser i uendelig. Noen ganger er vi interessert i å studere oppførselen til en funksjon når x blir større og større, det vil si "når x går not uendelig". (Her er det igjen viktig å være klar over at "uendelig", selv om det blir gitt et symbol, nemlig ∞, *ikke* er et tall, men brukes kun for å betegne oppførsler til størrelser.) En funksjon kan oppføre seg på flere forskjellige måter: den kan vokse seg større og større i positiv eller negativ retning (som $f(x) = x$ eller $f(x) = -x$), den kan svinge og svinge uten å nærme seg noen bestemt verdi (som $f(x) = \cos x$ eller $f(x) = \sin x$), eller den kan "slakke" mot en bestemt verdi L som i figuren nedenunder:



f NÆRMER SEG L NÅR x GÅR NOT UENDELIG.

Vi skriver $\lim\limits_{x\to\infty} f(x) = L$ og mener med dette, litt uformelt formulert, at

Vi kan sørge for at $f(x)$ er så nær vi vil L så lenge x er stor nok.

Som før, betyr "$f(x)$ så nær vi vil L" at $|f(x) - L| < \epsilon$ for enhver gitt $\epsilon > 0$. At "x er stor nok", betyr "$x > S$ for en S som avhenger av ϵ", se figuren.

Selvsagt må vi kreve at definisjonsmengden $D(f)$ ikke er oppad begrenset for at dette skal gi mening.

Den formelle definisjonen blir da:

Definisjon 9.10.17: Grenser mot uendelig

La $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$ som ikke er oppad begrenset. Vi sier at *grensen til f er L når punktet x går mot ∞ (uendelig)* (uttrykt som $\lim\limits_{x\to\infty} f(x) = L$), hvis det til ethvert reelt tall $\epsilon > 0$ finnes et reelt tall S slik at

$$|f(x) - L| < \epsilon \quad \text{når} \quad x > S \quad \text{og} \quad x \in D(f). \tag{277}$$

Vi får en helt tilsvarende definisjon for grenser mot $-\infty$:

Definisjon 9.10.18: Grenser mot minus uendelig

La $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$ som ikke er nedad begrenset. Vi sier at *grensen til f er L når punktet x går mot $-\infty$* (uttrykt som $\lim_{x\to-\infty} f(x) = L$), hvis det til ethvert reelt tall $\epsilon > 0$ finnes et reelt tall S slik at

$$|f(x) - L| < \epsilon \text{ når } x < S \text{ og } x \in D(f). \tag{278}$$

Igjen har vi et følge-kriterium:

Setning 9.10.19: Følge-kriteriet for grense mot uendelig

La $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$ som ikke er oppad (henholdsvis nedad) begrenset.

Da er $\lim_{x\to\infty} f(x) = L$ (hhv. $\lim_{x\to-\infty} f(x) = L$) hvis og bare hvis $\lim_{n\to\infty} f(a_n) = L$ for alle følger $\{a_n\}$ i $D(f)$ som divergerer mot ∞ (hhv. $-\infty$).

Bevis

Beviset er igjen veldig likt beviset for Setning 9.4.1 og overlates til Oppgave 9.68. $\square$

Som konsekvens gjelder Setning 9.10.10 om entydighet av grenser også for grenser mot uendelig.

Bruk av disse definisjonen for å vise grenser mot uendelig er litt (men ikke så mye) annerledes enn bruken for å vise de andre grensene.

Eksempel 9.10.20. La $r \in \mathbb{R}^+$. Vis at $\lim_{x\to\infty} \frac{1}{x^r} = 0$.

Løsning. Gitt $\epsilon > 0$. Vi har, for $x \in \mathbb{R}^+$:

$$\begin{aligned}
\left|\tfrac{1}{x^r}\right| < \epsilon &\iff \tfrac{1}{\epsilon} < x^r && (x^r > 0)\\
&\iff \ln(1/\epsilon) < \ln x^r && (\ln \text{ strengt voksende (Setn. 9.5.14))}\\
&\iff \ln(1/\epsilon) < r \ln x && \text{(Oppgave 9.18(e))}\\
&\iff \tfrac{\ln(1/\epsilon)}{r} < \ln x &&\\
&\iff e^{\frac{\ln(1/\epsilon)}{r}} < e^{\ln x} && (e \mapsto e^x \text{ strengt voks. (Merkn. 7.3.11))}\\
&\iff e^{\frac{\ln(1/\epsilon)}{r}} < x && ((230))
\end{aligned}$$

Vi setter da $S = e^{\frac{\ln(1/\epsilon)}{r}}$. Vi har da vist at når $x > S$, da er $\left|\frac{1}{x^r}\right| < \epsilon$. Dette viser at $\lim_{x\to\infty} \frac{1}{x^r} = 0$. $\square$

Eksempel 9.10.21. La $a \in \mathbb{R}^+$. Vis at

(i) $\lim_{x\to\infty} a^x = 0$ hvis $0 < a < 1$,

(ii) $\lim_{x\to-\infty} a^x = 0$ hvis $a > 1$.

LØSNING. Anta først $a > 1$. Gitt $\epsilon > 0$. Vi har, for $x \in \mathbb{R}^+$:

$$\begin{aligned} |a^x| < \epsilon &\iff a^x < \epsilon && (a^x > 0) \\ &\iff \log_a a^x < \log_a \epsilon && (x \mapsto \log_a x \text{ strengt voks. (Setn. 9.5.14)}) \\ &\iff x < \log_a \epsilon && ((231)) \end{aligned}$$

Vi setter da $S = \log_a \epsilon$. Vi har da vist at når $x < S$, da er $|a^x| < \epsilon$. Dette viser at $\lim_{x\to-\infty} a^x = 0$.

Tilfellet $a < 1$ er nesten helt identisk: vi får snudd ulikheten i midterste linje av ekvivalensrekken siden $x \mapsto \log_a x$ er strengt avtagende i dette tilfellet (Setning 9.5.14), slik at vi får

$$|a^x| < \epsilon \iff x > \log_a \epsilon.$$

Vi setter da fremdeles $S = \log_a \epsilon$ og har da vist at når $x > S$, da er $|a^x| < \epsilon$. Dette viser at $\lim_{x\to\infty} a^x = 0$. □

EKSEMPEL 9.10.22. Vis at $\lim_{x\to\infty} \cos x$ ikke eksisterer.

LØSNING. Betrakt følgen $\{x_n\} = \{\pi n\}$. Da er $\{\cos x_n\} = \{(-1)^n\}$ divergent. Ved følge-kriteriet 9.10.19 eksisterer ikke grensen $\lim_{x\to\infty} \cos x$. □

Egenskaper ved grenser. For å finne grenser til funksjoner generelt, er det nyttig å ha generelle resultater og regler for beregning av grenser istedenfor å bruke definisjonene hver gang. I praksis bruker vi derfor først den formelle definisjonen av grenseverdi til å bevise en del resultater om grenser, som vi så bruker for å beregne grenseverdier.

Vi oppsummerer her de viktigste resultatene som gjør oss i stand til å beregne grenser i mange tilfeller.

Setning 9.10.23: Grensesetningene

Anta at grensene $\lim_{x\to a} f(x)$ og $\lim_{x\to a} g(x)$ eksisterer og $k \in \mathbb{R}$. Da gjelder:

(i) $\lim_{x\to a} \Big(f(x) \pm g(x)\Big) = \lim_{x\to a} f(x) \pm \lim_{x\to a} g(x)$;

(ii) $\lim_{x\to a} f(x)g(x) = \Big(\lim_{x\to a} f(x)\Big)\Big(\lim_{x\to a} g(x)\Big)$;

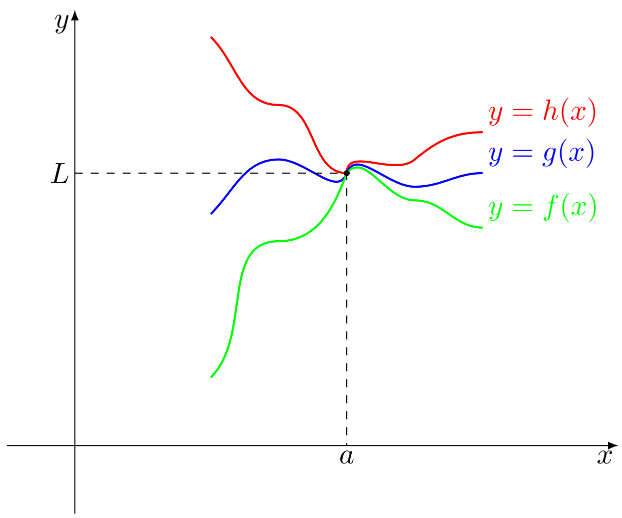
(iii) $\lim_{x\to a} \dfrac{f(x)}{g(x)} = \dfrac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)}$, dersom $\lim_{x\to a} g(x) \neq 0$;

(iv) $\lim_{x\to a} kf(x) = k \lim_{x\to a} f(x)$.

Bevis
Dette er en umiddelbar konsekvens av *følge-kriteriet for grenser* (Setning 9.10.14) og de tilsvarende regnereglene for grenser av følger (Setning 7.4.8). Alternativt kan setningen vises direkte ved ϵ–δ definisjonen av grenser, jf. Oppgave 9.63. □

Setning 9.10.24: Skvisesetningen

Anta at $f(x) \leq g(x) \leq h(x)$ for alle x i en punktert omegn om a og at $\lim_{x\to a} f(x) = \lim_{x\to a} h(x) = L$. Da vil også $\lim_{x\to a} g(x) = L$.



Skvisesetningen for funksjoner

Bevis
Dette er en umiddelbar konsekvens av *følge-kriteriet for grenser* (Setning 9.10.14) og *skvisesetningen for følger* (Setning 7.4.13). Alternativt kan setningen vises direkte ved ϵ–δ definisjonen av grenser, jf. Oppgave 9.63. □

Setning 9.10.25: Bevaring av orden

Dersom $f(x) \leq g(x)$ for alle x i en punktert omegn om a og begge grensene $\lim_{x\to a} f(x)$ og $\lim_{x\to a} g(x)$ eksisterer, da er

$$\lim_{x\to a} f(x) \leq \lim_{x\to a} g(x).$$

Bevis
Dette er en umiddelbar konsekvens av *følge-kriteriet for grenser* (Setning 9.10.14) og det tilsvarende utsagnet at grenser av følger bevarer orden (Setning 7.4.14). Alternativt kan setningen vises direkte ved ϵ–δ definisjonen av grenser, jf. Oppgave 9.63. □

Setning 9.10.26: Grense av sammensatte funksjoner

Hvis f er kontinuerlig i L og $\lim_{x\to a} g(x) = L$, da er

$$\lim_{x\to a} f(g(x)) = f(L). \tag{279}$$

Bevis

Vi bruker igjen følge-kriteriet for grenser av funksjoner (Setning 9.10.14), og vil derfor vise at for enhver følge $\{x_n\}$ i definisjonsmengden til $f \circ g$ som konvergerer mot a, vil $f(g(x_n))$ konvergere mot $f(L)$.

Siden $\lim_{x\to a} g(x) = L$, har vi ved følge-kriteriet (Setning 9.10.14) anvendt på g at $g(x_n) \to L$. Siden f er kontinuerlig, har vi ved følge-kriteriet for kontinuitet (Setning 9.4.1 brukt på f og følgen $\{g(x_n)\}$) at $f(g(x_n)) \to f(L)$. Dette fullfører beviset.

Alternativt kan setningen vises direkte ved ϵ–δ definisjonen av grenser, jf. Oppgave 9.63. $\square$

Vi husker (279) kanskje best som :

$$\lim_{x\to a} f(g(x)) = f(\lim_{x\to a} g(x)),$$

det vil si

$$\boxed{\text{grensen kan "trekkes innenfor" en kontinuerlig funksjon.}}$$

Dersom $\lim_{x\to a} g(x))$ eksisterer, så har vi for eksempel at

$$\begin{aligned}
\lim_{x\to c} g(x)^r &= \left(\lim_{x\to c} g(x)\right)^r, \quad r \in \mathbb{R}^+, \\
\lim_{x\to c} a^{g(x)} &= a^{\left(\lim_{x\to c} g(x)\right)}, \quad a \in \mathbb{R}^+, \\
\lim_{x\to c} \log_a |g(x)| &= \log_a \left|\lim_{x\to c} g(x)\right|, \\
\lim_{x\to c} \sin g(x) &= \sin\left(\lim_{x\to c} g(x)\right),
\end{aligned}$$

og så videre.

Merknad 9.10.27. Alle Setningene 9.10.23, 9.10.25, 9.10.24 og 9.10.26 er også gyldige for ensidige grenser og grenser i uendelig, ved å bytte ut $\lim_{x\to a}$ med $\lim_{x\to a^-}$, $\lim_{x\to a^+}$, $\lim_{x\to -\infty}$, $\lim_{x\to\infty}$. I Setningene 9.10.25 og 9.10.24 bytter man ut "en punktert omegn om a" med et intervall på formene:

- $(a-\delta, a) \cap D(f)$ for en $\delta > 0$ i tilfellet $\lim_{x\to a^-}$,
- $(a, a+\delta) \cap D(f)$ for en $\delta > 0$ i tilfellet $\lim_{x\to a^+}$,
- $(-\infty, c) \cap D(f)$ for en $c \in \mathbb{R}$ i tilfellet $\lim_{x\to -\infty}$,
- $(c, \infty) \cap D(f)$ for for en $c \in \mathbb{R}$ i tilfellet $\lim_{x\to\infty}$.

Ved hjelp av resultatene ovenfor, sammen med Setning 9.10.5, kan vi finne grensen til mange kjente funksjoner. La oss se på noen eksempler og starte med ett som også belyser feil bruk av grensesetningene:

EKSEMPEL 9.10.28. La $n \in \mathbb{Z}^+$. Finn $\lim_{x\to 0} x^n \sin\left(\frac{1}{x}\right)$.

FEIL LØSNING (MEN MED RETT SVAR). Ved produktregelen for grenser, Setning 9.10.23(ii), har vi at

$$\lim_{x\to 0} x^n \sin\left(\frac{1}{x}\right) = \lim_{x\to 0} x^n \cdot \lim_{x\to 0} \sin\left(\frac{1}{x}\right) = 0 \cdot \lim_{x\to 0} \sin\left(\frac{1}{x}\right) = 0.$$

□

Hvor ligger feilen? Produktregelen i grensesetningene forutsetter at *begge grensene i produktet eksisterer*, men i dette tilfellet eksisterer ikke $\lim_{x\to 0} \sin\left(\frac{1}{x}\right)$, som vi så i Eksempel 9.10.8, slik at regelen ikke kan brukes.

RIKTIG LØSNING. I tilfellet $n = 2$ kan vi bruke Eksempel 9.2.5: der viste vi nemlig at funksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & x \neq 0, \\ 0, & x = 0 \end{cases}$$

er kontinuerlig i 0, og dette sammen med Setning 9.10.5 gir oss at $\lim_{x\to 0} x^2 \sin\left(\frac{1}{x}\right) = 0$. Et helt tilsvarende ϵ–δ resonnement som i Eksempel 9.2.5 kan brukes til å vise at funksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved

$$f(x) = \begin{cases} x^n \sin\left(\frac{1}{x}\right), & x \neq 0, \\ 0, & x = 0 \end{cases}$$

er kontinuerlig i 0, slik at $\lim_{x\to 0} x^n \sin\left(\frac{1}{x}\right) = 0$, uansett $n \in \mathbb{Z}^+$. Dette beviset er så å si identisk med et direkte ϵ–δ bevis for grensen.

Vi kan imidlertid unngå dette ved å bruke *Skvisesetningen* 9.10.24, som vi nå viser: Vi har

$$-1 \leq \sin\left(\frac{1}{x}\right) \leq 1, \quad x \neq 0$$

Ganger vi med x^n må vi passe på å snu ulikhetene når $x < 0$ og n er et oddetall. Vi må derfor betrakte tilfellene $x > 0$ og $x < 0$ hver for seg. Vi kan imidlertid unngå denne todelingen ved å starte med

$$0 \leq \left|\sin\left(\frac{1}{x}\right)\right| \leq 1$$

og gange med $|x^n|$:

$$0 \leq \left|x^n \sin\left(\frac{1}{x}\right)\right| \leq |x^n|\,.$$

Dette er ekvivalent med

$$-|x^n| \le x^n \sin\left(\frac{1}{x}\right) \le |x^n|\,.$$

Siden $x \to |x^n|$ er en kontinuerlig funksjon (fordi den er sammensetningen av absoluttverdifunksjonen og en polynomiell funksjon), har vi ved Setning 9.10.5 at $\lim_{x\to 0} |x^n| = 0$. *Skvisesetningen* 9.10.24 gir dermed at også

$$\lim_{x\to 0} x^n \sin\left(\frac{1}{x}\right) = 0.$$

□

EKSEMPEL 9.10.29. Finn $\lim_{x\to 1} \dfrac{2x^2+8x-10}{x^2+2x-3}$.

LØSNING. Vi merker oss at både teller og nevner er kontinuerlige funksjoner, som går mot null, slik at vi ikke kan bruke kvotientregelen i Setning 9.10.23(iii). Mer nøyaktig er både teller og nevner null for $x = 1$, og dette betyr jo i henhold til *Faktorteoremet* (Følgesetning 4.1.9) at $x - 1$ er faktor i både teller og nevner. Vi forkorter og finner:

$$\begin{aligned}\frac{2x^2+8x-10}{x^2+2x-3} &= \frac{2(x^2+4x-5)}{x^2+2x-3} = \frac{2(x-1)(x+5)}{(x-1)(x+3)} \\ &= \frac{2\cancel{(x-1)}(x+5)}{\cancel{(x-1)}(x+3)} \overset{x\neq 1}{=} \frac{2(x+5)}{x+3}\end{aligned}$$

Forkortningen i siste likhet kan gjøres såfremt $x \neq 1$; med andre ord, funksjonene gitt ved $x \mapsto \frac{2x^2+8x-10}{x^2+2x-3}$ og $x \mapsto \frac{2(x+5)}{x+3}$ er nøyaktig samme funksjon når $x \neq 1$ (men sistnevnte funksjon er også definert i $x = 1$, mens den første er det ikke). Siden definisjon av grense ikke avhenger av det som skjer i $x = 1$, kan vi altså trygt forkorte uttrykk på denne måten når vi skal regne ut grensen når $x \to 1$. Siden $\lim_{x\to 1} 2(x+5) = 2\cdot(1+5) = 12$ og $\lim_{x\to 1}(x+3) = 1+3 = 4$ (ved reglene (i), (ii) og (iv) i Setning 9.10.23), finner vi ved kvotientregelen i Setning 9.10.23(iii) at

$$\lim_{x\to 1} \frac{2x^2+8x-10}{x^2+2x-3} = \lim_{x\to 1} \frac{2(x+5)}{x+3} = \frac{12}{4} = 3.$$

Selvsagt kan vi også bruke at funksjonen $x \to \frac{2(x+5)}{x+3}$ er kontinuerlig i 1 og har dermed grense 3 ved Setning 9.10.5. □

EKSEMPEL 9.10.30. Finn $\lim_{x\to 0} \dfrac{\sqrt{x+25}-5}{x}$.

LØSNING. Vi merker igjen at både teller og nevner går mot 0. Vi omskriver ved å gange både teller og nevner med $\sqrt{x+25}+5$:

$$\begin{aligned}\frac{\sqrt{x+25}-5}{x} &= \frac{(\sqrt{x+25}-5)(\sqrt{x+25}+5)}{x(\sqrt{x+25}+5)} = \frac{(\sqrt{x+25})^2-5^2}{x(\sqrt{x+25}+5)}\\ &= \frac{|x+25|-25}{x(\sqrt{x+25}+5)} \overset{x\geq -25}{=} \frac{x+25-25}{x(\sqrt{x+25}+5)}\\ &= \frac{\cancel{x}}{\cancel{x}\sqrt{x+25}+5)} \overset{x\neq 0}{=} \frac{1}{\sqrt{x+25}+5}.\end{aligned}$$

Merk at når vi "løste opp" absoluttverditegnet, så måtte vi anta at vi hadde $x+25 \geq 0$, hvilket vi godt kan gjøre når vi beregner grensen når $x \to 0$, for da er x i en liten punktert omegn om 0. Siden funksjonen $x \mapsto \frac{1}{\sqrt{x+25}+5}$ er kontinuerlig, er

$$\lim_{x\to 0} \frac{1}{\sqrt{x+25}+5} = \frac{1}{\sqrt{0+25}+5} = \frac{1}{10}$$

ved Setning 9.10.5. Dermed har vi

$$\lim_{x\to 0} \frac{\sqrt{x+25}-5}{x} = \lim_{x\to 0} \frac{1}{\sqrt{x+25}+5} = \frac{1}{10}.$$

□

EKSEMPEL 9.10.31. Vis at $\lim_{x\to 0} \frac{\cos x - 1}{x} = 0$.

LØSNING. Vi skriver

$$\begin{aligned}\frac{1-\cos x}{x} &= \frac{(1-\cos x)(1+\cos x)}{x(1+\cos x)} = \frac{1-\cos^2 x}{x(1+\cos x)}\\ &= \frac{\sin^2 x}{x(1+\cos x)} = \frac{\sin x}{x} \cdot \frac{\sin x}{1+\cos x}.\end{aligned}$$

Siden funksjonene $x \mapsto \sin x$ og $x \mapsto 1+\cos x$ er kontinuerlige, og sistnevnte er $\neq 0$ i 0, gir Setning 9.5.1 at funksjonen $x \mapsto \frac{\sin x}{1+\cos x}$ er kontinuerlig i 0. Dermed er $\lim_{x\to 0} \frac{\sin x}{1+\cos x} = \frac{\sin 0}{1+\cos 0} = 0$. Siden funksjonen

$$x \mapsto \begin{cases} \frac{\sin x}{x}, & x \neq 0,\\ 1, & x = 0\end{cases}$$

er kontinuerlig i 0 ved Merknad 9.2.9, er $\lim_{x\to 0} \frac{\sin x}{x} = 1$. Ved Setning 9.10.23(ii) har vi dermed

$$\begin{aligned}\lim_{x\to 0} \frac{\cos x - 1}{x} &= \lim_{x\to 0} \left(\frac{\sin x}{x} \cdot \frac{\sin x}{1+\cos x}\right)\\ &= \lim_{x\to 0} \left(\frac{\sin x}{x}\right) \lim_{x\to 0} \left(\frac{\sin x}{1+\cos x}\right) = 1 \cdot 0 = 0.\end{aligned}$$

□

Vi har merket flere ganger at tilfellet $\lim_{x\to a} \frac{f(x)}{g(x)}$ med $\lim_{x\to a} g(x) = 0$ ikke dekkes av Setning 9.10.23. I de siste tre eksemplene så vi at vi imidlertid kunne beregne grensene likevel, ved å jobbe litt med uttrykkene. I alle disse tilfellene hadde vi at *både* f og g gikk mot null. Dersom f ikke går mot null, eksisterer ikke grensen ved følgende:

Observasjon 9.10.32

Hvis $\lim_{x\to a} g(x) = 0$ og $\lim_{x\to a} f(x) \neq 0$, da eksisterer ikke $\lim_{x\to a} \frac{f(x)}{g(x)}$.

Bevis
Vi argumenterer ved selvmotsigelse.

Anta at grensen eksisterer, si $\lim_{x\to a} \frac{f(x)}{g(x)} = L$. Da gir Setning 9.10.23(ii) at

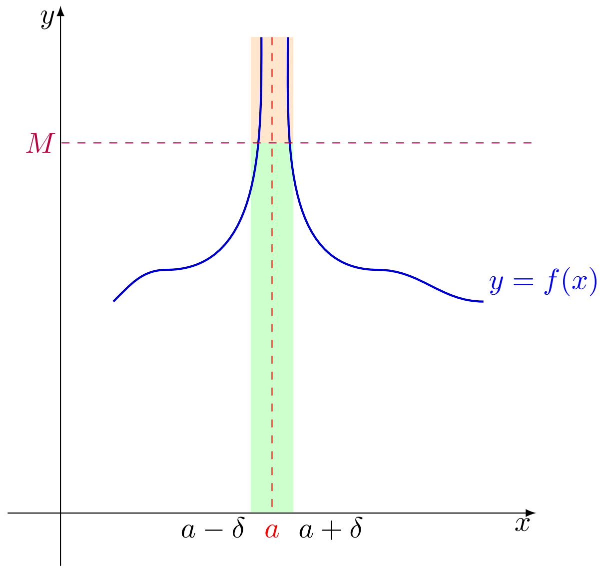
$$\lim_{x\to a} f(x) = \lim_{x\to a} \frac{f(x)\cdot g(x)}{g(x)} = \lim_{x\to a} \left(\frac{f(x)}{g(x)} \cdot g(x)\right) = L \cdot 0 = 0,$$

en selvmotsigelse. □

Merknad 9.10.33. Dersom både $\lim_{x\to a} g(x) = \lim_{x\to a} f(x) = 0$, kan grensen $\lim_{x\to a} \frac{f(x)}{g(x)}$ både eksistere og ikke eksistere. Vanlige angrepsmetoder er å forsøke å faktorisere (som i Eksemplene 9.10.29 og 9.10.30) eller bruke l'Hôpitals regel, som vi ikke skal betrakte her (se [**AE**, §4.3] eller [**Li**, §6.3]).

Uendelige grenser. I Eksemplene 9.10.8 og 9.10.16 så vi grenser som ikke eksisterte: i det første tilfellet eksisterer ikke grensen fordi funksjonen fortsetter å svinge mellom 1 og -1 uansett hvor mye vi nærmer oss punktet, mens i det andre nærmer funksjonen seg to forskjellige verdier fra hver side.

Det finnes imidlertid andre måter en grense ikke eksisterer på, som at funksjonen vokser ubegrenset inn mot punktet, som på figuren nedenunder:



f går not uendelig når x går mot a.

Selv om grensen ikke eksisterer, er dette likevel en oppførsel som vi ønsker å betegne med et eget begrep.

Vi skriver $\lim_{x \to a} f(x) = \infty$ og mener med dette, litt uformelt formulert, at

Vi kan sørge for at $f(x)$ er så stor vi vil så lenge x er nær nok a (og $x \neq a$).

Den formelle definisjonen er:

Definisjon 9.10.34: Uendelig grense

La $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$ og la $a \in \mathbb{R}$ være et opphopningspunkt til $D(f)$. Vi sier at grensen til f er ∞ når punktet x nærmer seg punktet a (uttrykt som $\lim_{x \to a} f(x) = \infty$), hvis det til ethvert reelt tall M finnes et reelt tall $\delta > 0$ slik at

$$f(x) > M \quad \text{når} \quad 0 < |x - a| < \delta \quad \text{og} \quad x \in D(f). \tag{280}$$

Vi får en helt tilsvarende definisjon for grense $-\infty$:

Definisjon 9.10.35: Negativ uendelig grense

La $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$ og la $a \in \mathbb{R}$ være et opphopningspunkt til $D(f)$. Vi sier at grensen til f er $-\infty$ når punktet x nærmer seg punktet a (uttrykt som $\lim_{x \to a} f(x) = \infty$), hvis det til ethvert reelt tall M finnes et reelt tall $\delta > 0$ slik at

$$f(x) < M \quad \text{når} \quad 0 < |x - a| < \delta \quad \text{og} \quad x \in D(f). \tag{281}$$

Som vanlig har vi også et følge-kriterium:

Setning 9.10.36: Følge-kriteriet for uendelig grense

La $f : D(f) \to \mathbb{R}$ være en funksjon med definisjonsmengde $D(f) \subset \mathbb{R}$ og la $a \in \mathbb{R}$ være et opphopningspunkt til $D(f)$. Da er $\lim_{x\to a} f(x) = \infty$ (henholdsvis $-\infty$) hvis og bare hvis $\lim_{n\to\infty} f(a_n) = \infty$ (hhv. $-\infty$) for alle følger $\{a_n\}$ i $D(f) \setminus \{a\}$ som konvergerer mot a.

BEVIS
Beviset er veldig likt beviset for Setning 9.4.1 og overlates til Oppgave 9.68. □

EKSEMPEL 9.10.37. La n være et positivt partall. Vis at $\lim_{x\to 0} \frac{1}{x^n} = \infty$.

LØSNING. Gitt $M \in \mathbb{R}$. Vi har

$$\frac{1}{x^n} > |M| \iff \frac{1}{\sqrt[n]{|M|}} > |x|.$$

Vi har dermed vist at $\frac{1}{x^n} > M$ såfremt $0 < |x| < \frac{1}{\sqrt[n]{|M|}}$, som viser grensen. □

Det finnes tilsvarende opplagte definisjoner for ensidige uendelige grenser $\lim_{x\to a^+} f(x) = \infty$ og $\lim_{x\to a^-} f(x) = \infty$, og likeledes for $-\infty$. I samme ånd som Eksempel 9.10.37, kan man for eksempel vise at

$$\lim_{x\to 0^-} \frac{1}{x^n} = -\infty \quad \text{og} \quad \lim_{x\to 0^+} \frac{1}{x^n} = \infty \quad \text{for } n \text{ (positivt) oddetall,} \tag{282}$$

jf. Oppgave 9.72.

EKSEMPEL 9.10.38. Betrakt tangensfunksjonen restrisert til $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$:

$$\begin{aligned} f : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) &\longrightarrow \mathbb{R} \\ x &\mapsto \tan x = \frac{\sin x}{\cos x}, \end{aligned}$$

som vi studerte i Eksempel 9.6.11. Der viste vi at uansett gitt M, finnes en $\delta' > 0$ slik at

$$f(x) \geq M, \quad \text{for } x \in \left(\frac{\pi}{2} - \delta', \frac{\pi}{2}\right).$$

Dette viser at $\lim_{x\to\frac{\pi}{2}} f(x) = \infty$, eller, om vi ikke spesifiserer den restriserte definisjonsmengden, $\lim_{x\to\frac{\pi}{2}^-} \tan x = \infty$. På helt tilsvarende måte har vi $\lim_{x\to-\frac{\pi}{2}} f(x) = \lim_{x\to-\frac{\pi}{2}^+} \tan x = -\infty$.

Siden dette betyr at $f(x) \to \infty$ (henholdsvis $-\infty$) når $x \to \frac{\pi}{2}$ (hhv. $x \to -\frac{\pi}{2}$), får vi også grensen $\lim_{x\to\infty} \tan^{-1} x = \dfrac{\pi}{2}$ (hhv. $\lim_{x\to-\infty} \tan^{-1} x = -\dfrac{\pi}{2}$) (jf. Merknad 9.6.12).

Den siste typen grenseverdi er den hvor funksjonen går mot ∞ eller $-\infty$ når x går mot ∞ eller $-\infty$. Vi skriver ned bare én av de fire variantene:

Definisjon 9.10.39: Uendelig grense mot uendelig

Vi sier at grensen til f er ∞ når punktet x går mot ∞ (uttrykt som $\lim_{x\to\infty} f(x) = \infty$), hvis det til ethvert reelt tall M finnes et reelt tall S slik at

$$f(x) > M \quad \text{når} \quad x > S. \tag{283}$$

Eksempel 9.10.40. La $n \in \mathbb{Z}^+$. Vis at $\lim_{x\to\infty} x^n = \infty$.

Løsning. Gitt $M > 0$. Dersom $x > \sqrt[n]{M}$, da er $x^n > M$ og vi er ferdig. □

På tilsvarende måte har vi

$$\lim_{x\to-\infty} x^n = \begin{cases} \infty, & \text{hvis } n \text{ er (positivt) partall,} \\ -\infty, & \text{hvis } n \text{ er (positivt) oddetall,} \end{cases} \tag{284}$$

jf. Oppgave 9.73.

Eksempel 9.10.41. La $a \in \mathbb{R}^+$. Vis at

(i) $\lim_{x\to\infty} a^x = \infty$ hvis $a > 1$,

(ii) $\lim_{x\to-\infty} a^x = \infty$ hvis $0 < a < 1$.

Løsning. Gitt $M \in \mathbb{R}$. Vi har, for $x \in \mathbb{R}^+$:

$$\begin{aligned} a^x > |M| &\iff \log_a a^x > \log_a |M| && (x \mapsto \log_a x \text{ str. voks. (Setn. 9.5.14)}) \\ &\iff x > \log_a |M| && ((231)) \end{aligned}$$

Vi setter $S = \log_a |M|$. Vi har da vist at når $x > S$, da er $a^x > M$. Dette viser at $\lim_{x\to\infty} a^x = \infty$ og dermed (i).

Beviset for (ii) er nesten helt likt. □

Det er ekstra viktig å være klar over at ∞ og $-\infty$ *ikke* er tall, slik at vi ikke kan bruke vanlige regneregler på ∞ og $-\infty$. Det er også viktig å merke seg at *grensesetningene* 9.10.23 **ikke** overføres på uendelige grenser. For eksempel er

$$\lim_{x\to 0}\frac{x^2+1}{x^2}=\lim_{x\to 0^+}\left(1+\frac{1}{x^2}\right)=\infty$$

ved Eksempel 9.10.37, siden $1+\frac{1}{x^2}>\frac{1}{x^2}$. Men

$$\lim_{x\to 0}\left(\frac{x^2+1}{x^2}-\frac{1}{x^2}\right)=\lim_{x\to 0}\frac{x^2+1-1}{x^2}=\lim_{x\to 0}\frac{x^2}{x^2}=\lim_{x\to 0}1=1,$$

og ikke "$\infty-\infty=0$"!

Det finnes imidlertid noen regler som gjelder for beregning av grenser som involverer uendelige grenser, men vi overlater disse til Oppgavene 9.75-9.78.

Oppgaver

Oppgaver til §9.1-9.5

OPPGAVE 9.1. Vis at funksjonene er kontinuerlige i de angitte punktene ved å bruke den formelle definisjonen:

(a) $f:\mathbb{R}\to\mathbb{R}$, $f(x)=3x-5$, i punktet 1;
(b) $f:\mathbb{R}\to\mathbb{R}$, $f(x)=2x^2+4$, i punktet 2;
(c) $f:\mathbb{R}\to\mathbb{R}$, $f(x)=x^3$, i punktet 1;
(d) $f:\mathbb{R}\setminus\{0\}\to\mathbb{R}$, $f(x)=\frac{1}{x}$ i punktet -1.
(e) $f:\mathbb{R}\to\mathbb{R}$, $f(x)=\begin{cases} x\sin\left(\frac{1}{x}\right), & x\neq 0,\\ 0, & x=0,\end{cases}$ i punktet 0.

OPPGAVE 9.2. Vis at kvadratrotfunksjonen $f:[0,\infty)\to\mathbb{R}$ gitt ved $f(x)=\sqrt{x}$ er kontinuerlig i alle punkter ved å bruke den formelle definisjonen.

(Hint: for fiksert $a\in[0,\infty)$ er

$$|\sqrt{x}-\sqrt{a}|=\sqrt{|\sqrt{x}-\sqrt{a}|^2}\leq\sqrt{|\sqrt{x}-\sqrt{a}|\,|\sqrt{x}+\sqrt{a}|}=\sqrt{|\sqrt{x}-\sqrt{a}|},$$

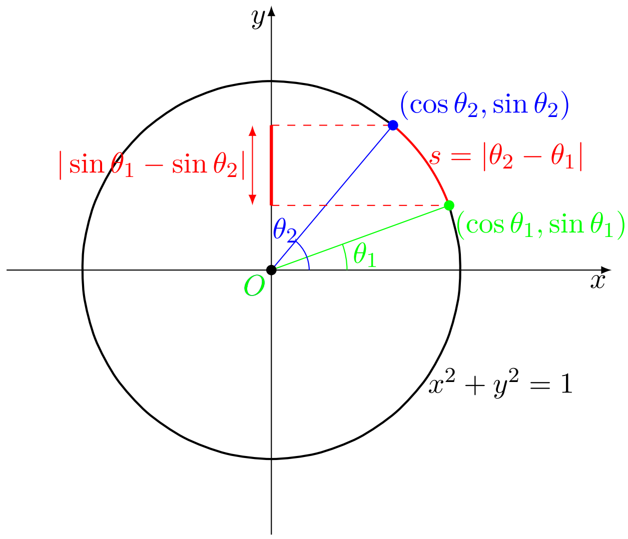
for alle $x\in[0,\infty)$.)

OPPGAVE 9.3. (a) Vis at for alle $\theta_1,\theta_2\in\mathbb{R}$ gjelder

$$|\sin\theta_1-\sin\theta_2|\leq|\theta_1-\theta_2|,$$

ved å

(i) bruke trigonometriske betraktninger og figuren



(ii) bruke vinkelsumidentiteten

$$\sin\theta_1 - \sin\theta_2 = 2\cos\left(\frac{\theta_1+\theta_2}{2}\right)\sin\left(\frac{\theta_1-\theta_2}{2}\right)$$

og at $|\sin x| \leq |x|$ for alle $x \in \mathbb{R}$ ved Merknad 9.2.7.

(b) Bruk resultatet fra (a) og den formelle definisjonen av kontinuitet til å vise at sinusfunksjonen $\sin : \mathbb{R} \to \mathbb{R}$ er kontinuerlig i alle punkter.

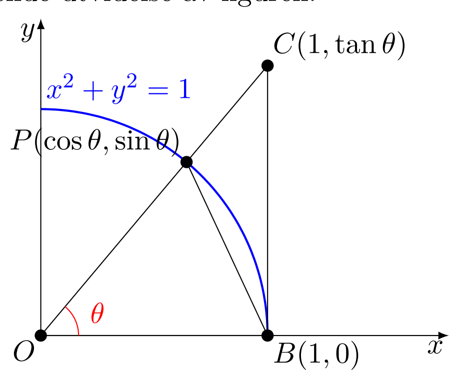
Oppgave 9.4. Vi skal i denne oppgaven vise at funksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved

$$f(x) = \begin{cases} \frac{\sin x}{x}, & x \neq 0, \\ 1, & x = 0 \end{cases}$$

er kontinuerlig i 0, som hevdet i Merknad 9.2.9.

(a) Bruk figuren fra Eksempel 9.2.6 på side 332 til å konkludere at $1 - \cos\theta \leq \theta$ for $\theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.

(b) Betrakt følgende utvidelse av figuren:



Bruk at

Areal trekant $OPB \leq$ Areal sirkelsektor $OPB \leq$ Areal trekant OCB

til å konkludere at

$$|\sin\theta| \le |\theta| \le |\tan\theta|$$

for $\theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ og utled fra dette at

$$\cos\theta \le \frac{\sin\theta}{\theta} \le 1$$

for $\theta \in \left(-\frac{\pi}{2}, 0\right) \cup \left(0, \frac{\pi}{2}\right)$.

(c) Bruk ulikhetene fra (a) og (b) til å vise at f er kontinuerlig i 0.

Oppgave 9.5. En delmengde $A \subset \mathbb{R}$ kalles *åpen* dersom det for hver $a \in A$ finnes en $\delta > 0$ slik at $(a-\delta, a+\delta) \subset A$.

La $f : \mathbb{R} \to \mathbb{R}$ og $a \in \mathbb{R}$.

(a) Vis at f er kontinuerlig i a hvis og bare hvis det for enhver åpen mengde V slik at $f(a) \in V$ finnes en åpen mengde U slik at $a \in U \subset f^{-1}(V)$.
(b) Vis at f er kontinuerlig hvis og bare hvis $f^{-1}(V) \subset \mathbb{R}$ er åpen for alle åpne mengder $V \subset \mathbb{R}$.

Vi kan generalisere dette til funksjoner $f : X \to \mathbb{R}$, der $X \subset \mathbb{R}$ muligens er en *ekte* delmengde (det vil si $X \neq \mathbb{R}$). Vi definerer en delmengde $A \subset X$ til å være *åpen i* X dersom det for hver $a \in A$ finnes en $\delta > 0$ slik at $(a-\delta, a+\delta) \cap X \subset A$.

(c) Vis at $f : X \to \mathbb{R}$ er kontinuerlig i a hvis og bare hvis det for enhver åpen mengde $V \subset \mathbb{R}$ slik at $f(a) \in V$ finnes en åpen delmengde U av X slik at $a \in U \subset f^{-1}(V)$.
(d) Vis at $f : X \to \mathbb{R}$ er kontinuerlig hvis og bare hvis $f^{-1}(V)$ er åpen i X for alle åpne mengder V i $\mathbb{R}$.

Oppgave 9.6. La $X \subset \mathbb{R}$. Et punkt a kalles et *isolert punkt i* X dersom det finnes en δ-omegn $(a-\delta, a+\delta)$ om a slik at $(a-\delta, a+\delta) \cap X = \{a\}$.

Vis at enhver funksjon $f : X \to \mathbb{R}$ med $X \subset \mathbb{R}$ er kontinuerlig i ethvert isolert punkt i X.

Oppgave 9.7. La $f, g : \mathbb{R} \to \mathbb{R}$ være kontinuerlige funksjoner som er like i alle rasjonale punkter (med andre ord har vi $f(x) = g(x)$ for alle $x \in \mathbb{Q}$). Vis at $f(x) = g(x)$ for alle $x \in \mathbb{R}$. Vi sier at *en kontinuerlig funksjon er entydig bestemt av sine verdier i rasjonale punkter.*

Oppgave 9.8. La $f : \mathbb{R} \to \mathbb{R}$ være definert ved

$$f(x) = \begin{cases} x, & x \text{ rasjonal;} \\ 0, & x \text{ irrasjonal.} \end{cases}$$

Avgjør i hvilke punkter f er kontinuerlig.

Oppgave 9.9. Vi sier at et reellt tall r er *dyadisk* dersom $r = \frac{m}{2^n}$. Definér en funksjon $f : \mathbb{R} \to \mathbb{R}$ ved

$$f(x) = \begin{cases} 0, & \text{hvis } x \text{ ikke er dyadisk,} \\ \frac{1}{2^n}, & \text{hvis } x = \frac{m}{2^n} \text{ er dyadisk, med } m \text{ oddetall.} \end{cases}$$

Finn ut i hvilke punkter f er kontinuerlig.

Oppgave 9.10. Bevis Setningene 9.5.1 og 9.5.2 ved hjelp av ϵ–δ-definisjonen.

Oppgave 9.11. Finn definisjonsmengdene til funksjonene nedenunder og avgjør i hvilke punkter funksjonene er kontinuerlige og i hvilke punkter de er diskontinuerlige.

(a) $f(x) = \frac{x+1}{|x+1|}$.

(b) $f(x) = \begin{cases} \sqrt{x}, & x \geq 0, \\ x - 1, & x < 0. \end{cases}$

(c) $f(x) = \begin{cases} x \cos\left(\frac{1}{x^2}\right), & x \neq 0, \\ 0, & x = 0. \end{cases}$

(d) $f(x) = \begin{cases} \cos\left(\frac{1}{x}\right), & x \neq 0, \\ 0, & x = 0. \end{cases}$

Oppgave 9.12. La $a \in \mathbb{R}$ og $V \subset \mathbb{R}$ være et intervall som innholder a. La $f : V \to \mathbb{R}$ være kontinuerlig med $f(a) = 0$ og $g : V \setminus \{a\} \to \mathbb{R}$ være begrenset. Vis at funksjonen $h : V \to \mathbb{R}$ gitt ved

$$h(x) = \begin{cases} f(x)g(x), & \text{hvis } x \in V \setminus \{a\}, \\ 0, & \text{hvis } x = a \end{cases}$$

er kontinuerlig.

Oppgave 9.13. La $f : I \to \mathbb{R}$ være en funksjon, med $I \subset \mathbb{R}$ et intervall. Alt vi vet om f er at f kan tilnærmes ved et polynom med den grad av nøyaktighet vi måtte ønske. Formelt kan vi skrive dette som: For ethvert reellt tall $\epsilon > 0$ er det mulig å skrive f på formen

$$f(x) = P_\epsilon(x) + h_\epsilon(x),$$

der P_ϵ er et polynom og h_ϵ oppfyller $|h_\epsilon(x)| \leq \epsilon$ for alle $x \in I$.

Vis at f er kontinuerlig på I.

Oppgave 9.14. Utfør detaljene i beviset for Setning 9.5.4 i tilfellet b er et endepunkt i I.

Oppgave 9.15. Vi skal i denne oppgaven bevise Følgesetning 9.5.10. Vi har allerede vist i Oppgave 9.3 at sinusfunksjonen $\sin : \mathbb{R} \to \mathbb{R}$ er kontinuerlig og skal bruke dette i denne oppgaven. Vi viser til [**AE**, §P.7 og §3.5] for definisjonene av alle trigonometriske funksjoner og deres inverser.

(a) Bruk identiteten $\cos x = \sin\left(\frac{\pi}{2} - x\right)$ til å konkludere at cos er kontinuerlig. Bruk lignende sammenhenger til å konkludere at alle andre trigonometriske funksjoner er kontinuerlige.

(b) Konkludér at alle inverse trigonometriske funksjoner er kontinuerlige.

Oppgave 9.16. La $\{x_1, x_2, x_3, \ldots\} \subset \mathbb{R}$ være en tellbar delmengde av $\mathbb{R}$ slik at $x_n \to x_0$. La $\{y_n\}$ være en følge av reelle tall slik at $y_n \to y_0$. Sett $X = \{x_0, x_1, x_2, \ldots\}$ og definér funksjonen

$$\begin{aligned} F : X &\longrightarrow \mathbb{R} \\ x_i &\mapsto y_i. \end{aligned}$$

Vis at f er kontinuerlig.

Oppgave 9.17. Gi et eksempel på en funksjon $f : X \to \mathbb{R}$, med $X \subset \mathbb{R}$, og en Cauchy-følge $\{x_n\}$ i X slik at $\{f(x_n)\}$ ikke er en Cauchy-følge. (Med andre ord: en kontinuerlig funksjon bevarer ikke nødvendigvis Cauchy-følger.)

Oppgave 9.18. La $a \in \mathbb{R}^+ \setminus \{1\}$ og $x, y \in \mathbb{R}^+$. Utled følgende egenskaper til logaritmer (jf. Definisjon 9.5.13):

(a) $\log_a a = 1$;
(b) $\log_a 1 = 0$;
(c) $\log_a(xy) = \log_a x + \log_a y$;
(d) $\log_a \left(\frac{x}{y}\right) = \log_a x - \log_a y$;
(e) $\log_a(x^r) = r \log_a x$.

Oppgave 9.19. Vi skal i denne oppgaven vise at

$$\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x \text{ for alle } x \in \mathbb{R}. \tag{285}$$

Vi skal først vise følgende: La $\{x_n\}$ være en følge av reelle tall slik at $\lim_{n\to\infty} x_n = 0$, med $x_n \neq 0$ og $x_n > -1$ for alle $n \in \mathbb{Z}^+$. Da er

$$\lim_{n\to\infty} (1 + x_n)^{\frac{1}{x_n}} = e. \tag{286}$$

Velg $m \in \mathbb{Z}^+$ slik at for alle $n \geq m$ er $|x_n| \leq 1$, det vil si $\frac{1}{|x_n|} \geq 1$.

Velg så $k_n \in \mathbb{Z}^+$ slik at

$$k_n \leq \frac{1}{|x_n|} < k_n + 1, \text{ altså } \frac{1}{k_n + 1} < |x_n| \leq \frac{1}{k_n}.$$

(a) Begrunn at $\lim_{n\to\infty} \left(1 + \frac{1}{k_n}\right)^{k_n} = e$.
(b) Begrunn at $\left(1 + \frac{1}{k_n+1}\right)^{k_n} < (1 + |x_n|)^{\frac{1}{|x_n|}} < \left(1 + \frac{1}{k_n}\right)^{k_n+1}$.
(c) Begrunn at $\left(1 - \frac{1}{k_n}\right)^{k_n+1} < (1 - |x_n|)^{\frac{1}{|x_n|}} < \left(1 - \frac{1}{k_n+1}\right)^{k_n}$ og utled fra dette at $\left(1 + \frac{1}{k_n}\right)^{k_n} < (1 - |x_n|)^{-\frac{1}{|x_n|}} < \left(1 + \frac{1}{k_n-1}\right)^{k_n+1}$.
(d) Utled fra (b) og (c) at $\left(1 + \frac{1}{k_n+1}\right)^{k_n} < (1 + x_n)^{\frac{1}{x_n}} < \left(1 + \frac{1}{k_n-1}\right)^{k_n+1}$.
(e) Utled (286) fra (a) og (d). Hint: bruk at

$$\left(1 + \frac{1}{k_n + 1}\right)^{k_n} = \left(1 + \frac{1}{k_n + 1}\right)^{k_n+1} \left(1 + \frac{1}{k_n + 1}\right)^{-1}.$$

(f) Utled (285). Hint: $(1+\frac{x}{n})^n = \left[(1+\frac{x}{n})^{\frac{n}{x}}\right]^x$.

Oppgave 9.20. La $\{x_n\}$ være en følge av reelle tall slik at $\lim_{n\to\infty} x_n = 0$, med $x_n \neq 0$ og $x_n > -1$ for alle $n \in \mathbb{Z}^+$. Bruk resultatet fra Oppgavene 9.18 og 9.19 til å vise følgende:

(a) $\displaystyle\lim_{n\to\infty} \frac{\log_a(1+x_n)}{x_n} = \log_a e$;

(b) For alle $a \in \mathbb{R}^+$ er $\displaystyle\lim_{n\to\infty} \frac{a^{x_n}-1}{x_n} = \ln a$.

Oppgave 9.21. Vis følgende ulikheter:

(a) $e^x \geq 1 + x$ for alle $x \in \mathbb{R}$. (Hint: ta utgangspunkt i (285) utledet i Oppgave 9.19 og *Bernoullis ulikhet* fra Oppgave 7.13);

(b) $\ln x \leq x - 1$ for alle $x \in \mathbb{R}^+$.

Oppgaver til §9.6–9.7

Oppgave 9.22. Bevis Hjelpesetning 9.6.5. (Hint: Hvis $f(c) > d$, anvend ϵ–δ definisjonen av kontinuitet med $\epsilon = f(c) - d > 0$. Tilsvarende for $f(c) < d$.)

Oppgave 9.23. Vis at ligningen $\sin x = 2x - 1$ har en løsning i intervallet $(0, \frac{\pi}{2})$.

Oppgave 9.24. Vis at ligningen $x = \tan x$ har en løsning i hvert intervall $\left(\left(n-\frac{1}{2}\right)\pi, \left(n+\frac{1}{2}\right)\pi\right)$ for alle $n \in \mathbb{Z}$.

Oppgave 9.25. (a) La $f : [x_1, x_2] \to \mathbb{R}$ være en kontinuerlig funksjon slik at $f(x_1) \geq x_1$ og $f(x_2) \leq x_2$, eller slik at $f(x_1) \leq x_1$ og $f(x_2) \geq x_2$. Vis at f da har et fikspunkt i intervallet $[x_1, x_2]$. (Se Definisjon 9.8.3 for definisjon av et fikspunkt. Merk også at f ikke nødvendigvis er en kontraksjon, så du kan ikke bruke *fikspunktteoremet* 9.8.11.)

(b) Du er på fest med noen studenter fra HF- og SV-fakultetet som lurer på hva du har lært de første månedene på universitetet. I lommen har du et smalt og tynt elastisk bånd. Du legger båndet på bordet og strekker det ved å trekke den ene enden mot høyre og den andre enden mot venstre og forteller at du kan bevise at et eller annet punkt på båndet har forblitt i sin opprinnelige posisjon. Hvorfor? (Hint: hva har dette spørsmålet med (a) å gjøre?)

Oppgave 9.26. En munk starter fra sitt kloster klokken 12:00 og går opp en fjellsti. Han når fjelltoppen klokken 24:00. Der mediterer han og sover til neste morgen. Så starter han nedstigningen presis klokken 12:00, tar nøyaktig samme vei ned, og kommer ned til startpunktet ved klosteret klokken 24:00.

Vis at det finnes minst ett punkt på veien som han når ved samme tidspunkt på dagen både på vei opp og ned. (Vi går ut fra at munken går på en kontinuerlig måte.)

Oppgave 9.27. Gi eksempler på at konklusjonen i Følgesetning 9.6.2 ikke holder hvis vi fjerner betingelsene

(a) f kontinuerlig;
(b) I intervall.

Oppgave 9.28. Vis, gjennom et eksempel, at *skjæringssetningen* (Teorem 9.6.1) ikke gjelder dersom vi fjerner betingelsen om at funksjonen skal være kontinuerlig.

Oppgave 9.29. Bevis Følgesetning 9.6.8.

Oppgave 9.30. Vi skal se på et alternativt bevis for *skjæringssetningen* (Teorem 9.6.1). Begynnelsen på beviset er likt beviset i teksten, men vi vil nå vise at $f(c) = d$ ved å vise ulikhetene $f(c) \geq d$ og $f(c) \leq d$. Dette medfører at $f(c) = d$ og fullfører beviset.

(a) Vi ser først på hva som skjer hvis vi nærmer oss c ovenfra. Betrakt følgen $\{x_n\}$ gitt ved $x_n = c + \frac{1}{n}$. Argumentér at $c < b$ og at $x_n \in [a, b]$ for stor nok n. Vis at $f(x_n) \to f(c)$ og at vi derfor har $f(c) \geq d$.
(b) Vi ser dernest på hva som skjer hvis vi nærmer oss c nedenfra. Vis at det for enhver stor nok $n \in \mathbb{Z}^+$ finnes en $y_n \in S$ slik at $c - \frac{1}{n} \leq y_n \leq c$. Vis at $f(y_n) \to f(c)$ og at vi derfor har $f(c) \leq d$.

Oppgave 9.31. Finnes det en kontinuerlig bijeksjon fra intervallet $(0, 1)$ til intervallet $[0, 1]$?

Oppgave 9.32. Finnes det en kontinuerlig funksjon $\mathbb{R} \to \mathbb{Q}$ som ikke er konstant?

Oppgave 9.33. Vis ved et eksempel at *ekstremalverdisetningen* (Teorem 9.7.2) ikke gjelder i $\mathbb{Q}$.

Oppgave 9.34. La $f : [a, b] \to \mathbb{C}$ være kontinuerlig. Betrakt funksjonen

$$\begin{array}{ccc} [a,b] & \longrightarrow & \mathbb{R} \\ x & \mapsto & |f(x)|. \end{array}$$

Vil denne oppnå maksimum og minimum?

Oppgave 9.35. I denne oppgaven skal vi utlede et litt annerledes bevis for *begrensningsteoremet* 9.7.9 ved å bruke "intervallhalveringsmetoden" som i bevisene for Setningene 7.6.1 og 9.5.12 (og Oppgave 7.45). Vi argumenterer ved selvmotsigelse og antar at f er ubegrenset på $[a, b]$.

(a) Start med å dele opp intervallet i to like store deler: intervallene $[a, \frac{a+b}{2}]$ og $[\frac{a+b}{2}, b]$. Siden f er ubegrenset, må den også være ubegrenset på minst ett av disse delintervallene; kall dette $[a_1, b_1]$ (ved å eventuelt velge ett av dem). Fortsett prosedyren og argumentér at man får en uendelig kjede av intervaller

$$[a, b] \supset [a_1, b_1] \supset [a_2, b_2] \supset \cdots \supset [a_n, b_n] \supset \cdots$$

av lengde $b_n - a_n = \frac{b-a}{2^n}$ hvor f er ubegrenset.

(b) Argumentér at begge følgene $\{a_n\}$ og $\{b_n\}$ konvergerer mot samme punkt i $[a, b]$. Kall dette punktet c. (Hint: ta utgangspunkt i beviset for Setning 7.6.1.)
(c) Argumentér at det for enhver $n \in \mathbb{Z}^+$ finnes et punkt $c_n \in [a, b]$ slik at $|f(c_n)| \geq n$.
(d) Argumentér at også c_n må konvergere mot c.
(e) Utled en selvmotsigelse. (Hint: dette ligner på tilsvarende punkt i beviset for *begrensningsteoremet* 9.7.9 gitt i teksten.)

Oppgave 9.36. I denne oppgaven skal vi utlede et litt annerledes bevis for *ekstremalverdisetningen* 9.7.2 ved å fortsette argumentasjonen fra forrige oppgave. La $M = \sup V(f)$ som i beviset gitt i teksten.

(a) Start igjen med å dele opp intervallet i to like store deler. Argumentér at siden f har supremum M på $[a, b]$, må den ha supremum M på minst ett av de to delintervallene; kall dette $[a_1, b_1]$. Fortsett prosedyren og argumentér at man får en uendelig kjede av intervaller

$$[a, b] \supset [a_1, b_1] \supset [a_2, b_2] \supset \cdots \supset [a_n, b_n] \supset \cdots$$

av lengde $b_n - a_n = \frac{b-a}{2^n}$ hvor f har supremum lik M.
(b) Argumentér at begge følgene $\{a_n\}$ og $\{b_n\}$ konvergerer mot samme punkt i $[a, b]$. Kall dette punktet c.
(c) Vis at $f(c) = M$.

Oppgave 9.37. I denne oppgaven skal vi utlede nok et bevis for *ekstremalverdisetningen* 9.7.2, som bygger på *skjæringssetningen* (Teorem 9.6.1) og stoffet fra §7.7.

La $f : [a, b] \to \mathbb{R}$ være kontinuerlig. Ved Følgesetning 9.6.2 vet vi at verdimengden $J = f([a, b])$ er et intervall. For å vise *ekstremalverdisetningen* er det tilstrekkelig å vise at J er lukket og begrenset. Ved Følgesetning 7.7.8 er det nok å vise at *enhver delfølge i J har en delfølge som konvergerer i J.* Vis dette siste utsagnet.

Oppgaver til §9.8

Oppgave 9.38. Verifisér påstanden om at funksjonen f i Merknad 9.8.12 oppfyller betingelsen (246), men har ingen fikspunkter.

Oppgave 9.39. La f være en funksjon som oppfyller betingelsene i *fikspunktteoremet* 9.8.11. Anta at vi ønsker å approksimere fikspunktet r til f med et ledd i følgen $\{x_n\}$ generert av fikspunktiterasjon med en gitt tålegrense F.

Vis at $|x_n - r| < F$ så lenge

$$n > \frac{F}{\ln K(1-K)|x_1 - x_0|}.$$

OPPGAVE 9.40. Utfør fikspunktiterasjon med et valgt startpunkt på funksjonen i (259) for å finne tilnærmede verdier på de to løsningene til ligningen (249) som ligger i hvert av intervallene $(-3,-2)$ og $(2,3)$. Begrunn at du kan være sikker på at følgen generert ved fikspunktiterasjon konvergerer og drøft nøyaktigheten av tilnærmingene.

OPPGAVE 9.41. Denne oppgaven forutsetter kjennskap til derivasjon og spesielt til *Sekantsetningen/Middelverdisetningen*[7] (jf. [**AE**, Thm. 11 i §2.8] eller [**Li**, Teorem 6.2.3]).

Anta $I \subset \mathbb{R}$ er et intervall og $f : I \to I$ er en deriverbar funksjon slik at $|f'(x)| \leq K$ for en konstant $K < 1$ for alle $x \in I$. Vis at f er en kontraksjon.

OPPGAVE 9.42. (a) Vis at ligningen

$$x^3 + 4x^2 - 10 = 0 \tag{287}$$

har en reell løsning i intervallet $[1,2]$.

(b) Vis at ligningen kan omformes til hver av ligningene

- $x = g_1(x) = x - x^3 - 4x^2 + 10$,
- $x = g_2(x) = \sqrt{\frac{10}{x} - 4x}$,
- $x = g_3(x) = \frac{1}{2}\sqrt{10 - x^3}$,
- $x = g_4(x) = \sqrt{\frac{10}{4+x}}$,
- $x = g_5(x) = x - \frac{x^3+4x^2-10}{3x^2+8x}$.

(c) Utfør fikspunktiterasjon med startpunkt $x_0 = 1.5$ på hver av funksjonene $g_1, \ldots, g_5$ ovenfor. (Bruk en datamaskin.)

(d) Finn ut i hvilke tilfeller du kan være sikker på at følgen generert ved fikspunktiterasjon konvergerer og divergerer. (Bruk gjerne resultatet fra Oppgave 9.41.)

(e) Drøft nøyaktigheten i forskjellige tilnærminger på løsningen av ligningen (287) du oppnår i de tilfellene vi har konvergens.

OPPGAVE 9.43. I Eksemplene 9.6.7 og 9.6.9 viste vi at ligningen $x = \cos x$ har nøyaktig én løsning, og at denne ligger i intervallet $[0,1]$.

(a) Vis at funksjonen $[0,1] \to [0,1]$ definert ved $x \mapsto \cos x$ er en kontraksjon. (Hint: Bruk vinkelsumidentiteten

$$\cos\theta_1 - \cos\theta_2 = -2\sin\left(\frac{\theta_1+\theta_2}{2}\right)\sin\left(\frac{\theta_1-\theta_2}{2}\right)$$

og at $|\sin x| \leq |x|$ for alle $x \in \mathbb{R}$ ved Merknad 9.2.7.)

(b) Finn tilnærmede verdier for løsningen på ligningen $x = \cos x$ ved hjelp av fikspunktiterasjon og drøft nøyaktigheten.

[7] *Mean Value Theorem* på engelsk.

Oppgaver til §9.9

Oppgave 9.44. Gå gjennom det du gjorde i Oppgavene 9.2 og 9.3 og vis at kvadratrotfunksjonen $x \mapsto \sqrt{x}$ og sinusfunksjonen $x \mapsto \sin x$ er uniformt kontinuerlige på sine definisjonsmengder, som hevdet i teksten.

Oppgave 9.45. La $f : A \to \mathbb{R}$ med $A \subset \mathbb{R}$. Vis følgende kriterium: Hvis det for hver $h > 0$ er slik at $|f(x+h) - f(x)|$ er ubegrenset på A, så er f ikke uniformt kontinuerlig på A.

Oppgave 9.46. Bevis Setning 9.9.10.

Oppgave 9.47. Bevis Setning 9.9.12.

Oppgave 9.48. Avgjør om $\cos : \mathbb{R} \to \mathbb{R}$ er uniformt kontinuerlig.

Oppgave 9.49. Avgjør om $\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}$ og $\tan^{-1} : \mathbb{R} \to \mathbb{R}$ er uniformt kontinuerlige eller ikke. (Hint til den ene delen: bruk identiteten $\tan^{-1} x - \tan^{-1} y = \tan^{-1}\left(\frac{x-y}{1+xy}\right)$.)

Oppgave 9.50. Avgjør om eksponensialfunksjonen $x \mapsto e^x$ er uniformt kontinuerlig på $(-\infty, 0)$, $(0, \infty)$ og hele $\mathbb{R}$, henholdsvis.

Oppgave 9.51. Vis at funksjonen $f : (0, \infty) \to \mathbb{R}$ gitt ved $f(x) = \sin\left(\frac{1}{x}\right)$ er uniformt kontinuerlig på ethvert intervall $[a, \infty)$, for $a > 0$.

Oppgave 9.52. Benytt følge-kriteriet for uniform kontinuitet (Setning 9.9.28) til å vise at $x \mapsto x^2$ ikke er uniformt kontinuerlig på $[0, \infty)$.

Oppgave 9.53. Avgjør om funksjonen $x \mapsto x \ln x$ er uniformt kontinuerlig på $(0, 1)$, $(1, \infty)$ og hele $(0, \infty)$, henholdsvis.

Oppgave 9.54. Avgjør om funksjonen $f : \mathbb{R} \to \mathbb{R}$ gitt ved $f(x) = x \cos x$ er uniformt kontinuerlig.
(Hint: betrakt følgene $x_n = \frac{\pi}{2} + 2n\pi$ og $y_n = x_n + \frac{1}{n}$.)

Oppgave 9.55. Avgjør om $f : \mathbb{R} \to \mathbb{R}$ gitt ved $f(x) = x^2 \tan^{-1} x$ er uniformt kontinuerlig.

Oppgave 9.56. Bevis Følgesetning 9.9.16.

Oppgave 9.57. Fleip eller fakta: dersom $f : (a, b) \to \mathbb{R}$ er kontinuerlig i $c \in (a, b)$ og uniformt kontinuerlig på (a, c) og (c, b), da er f uniformt kontinuerlig på (a, b).

Oppgave 9.58. Vi vil gi et alternativt bevis for *Heine-Cantor teoremet* 9.9.13 som bygger på stoffet i §7.7.

Vi argumenterer ved selvmotsigelse og antar at $f : [a, b] \to \mathbb{R}$ er kontinuerlig, men *ikke* uniformt kontinuerlig.

(a) Vis at det finnes en $\epsilon \in \mathbb{R}^+$ slik at vi for alle $n \in \mathbb{Z}^+$ kan finne $x_n, y_n \in [a, b]$ som oppfyller

$$|x_n - y_n| < \frac{1}{n} \text{ og } |f(x_n) - f(y_n)| \geq \epsilon.$$

(b) Begrunn at $\{x_n\}$ har en delfølge $\{x_{n_k}\}$ som konvergerer mot en $c \in [a, b]$ og at følgen $\{y_{n_k}\}$ også konvergerer mot c.

(c) Begrunn at $\lim_{n\to\infty} f(x_{n_k}) = \lim_{n\to\infty} f(y_{n_k}) = f(c)$.

(d) Begrunn at resultatene i (a) og (c) er motstridende.

Oppgave 9.59. Bevis Setning 9.9.20.

Oppgave 9.60. Denne oppgaven forutsetter kjennskap til derivasjon og spesielt til *Sekantsetningen/Middelverdisetningen*[8] (jf. [**AE**, Thm. 11 i §2.8] eller [**Li**, Teorem 6.2.3]).

La $I \subset \mathbb{R}$ være et intervall. Vis at en deriverbar funksjon $f : I \to \mathbb{R}$ med begrenset derivert er uniformt kontinuerlig.

Oppgave 9.61. Bevis Setning 9.9.28. (Hint: Oppgave 9.58(a) kan være til hjelp.)

Oppgave 9.62. Vis at dersom $f : X \to \mathbb{R}$, med $X \subset \mathbb{R}$, er uniformt kontinuerlig og $\{x_n\}$ er en Cauchy-følge, da er $\{f(x_n)\}$ fremdeles en Cauchy-følge. Vi sier at *uniformt kontinuerlige funksjoner bevarer Cauchy-følger*, noe som ikke nødvendigvis er sant for kontinuerlige funksjoner, jf. Oppgave 9.17.

Oppgaver til §9.10

Oppgave 9.63. Gi bevis for Setningene 9.10.23–9.10.26 ved hjelp av ϵ-δ-definisjonen for grenser.

Oppgave 9.64. Bevis Setning 9.10.2.

Oppgave 9.65. Vis at hvis $a \in \mathbb{R}$ er et opphopningspunkt/grensepunkt til $X \subset \mathbb{R}$, så inneholder enhver omegn om a uendelig mange punkter i X.

Oppgave 9.66. Vis at enhver uendelig og begrenset delmengde av $\mathbb{R}$ inneholder minst ett opphopningspunkt.

Oppgave 9.67. En delmengde $A \subset \mathbb{R}$ kalles *lukket* dersom følgende holder: for enhver følge $\{x_n\}$ i A slik at x_n konvergerer mot $x \in \mathbb{R}$, da er $x \in A$ (jf. Oppgave 7.28).

Vis at A er lukket hvis og bare hvis A inneholder alle sine opphopningspunkter.

Oppgave 9.68. Bevis Setning 9.10.36 (og, om du gidder, også Setningene 9.10.14 og 9.10.19). Hint: Mim etter beviset for Setning 9.4.1.

[8] *Mean Value Theorem* på engelsk.

Oppgave 9.69. Finn følgende grenser:

(a) $\lim_{x\to 0}(1+x)^{\frac{1}{x}}$;

(b) $\lim_{x\to 0}\dfrac{\log_a(1+x)}{x}$;

(c) $\lim_{x\to 0}\dfrac{a^x-1}{x}$.

(Hint: bruk (286) i Oppgave 9.19 og Oppgave 9.20.)

Oppgave 9.70. Vis at dersom $\lim_{x\to a} f(x)$ eksisterer, da finnes en omegn $(x-\delta, x+\delta)$ om a slik at $f|_{D(f)\cap(x-\delta,x+\delta)}$ er begrenset.

Oppgave 9.71. Vis at dersom $f:(a,b)\to\mathbb{R}$ er voksende og nedad (henholdsvis oppad) begrenset, da eksisterer $\lim_{x\to a^+} f(x)$ (hhv. $\lim_{x\to b^-} f(x)$).

Oppgave 9.72. Vis (282).

Oppgave 9.73. Vis (284).

Oppgave 9.74. La $a\in\mathbb{R}^+\setminus\{0\}$. Finn grensene $\lim_{x\to 0^+}\log_a x$ og $\lim_{x\to\infty}\log_a x$. (Hint: del opp i tilfellene $a>1$ og $a<1$.)

Oppgave 9.75. Anta at $\lim_{x\to a} f(x)=\infty$, $\lim_{x\to a} g(x)=L$ og $k\in\mathbb{R}$. Vis følgende grenser:

(a) $\lim_{x\to a}\Big(f(x)\pm g(x)\Big)=\infty$

(b) $\lim_{x\to a} f(x)g(x)=\begin{cases}\infty, & \text{hvis } L>0\\ -\infty, & \text{hvis } L<0\end{cases}$

(c) $\lim_{x\to a}\dfrac{f(x)}{g(x)}=\begin{cases}\infty, & \text{hvis } L>0\\ -\infty, & \text{hvis } L<0\end{cases}$

(d) $\lim_{x\to a} kf(x)=\begin{cases}\infty, & \text{hvis } k>0\\ -\infty, & \text{hvis } k<0\end{cases}$

(e) $\lim_{x\to a}\dfrac{g(x)}{f(x)}=0$.

(Hvis $\lim_{x\to a} f(x)=-\infty$ og $\lim_{x\to a} g(x)=L$, da gjelder de samme konklusjonene med motsatt fortegn. Resultatene i oppgaven gjelder også for ensidige grenser og grenser i uendelig, med endringene angitt i Bemerkning 9.10.27.)

Oppgave 9.76. (a) Vis at dersom $\lim_{x\to a} f(x)=0$ og $f(x)>0$ (henholdsvis <0) i en punktert omegn om a, da er

$$\lim_{x\to a}\frac{1}{f(x)}=\infty \ \ (\text{hhv. } -\infty).$$

(Resultatet gjelder også for ensidige grenser og grenser i uendelig, med endringene angitt i Bemerkning 9.10.27.)

(b) Vis at betingelsen om at $f(x) > 0$ (hhv. < 0) i en punktert omegn om a er nødvendig ved å betrakte funksjonen

$$\lim_{x \to 0} \frac{1}{x^2 \sin\left(\frac{1}{x}\right)}.$$

Oppgave 9.77. La

$$P(x) = a_n x^n + \cdots + a_1 x + a_0, \quad a_0, \ldots, a_n \in \mathbb{R}, \; a_n \neq 0,$$

være et polynom.

(a) Vis at

$$\lim_{x \to \infty} P(x) = \begin{cases} \infty & \text{hvis } a_n > 0 \\ -\infty & \text{hvis } a_n < 0. \end{cases}$$

(b) Vis at

$$\lim_{x \to -\infty} P(x) = \begin{cases} \infty & \text{hvis } a_n > 0 \text{ og } n \text{ er partall,} \\ -\infty & \text{hvis } a_n < 0 \text{ og } n \text{ er partall,} \\ -\infty & \text{hvis } a_n > 0 \text{ og } n \text{ er oddetall,} \\ \infty & \text{hvis } a_n < 0 \text{ og } n \text{ er oddetall.} \end{cases}$$

Oppgave 9.78. La

$$R(x) = \frac{a_n x^n + \cdots + a_1 x + a_0}{b_m x^m + \cdots + b_1 x + b_0}$$

være en rasjonal funksjon (med $a_n, b_m \neq 0$). Finn $\lim_{x \to \pm\infty} R(x)$ og $\lim_{x \to \pm -\infty} R(x)$ ved å dele opp studiet i tilfellene $n < m$, $n = m$ og $n > m$.

Oppgave 9.79. Diskutér følgende påstand: En kontinuerlig funksjon $f : (a, b) \to \mathbb{R}$ er begrenset hvis og bare hvis vi har $\lim_{x \to a^+} f(x) \neq \pm\infty$ og $\lim_{x \to b-} f(x) \neq \pm\infty$.

Oppgave 9.80. La $f : [a, \infty) \to \mathbb{R}$ være kontinuerlig slik at $\lim_{x \to \infty} f(x)$ eksisterer. Vis at f er uniformt kontinuerlig.

KAPITTEL 10

Euklidske rom, metriske rom og topologiske rom

Matematikere har alltid vært opptatt av å sile ut situasjoner og mønstre som har noe til felles for å forsøke å bygge opp teorier som omfatter mange av disse mønstrene på en gang. På denne måten evner man å "se skogen for bare trær" og unngår å måtte behandle mange spesialtilfeller separat. Et godt eksempel, som vi allerede har sett, er begrepene *ringer* og *kropper* i algebra, som generaliserer heltallene på den ene siden, og de rasjonale, reelle og komplekse tallene på den andre siden, og har gjort matematikere i stand til å utvikle teori som kan anvendes på mange forskjellige bruksområder.

Et annet eksempel er teorien for følger og kontinuitet av funksjoner som vi har utviklet så langt. Vi har allerede sett at $\mathbb{R}$ og $\mathbb{C}$ oppfører seg svært likt og at mye av stoffet kun avhenger av at vi har et veldefinert avstandsbegrep på disse mengdene. Spesielt hendig er det at avstandsbegrepet oppfyller *trekantulikheten*, som vi har brukt gjentatte ganger i mange bevis. Mye av det vi har snakket om så langt fungerer like godt på vilkårlige mengder hvor vi har et avstandsbegrep. Dette er bakgrunnen for definisjonen av *metriske rom* (Definisjon 7.2.4), som er mengder med et slikt avstandsbegrep, som oppfyller visse aksiomer, deriblant trekantulikheten. Her kan vi definere konvergens av følger og kontinuitet og de aller fleste resultatene overføres.

Viktige eksempler på metriske rom er *euklidske rom*, som er mengden $\mathbb{R}^n$ av alle ordnede n-tupler $(x_1, \ldots, x_n)$ av reelle tall, utstyrt med det euklidske avstandsbegrepet. Til disse rommene overføres så å si alt det vi har lært om funksjoner og følger i $\mathbb{R}$ og $\mathbb{C}$. Når dere i kursene *MAT112–Grunnkurs i Matematikk II* og *MAT212–Funksjoner av flere variable* studerer funksjoner av flere variable, er det nettopp funksjoner mellom euklidske rom dere studerer.

Euklidske rom går tilbake til Euklids *Elementer* fra ca. 300 f.Kr. og ble introdusert som en abstraksjon av det fysiske universet. I 1637 innførte René Descartes (1596–1650) kartesiske koordinater, som muliggjorde vekselvirkningen mellom algebra og ligningsløsning på den ene siden og geometri på den andre siden som vi er vant til i dag. Euklidske rom $\mathbb{R}^n$ med $n > 3$ ble imidlertid ikke studert før på 1800-tallet; de viktigste bidragsyterne i denne retningen var den sveitsiske matematikeren Ludwig Schläfli (1814–1895) og den tyske matematikeren Bernhard Riemann (1826–1866).

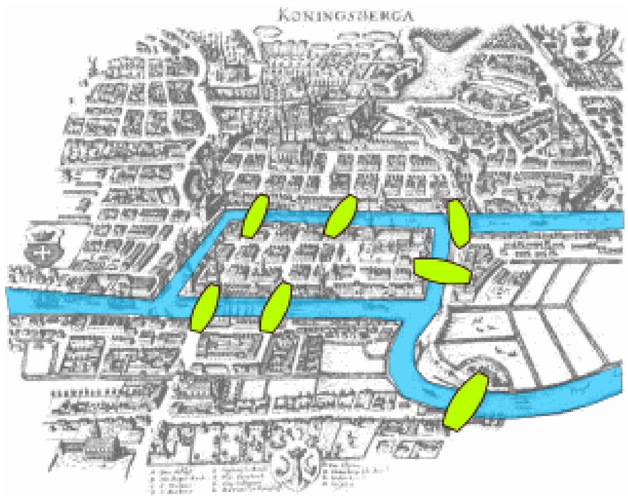
Definisjonen på metriske rom ble introdusert i 1906 av den franske matematikeren Maurice Fréchet (1878–1973) i sin doktoravhandling *Sur quelques points du calcul fonctionnel*. Denne avhandlingen er et kunststykke i matematisk abstraksjon hvor teoremer bevises ut i fra gjennomtenkte aksiomer og anvendes så i en rekke spesialtilfeller. Selve navnet "metrisk rom" ble imidlertid introdusert av den tyske matematikeren Felix Hausdorff (1868–1942) i sin innflytelsesrike bok *Grundzüge der Mengenlehre* om mengdeteori fra 1914. Både Fréchets og Hausdorffs arbeider bygde på mye av mengdelæren utviklet av Georg Cantor (1845–1918) mot slutten av 1800-tallet. Hausdorff var også en av grunnleggerne av en ytterligere generalisering av metriske rom, nemlig *topologiske rom*, et begrep også introdusert i boken fra 1914. Her blir avstandsbegrepet erstattet med en aksiomatisk definisjon av omegner (og åpne mengder) og man løsriver seg fra alle egenskaper som har med avstander å gjøre og er mer opptatt av objekters "form". Dessverre begikk Hausdorff selvmord i 1942 sammen med sin kone og hennes søster da de som jøder ble beordret til konsentrasjonsleiren Endenich.

Maurice Fréchet. Kilde: Wikipedia (public domain)

Felix Hausdorff. Kilde: Wikipedia (public domain)

For øvrig har både Fréchet og Hausdorff fått oppkalt en asteroide etter seg, nemlig *21537 Fréchet* og *24947 Hausdorff*.

Topologiske idéer har imidlertid en tidligere opprinnelse. Starten regnes ofte for å være det velkjente problemet om de "Syv broene i Königsberg". Byen Königsberg i Preussen (nåværende Kaliningrad i Russland) var på 1700-tallet delt i fire deler: den nordlige og sørlige siden av elven Pregel og to øyer midt i elven. Disse var forbundet med 7 broer, som vist i figuren:



De syv broene i Königsberg på Eulers tid. Kilde: Wikipedia (Creative Commons Attribution-Share Alike 3.0 Unported license).

Flere av byens innbyggere spurte seg om det var mulig å gå gjennom byen på en slik måte at man passerte hver bro bare én gang og endte turen tilbake i utgangspunktet. Ingen lyktes å finne en slik tur, men ingen klarte å vise at det var umulig.

Den sveitsiske matematikeren Leonhard Euler (1707–1783) fikk høre om problemet, og løste det i artikkelen *Solutio problematis ad geometriam situs pertinentis* i 1735. Han viste at det ikke fantes noen løsning på problemet, det vil si, at en slik tur ikke er mulig[1]. Euler innså at problemet ikke avhenger av broenes eksakte posisjon, men av antallet broer og landområder og hvordan de er forbundet. Dette innebar et skifte av fokus fra de *metriske* egenskapene til geometriske objekter i den greske (euklidske) geometrien til mer kvalitative egenskaper.

Ytterligere bidrag kom fra Augustin–Louis Cauchy (1789–1857), Ludwig Schläfli (1814–1895), Bernhard Riemann (1826–1866) og den italienske matematikeren Enrico Betti (1823–1892). Selve ordet "topologi" kommer fra de greske ordene "topos", som betyr "sted", og "logi", som betyr "lære", og ble først brukt av den tyske matematikeren Johann Benedict Listing (1808–1882) i sin artikkel *Vorstudien zur Topologie* fra 1847. Arbeidene til de fem sistnevnte matematikerne ble utvidet og sterkt forbedret

[1] I dag er imidlertid en slik tur mulig, siden to av broene ble bombet i stykker under andre verdenskrig.

av den franske matematikeren Henri Poincaré (1854–1912) i flere grunnleggende artikler i perioden 1895–1904, hvor han blant annet fremla den berømte Poincaréformodningen som vi snakket om i §1.4.

Vi skal i dette kapitlet først introdusere metriske og euklidske rom og deres egenskaper i §10.1, for så å se hvordan begreper vi har sett før, som konvergens av følger og kontinuitet, samt *skjæringssetningen* og *ekstremalverdisetningen*, overføres til disse rommene (§10.2–10.6). Ytterligere eksempler på metriske rom foruten euklidske rom gis i §10.7–10.9. Vi avslutter med en liten seksjon om topologiske rom (§10.10).

10.1. Grunnleggende egenskaper til metriske og euklidske rom

Vi har allerede bemerket at det sentrale i definisjonen av konvergens av følger og kontinuitet (samt grensebegrepet for funksjoner) er at vi har et avstandsbegrep i $\mathbb{R}$ og $\mathbb{C}$. Vi vil nå se at veldig mye av det vi har sagt tidligere kan generaliseres til mengder som har et slikt avstandsbegrep.

Metriske rom. Vi starter med å repetere Definisjon 7.2.4:

Definisjon 10.1.1: Metrikk og metrisk rom

La X være en mengde. En *metrikk*[a] på X er en funksjon $d : X \times X \to \mathbb{R}$ som oppfyller aksiomene

(M1) $\begin{cases} d(x,y) > 0 & \text{hvis } x \neq y \\ d(x,y) = 0, & \text{hvis } x = y \end{cases}$ (Positivitet)

(M2) $d(x,y) = d(y,x)$ (Symmetri)

(M3) $d(x,z) \leq d(x,y) + d(y,z)$ (Trekantulikhet)

En mengde X sammen med en metrikk d på X kalles et *metrisk rom*[b].

[a] *Metric* på engelsk.
[b] *Metric space* på engelsk.

Elementene i et metrisk rom kalles som oftest for *punkter*.

Merknad 10.1.2. Selv om bokstaven d (for "distance") ofte brukes for en metrikk, er også den greske bokstaven ρ ("ro") vanlig.

Det er viktig å merke seg at et metrisk rom ikke bare er en mengde X, men en mengde sammen med en metrikk d. Samme mengde kan ha flere forskjellige metrikker (jf. §10.7 for eksempler på forskjellige metrikker på $\mathbb{R}^n$). Derfor betegner vi vanligvis et metrisk rom som et par (X, d), der X er en mengde og d er en metrikk på X. Hvis metrikken er underforstått, sier vi imidlertid også kun at "X er et metrisk rom".

Begrepet *omegn* i $\mathbb{R}$ vi har sett tidligere har følgende opplagte generalisering til alle metriske rom:

Definisjon 10.1.3: Omegn

La (X, d) være et metrisk rom og $p \in X$. La $\delta \in \mathbb{R}^+$.

Mengden

$$\{x \in X \mid d(x, p) < \delta\},$$

ofte betegnet med $N_\delta(p)$ eller $B_\delta(p)$, kalles en *omegn om x*, eller mer presist en *δ-omegn om p* eller *omegn om p av radius δ*.

Mengden

$$\{x \in X \mid 0 < d(x, p) < \delta\} = N_\delta(p) \setminus \{p\},$$

ofte betegnet med $N_\delta(p)^*$ eller $B_\delta(p)^*$ kalles en *punktert omegn om p*, eller mer presist en *punktert δ-omegn om p* eller *punktert omegn om p av radius δ*.

En δ-omegn om p er altså ikke annet enn alle punkter i X i avstand mindre enn δ fra p. Noen ganger brukes begrepet *(åpen) ball* for en *omegn*.

Også egenskapen "begrenset" har sin generalisering til vilkårlige metriske rom:

Definisjon 10.1.4: Begrenset delmengde

La X være et metrisk rom med metrikk d. En delmengde $A \subset X$ er *begrenset* dersom en av følgende ekvivalente betingelser er oppfylt:

(i) Det finnes et punkt $x \in X$ og en $M \in \mathbb{R}$ slik at $d(x, a) \leq M$ for alle $a \in A$;

(ii) Det finnes en $M \in \mathbb{R}$ slik at $d(a, b) \leq M$ for alle $a, b \in A$.

Det overlates til Oppgave 10.1 å vise at de to betingelsene (i) og (ii) virkelig er ekvivalente.

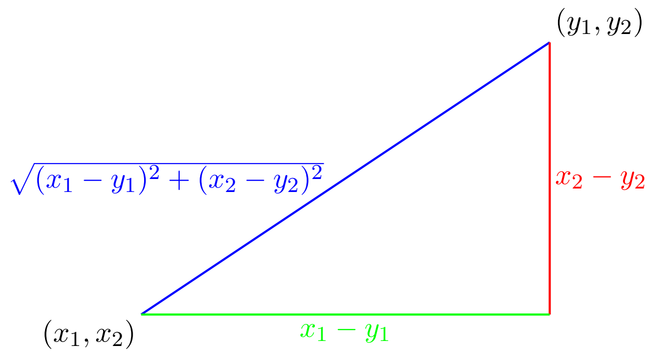
Vi har sett tidligere at både $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}$ og $\mathbb{N}$ alle er metriske rom, med metrikk definert ved $d(x, y) = |x - y|$. Det finnes mange andre eksempler, men vi vil i første omgang, også for å lette lesingen, konsentrere oss om *den euklidske metrikken* på $\mathbb{R}^n$. Vi skal se på mange flere eksempler på metriske rom i §10.7–10.9.

Den euklidske metrikken. Betrakt mengden $\mathbb{R}^2$, eller "planet", som er mengden av alle ordnede tallpar (x_1, x_2) med $x_1, x_2 \in \mathbb{R}$. Det vi "i dagliglivet" mener med avstanden mellom et punkt (eller vektor) i planet $\mathbf{x} = (x_1, x_2)$ og origo er lengden på det rette linjestykket mellom punktene, som er $\sqrt{x_1^2 + x_2^2}$, og betegnes gjerne med $|\mathbf{x}|$. Vi kan også definere sum og differanse mellom to punkter i planet $\mathbf{x} = (x_1, x_2)$ og $\mathbf{y} = (y_1, y_2)$ som

$$\mathbf{x} \pm \mathbf{y} = (x_1 \pm y_1, x_2 \pm y_2).$$

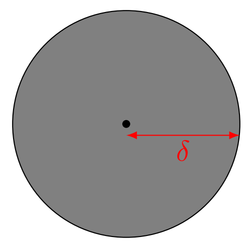
Avstanden mellom $\mathbf{x} = (x_1, x_2)$ og $\mathbf{y} = (y_1, y_2)$ er gitt ved lengden på det rette linjestykket mellom punktene, som er

$$|\mathbf{x} - \mathbf{y}| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}.$$



Avstanden mellom to punkter i planet

En δ-omegn om et punkt består av det indre av en sirkel av radius δ om punktet.



En δ-omegn i den euklidske metrikken i $\mathbb{R}^2$

Betrakt mengden $\mathbb{R}^3$, eller "rommet", som er mengden av alle ordnede talltripler (x_1, x_2, x_3) med $x_1, x_2, x_3 \in \mathbb{R}$. Det vi "i dagliglivet" mener med avstanden mellom et punkt i rommet $\mathbf{x} = (x_1, x_2, x_3)$ og origo er lengden på det rette linjestykket mellom punktene, som er $\sqrt{x_1^2 + x_2^2 + x_3^2}$, og betegnes igjen med $|\mathbf{x}|$. Igjen kan vi definere sum og differanse mellom to punkter i planet $\mathbf{x} = (x_1, x_2, x_3)$ og $\mathbf{y} = (y_1, y_2, y_3)$ som

$$\mathbf{x} \pm \mathbf{y} = (x_1 \pm y_1, x_2 \pm y_2, x_3 \pm y_3).$$

Avstanden mellom to punkter i rommet $\mathbf{x} = (x_1, x_2, x_3)$ og $\mathbf{y} = (y_1, y_2, y_3)$ er gitt ved lengden på det rette linjestykket mellom punktene, som er

$$|\mathbf{x} - \mathbf{y}| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}.$$

En δ-omegn om et punkt består av det indre av en ball av radius δ om punktet.

Slik kan vi fortsette: La $n \in \mathbb{Z}^+$ og betrakt mengden $\mathbb{R}^n$, som er mengden av alle ordnede n-tupler $(x_1, \ldots, x_n)$ med $x_1, \ldots, x_n \in \mathbb{R}$. Slike ordnede tupler kalles *punkter* eller *vektorer*. Vi definerer *normen*[2] til et punkt $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ til å være

$$|\mathbf{x}| = \sqrt{x_1^2 + \cdots + x_n^2},$$

[2] *Norm* på engelsk.

som er avstanden mellom $\mathbf{x}$ og origo $(0, \ldots, 0)$. For $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ og $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$, kan vi definere

$$\mathbf{x} \pm \mathbf{y} = (x_1 \pm y_1, \ldots, x_n \pm y_n)$$

og

$$d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}. \tag{288}$$

Da definerer d en metrikk på $\mathbb{R}^n$ (vises i Oppgave 10.5), kalt *den euklidske metrikken*. Mengden $\mathbb{R}^n$ utstyrt med denne metrikken kalles *euklidsk rom*[3]. Med mindre ikke noe annet er oppgitt, er det denne metrikken vi vil bruke i $\mathbb{R}^n$.

En δ-omegn om et punkt består av det indre av en "n-dimensjonal ball" av radius δ om punktet. Mer presist, for et punkt $\mathbf{a} = (a_1, \ldots, a_n)$ er

$$N_\delta(\mathbf{a}) = \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n \mid \sqrt{(x_1 - a_1)^2 + \cdots + (x_n - a_n)^2} < \delta \right\}.$$

Noen ganger kalles $N_\delta(\mathbf{a})$ en *åpen ball*[4] om $\mathbf{a}$, mens mengden

$$\left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n \mid \sqrt{(x_1 - a_1)^2 + \cdots + (x_n - a_n)^2} \leq \delta \right\}$$

kalles en *lukket ball*[5] om $\mathbf{a}$.

EKSEMPEL 10.1.5. Det er lett å se at dersom X er et metrisk rom med metrikk d_X, og Y er en annen mengde slik at det finnes en bijeksjon $F : X \to Y$, så får vi automatisk indusert en metrikk d_Y på Y ved

$$d_Y(y_1, y_2) = d_X(F^{-1}(y_1), F^{-1}(y_2)).$$

Betrakt $\mathbb{C}^n$, som er mengden av alle ordnede n-tupler $(z_1, \ldots, z_n)$, der $z_1, \ldots, z_n \in \mathbb{C}$. Vi kan skrive $z_j = x_j + iy_j$ for hver $j \in \{1, \ldots, n\}$. Vi har en bijeksjon

$$\begin{array}{ccc} \mathbb{C}^n & \longrightarrow & \mathbb{R}^{2n} \\ (z_1, \ldots, z_n) & \mapsto & (x_1, y_1, x_2, y_2, \ldots, x_n, y_n), \end{array}$$

og ved det vi nettopp sa, induserer den euklidske metrikken på $\mathbb{R}^n$ fra forrige eksempel en metrikk Δ på $\mathbb{C}^n$ som følger: for $\mathbf{z} = (x_1 + iy_1, \ldots, x_n + iy_n) \in \mathbb{C}^n$ og $\mathbf{w} = (a_1 + ib_1, \ldots, a_n + ib_n) \in \mathbb{C}^n$, er

$$\Delta(\mathbf{z}, \mathbf{w}) = \sqrt{\sum_{j=1}^{n} [(x_j - a_j)^2 + (y_j - b_j)^2]}.$$

Denne metrikken kalles igjen *den euklidske metrikken*.

[3] *Euclidean space* på engelsk.
[4] *Open ball* på engelsk.
[5] *Closed ball* på engelsk.

Spesialtilfellet $n = 1$ gir, med $z = x + iy$ og $w = a + ib$,

$$\Delta(z, w) = \sqrt{(x-a)^2 + (y-b)^2} = |z - w|,$$

som er den "vanlige" metrikken vi allerede har sett på $\mathbb{C}$ (Merknad 8.2.2).

EKSEMPEL 10.1.6. Vis at en delmengde $A \subset \mathbb{R}^n$ er begrenset hvis og bare hvis det finnes en $M \in \mathbb{R}$ slik at A er inneholdt i den lukkede ballen

$$\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_1^2 + \cdots + x_n^2 \leq M\}. \tag{289}$$

LØSNING. En mengde inneholdt i en lukket ball som i (289) oppfyller betingelsen (i) i Definisjon 10.1.4 (med det "faste punktet" x lik origo) og er dermed begrenset. Motsatt vei, vil en delmengde $A \subset \mathbb{R}^n$ som oppfyller betingelsen (ii) i Definisjon 10.1.4 spesielt oppfylle betingelsen med b lik origo, som viser at A er inneholdt i den lukkede ballen (289). □

I motsetning til $\mathbb{R}$ er ikke $\mathbb{R}^n$ en kropp for $n \geq 2$. Vi har sett at vi kan definere operasjoner $+$ og $-$ på $\mathbb{R}^n$, og det er lett å sjekke at disse oppfyller (A1)-(A4) i kroppsaksiomene (Definisjon 3.6.1), med nøytralt element $(0, \ldots, 0)$. Vi har imidlertid ingen multiplikasjon definert på $\mathbb{R}^n$. Vi kan likevel definere et par operasjoner som "ligner" på multiplikasjon, som dere sikkert husker fra skolen. Vi oppsummerer i følgende definisjon:

Definisjon 10.1.7: Operasjoner på $\mathbb{R}^n$

La $\mathbf{x} = (x_1, \ldots, x_n)$ og $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$, og la $c \in \mathbb{R}$. Vi definerer:

- $\mathbf{x} \pm \mathbf{y} = (x_1 \pm y_1, \ldots, x_n \pm y_n)$ (vektoraddisjon);
- $c\mathbf{x} = (cx_1, \ldots, cx_n)$ (skalarmultiplikasjon);
- $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i y_i$ (indreprodukt).

Merk at $\mathbf{x} \pm \mathbf{y} \in \mathbb{R}^n$, $c\mathbf{x} \in \mathbb{R}^n$, mens $\mathbf{x} \cdot \mathbf{y} \in \mathbb{R}$.

MERKNAD 10.1.8. Operasjonene *vektoraddisjon* og *skalarmultiplikasjon* gjør euklidske rom til $\mathbb{R}$-*vektorrom*, jf. Eksempel 4.5.6. Indreproduktet gjør euklidske rom til såkalte *indreproduktrom*, jf. Oppgave 10.6.

Det euklidske rommet $\mathbb{R}^n$ er heller ikke en ordnet mengde i utgangspunktet, selv om det finnes flere muligheter å definere ordensrelasjoner på $\mathbb{R}^n$ (som for eksempel *leksikografisk orden* definert i Eksempel 3.6.11). Disse er imidlertid ikke så viktige, nå når vi har definert en metrikk på $\mathbb{R}^n$, så vi kommer ikke til å gå nærmere inn på disse.

Underrom. Enhver delmengde av et metrisk rom er automatisk et metrisk rom: La X være et metrisk rom med metrikk d_X og $Y \subset X$ være en delmengde. Da kan man lett verifisere at funksjonen $d_Y : Y \times Y \to \mathbb{R}$ definert ved

$$d_Y(p, q) = d_X(p, q),$$

som er restriksjonen av $d_X : X \times X \to \mathbb{R}$ til $Y \times Y \subset X \times X$, oppfylle kriteriene til en metrikk på Y.

Definisjon 10.1.9: Indusert metrikk og underrom

Vi kaller metrikken d_Y for *den induserte metrikken*[a] eller *underromsmetrikken*[b] på Y og sier at Y *er et underrom*[c] *av* X, når Y har denne metrikken.

[a] *Induced metric* på engelsk.
[b] *Subspace metric* på engelsk.
[c] *Subspace* på engelsk.

EKSEMPEL 10.1.10. Siden metrikken i alle mengdene $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}$ og $\mathbb{N}$ er "den samme", det vil si, den som er indusert fra det største rommet $\mathbb{C}$, har vi at

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

er en kjede av underrom.

EKSEMPEL 10.1.11. For hver $n \leq m$ kan vi betrakte det euklidske rommet $\mathbb{R}^n$ som et underrom av $\mathbb{R}^m$ ved å identifisere et punkt $(x_1, \ldots, x_n) \in \mathbb{R}^n$ med punktet $(x_1, \ldots, x_n, 0, \ldots, 0) \in \mathbb{R}^m$; altså identifiserer vi $\mathbb{R}^n$ med delmengden

$$\mathbb{R}^n \times \underbrace{\{0, \ldots, 0\}}_{m-n \text{ ganger}} \subset \mathbb{R}^m.$$

I tilfellet $n = 1, m = 2$ betyr $\mathbb{R} \subset \mathbb{R}^2$ at vi identifiserer $\mathbb{R}$ med x-aksen i $\mathbb{R}^2$. Merk at dette samsvarer med hvordan vi representerer inklusjonen $\mathbb{R} \subset \mathbb{C}$ av de reelle tallene i det komplekse planet. I tilfellet $n = 2, m = 3$ betyr $\mathbb{R}^2 \subset \mathbb{R}^3$ at vi identifiserer $\mathbb{R}^2$ med xy-planet i $\mathbb{R}^3$.

10.2. Følger i metriske rom

Det meste av stoffet vi har gått gjennom om følger i $\mathbb{Q}$, $\mathbb{R}$ og $\mathbb{C}$ har sine opplagte generaliseringer til metriske rom og mer spesifikt til euklidske rom. Vi minner om at en følge i en vilkårlig mengde X er en funksjon $f : \mathbb{Z}^+ \to X$ (Definisjon 6.2.1). Vi definerer $x_n = f(n) \in X$ for alle $n \in \mathbb{Z}^+$, kaller disse for *leddene* i følgen og betegner følgen også som $\{x_n\}$.

Konvergens av en følge mot et punkt x betyr, som før, at vi kan få alle leddene langt nok ut i følgen så nær vi vil punktet x. Definisjonen er en opplagt generalisering av definisjonene for konvergens av følger i $\mathbb{Q}$, $\mathbb{R}$ og $\mathbb{C}$:

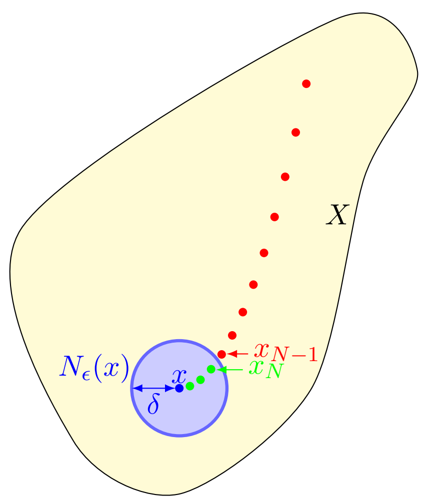
Definisjon 10.2.1: Konvergens av følge

La X være et metrisk rom med metrikk d. Følgen $\{x_n\}$ i X *konvergerer* mot $x \in X$, dersom det for ethvert reelt tall $\epsilon > 0$ finnes en $N \in \mathbb{Z}^+$ slik at $d(x_n, x) < \epsilon$ for alle $n \geq N$.

Vi skriver i så fall $\lim_{n\to\infty} x_n = x$ eller "$x_n \to x$ når $n \to \infty$" og kaller x for *grensen* til følgen.

Hvis følgen $\{x_n\}$ ikke konvergerer, sier vi at den *divergerer*.

At $d(x_n, x) < \epsilon$ betyr at $x_n \in N_\epsilon(x)$. Konvergens kan derfor visualiseres i følgende figur:



Følgen $\{x_n\}$ konvergerer mot x

Setning 10.2.2: Entydighet av grenser

Dersom en følge konvergerer mot x og x', da er $x = x'$.

Bevis

I prinsippet likt beviset for Setning 6.2.6 utledet i Oppgave 6.8. □

Definisjon 10.2.3: Begrenset følge

La X være et metrisk rom med metrikk d. En følge $\{x_n\}$ i X er *begrenset* dersom verdimengden til følgen er begrenset som delmengde av X, med andre ord: det finnes en $M \in \mathbb{R}$ slik at $d(x_n, x_m) \leq M$ for alle $m, n \in \mathbb{Z}^+$.

En følge er *ubegrenset* dersom den ikke er begrenset.

Setning 10.2.4

En konvergent følge er begrenset.

Bevis

I praksis likt beviset for Setning 6.2.11. □

Definisjon 10.2.5: Cauchy-følge

La X være et metrisk rom med metrikk d. En følge $\{x_n\}$ i X kalles en *Cauchy-følge* dersom det for ethvert reelt tall $\epsilon > 0$ finnes en $N \in \mathbb{Z}^+$ slik at $d(x_n, x_m) < \epsilon$ for alle $m, n \geq N$.

Setning 10.2.6

En Cauchy-følge i et metrisk rom er begrenset.

Bevis

I praksis likt beviset utledet for rasjonale følger i Oppgave 6.13. □

Setning 10.2.7

Enhver konvergent følge i et metrisk rom er en Cauchy-følge.

Bevis

I praksis likt beviset for Setning 6.3.3 for rasjonale følger. □

Vi har sett eksempler på metriske rom der alle Cauchy-følger konvergerer, nemlig $\mathbb{R}$ og $\mathbb{C}$ (jf. Setningene 7.6.1 og 8.5.12) og et eksempel på et metrisk rom der ikke alle Cauchy-følger konvergerer, nemlig $\mathbb{Q}$ (jf. Setning 6.3.4). Det motsatte av Setning 10.2.7 er altså ikke alltid sant. (Vi skal se på et eksempel til på dette i Setning 10.8.4.) Metriske rom som har den pene egenskapen at alle deres Cauchy-følger konvergerer har fått et eget navn:

Definisjon 10.2.8: Komplett metrisk rom

Et metrisk rom X kalles *komplett*[a] dersom alle Cauchy-følger i X konvergerer.

[a] *Complete* på engelsk.

Noen resultater for reelle følger har ingen generaliseringer til vilkårlige metriske rom. Det gjelder for eksempel *skvisesetningen* 7.4.13 og egenskapen *bevaring av orden* 7.4.14, fordi et vilkårlig metrisk rom ikke nødvendigvis er en ordnet mengde. Dessuten har regnereglene for grenser av følger i Setningene 7.4.8 (for reelle følger) og 8.5.4 (for komplekse følger) ingen generaliseringer til vilkårlige metriske rom, siden det generelt ikke finnes noen operasjoner som addisjon og multiplikasjon på dem.

Spesielle resultater for euklidske rom. I tilfellet euklidske rom finnes imidlertid flere resultater tilgjengelig.

Vi starter med å bemerke at konvergens av en følge i et euklidsk rom er ekvivalent med "konvergens av hver koordinatfølge":

Setning 10.2.9: Konvergens i euklidske rom

La $\mathbf{x}_k$ være en følge i det euklidske rommet $\mathbb{R}^n$ og la $\mathbf{x}_k = (x_{k,1}, \ldots, x_{k,n})$ for alle $k \in \mathbb{Z}^+$. Da konvergerer $\mathbf{x}_k$ mot $\mathbf{x} = (x_1, \ldots, x_n)$ hvis og bare hvis

$$\lim_{k \to \infty} x_{k,j} = x_j \quad \text{for alle} \quad j \in \{1, \ldots, n\}.$$

BEVIS
Overlates til Oppgave 10.8. □

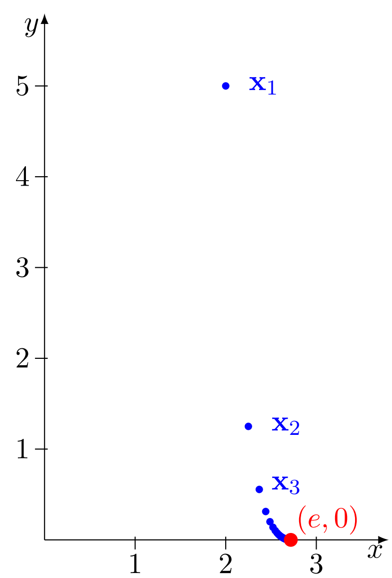
Vi har allerede sett et spesialtilfelle av dette resultatet i Setning 8.5.2, når vi identifiserer $\mathbb{C}$ med det euklidske rommet $\mathbb{R}^2$ (jf. også Eksempel 10.1.5).

EKSEMPEL 10.2.10. Finn grensen til følgen $\{\mathbf{x}_n\}$ i $\mathbb{R}^2$ gitt ved

$$\mathbf{x}_n = \left(\left(1 + \frac{1}{n}\right)^n, \frac{5}{n^2} \right),$$

om den finnes.

LØSNING. Vi vet at $\left(1 + \frac{1}{n}\right)^n \to e$ (Definisjon 7.5.7) og $\frac{5}{n^2} \to 0$. Ved Setning 10.2.9 er $\lim\limits_{n \to \infty} \mathbf{x}_n = (e, 0)$. □



FØLGEN $\left\{\mathbf{x}_n = \left(\left(1 + \frac{1}{n^2}\right)^n, \frac{5}{n}\right)\right\}$ KONVERGERER MOT $(e, 0)$

Regnereglene for grenser av følger i Setningene 7.4.8 og 8.5.4 generaliseres til euklidske rom:

Setning 10.2.11: Regler for grenser av følger i euklidske rom

La $\{\mathbf{x}_k\}$ og $\{\mathbf{y}_k\}$ være konvergente følger i det euklidske rommet $\mathbb{R}^n$ og $\{c_k\}$ en konvergent følge i $\mathbb{R}$. La $\mathbf{x} = \lim_{k\to\infty} \mathbf{x}_k$, $\mathbf{y} = \lim_{k\to\infty} \mathbf{y}_k$ og $c = \lim_{k\to\infty} c_k$. Da gjelder:

(i) $\lim_{k\to\infty} (\mathbf{x}_k + \mathbf{y}_k) = \mathbf{x} + \mathbf{y}$;

(ii) $\lim_{k\to\infty} (\mathbf{x}_k - \mathbf{y}_k) = \mathbf{x} - \mathbf{y}$;

(iii) $\lim_{k\to\infty} (c_k\mathbf{x}_k) = c\mathbf{x}$;

(iv) $\lim_{k\to\infty} (\mathbf{x}_k \cdot \mathbf{y}_k) = \mathbf{x} \cdot \mathbf{y}$.

Bevis

Dette er ganske så likt beviset for Setning 6.2.14, ved å erstatte metrikken på $\mathbb{Q}$ med den euklidske metrikken på $\mathbb{R}^n$. Alternativt er resultatet en nesten umiddelbar konsekvens av regnereglene i $\mathbb{R}$ (Setning 7.4.8) og Setning 10.2.9. $\square$

Dessuten har vi følgende generalisering av Setningene 7.6.1 og 8.5.12:

Setning 10.2.12

Euklidske rom er komplette.

Bevis

Bevisideen er temmelig lik bevisideen for Setning 7.6.1 og vi overlater detaljene til Oppgave 10.9. $\square$

Komplettering av metriske rom. Vi har sett hvordan vi kan ta det ikke-komplette metrisk rommet $\mathbb{Q}$ og "tette igjen" hull slik at vi får det komplett metriske rommet $\mathbb{R}$ som inneholder $\mathbb{Q}$ (Teorem 7.1.2), og det følger av konstruksjonen at $\mathbb{Q}$ er et underrom av $\mathbb{R}$. Vi så i tillegg at $\mathbb{Q}$ er tett i $\mathbb{R}$ (Setning 7.2.2). Dette er et spesialtilfelle av en mer generell konstruksjon. Før vi ser på det generelle resultatet, må vi definere hva vi mener med tetthet i et vilkårlig metrisk rom.

Definisjon 10.2.13: Tetthet

La X være et metrisk rom og $A \subset X$ en delmengde.

Vi sier at A er *tett*[a] *i* X dersom følgende holder: for enhver $x \in X$ og enhver omegn $N_\delta(x)$ er $N_\delta(x) \cap A \neq \emptyset$.

[a] *Dense* på engelsk.

Det er lett å se at dette er akkurat det samme som vi har sagt tidligere for $\mathbb{Q} \subset \mathbb{R}$. Se Oppgave 10.10 for en ekvivalent beskrivelse av tetthet. Resultatet som generaliserer konstruksjonen vår av $\mathbb{R}$ fra $\mathbb{Q}$ er følgende:

Teorem 10.2.14

La X være et metrisk rom. Da finnes et komplett metrisk rom X^* slik at X er et tett underrom av X^*.

Bevis
Bevisideen er temmelig lik bevisideen for Teorem 7.1.2: Punktene i X^* er ekvivalensklasser av Cauchyfølger i X på en slik måte at punktene i X representeres av konstante følger. Vi går ikke gjennom detaljene her. $\square$

Merknad 10.2.15. Rommet X^* i Teorem 10.2.14 kalles en *komplettering*[6] av X. Det kan finnes flere slike, men alle kan vises å være *isometriske*, et begrep vi nå forklarer nærmere.

To metriske rom X og Y med metrikker d_X og d_Y sies å være *isometriske*[7], dersom det finnes en bijeksjon $f : X \to Y$ slik at

$$d_Y(f(x), f(x')) = d_X(x, x') \text{ for alle } x, x' \in X.$$

Med andre ord bevarer funksjonen avstanden mellom alle punktene i tillegg til å gi en én-til-én korrespondanse mellom mengdene. En slik funksjon kalles en *isometri*[8].

I praksis kan vi jobbe med isometriske metriske rom "som om de var like"; eneste forskjellen er at objektene har forskjellige "navn".

At alle kompletteringer er isometriske uttrykkes som at *ethvert metrisk rom har en komplettering som er entydig opp til isometri.*

10.3. Åpne og lukkede mengder

En del begreper vi har blitt kjent med i $\mathbb{Q}$, $\mathbb{R}$ og $\mathbb{C}$ generaliseres på en helt naturlig måte til vilkårlige metriske rom. Det gjelder for eksempel begrepene *åpen* og *lukket* mengde. Stoffet i denne seksjonen vil være spesielt nyttig når dere jobber med funksjoner av flere variable i kursene *MAT112–Grunnkurs i matematikk II* og *MAT212–Funksjoner av flere variable.*

Åpne mengder. Følgende definisjon er den naturlige generaliseringen av definisjonen på en åpen delmengde av $\mathbb{R}$ gitt i Oppgave 7.28:

[6] *Completion* på engelsk.
[7] *Isometric* på engelsk.
[8] *Isometry* på engelsk.

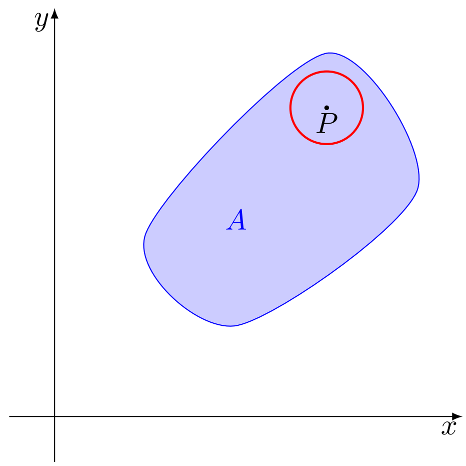
Definisjon 10.3.1: Indre punkt og åpen mengde

La X være et metrisk rom med metrikk d og $A \subset X$ en delmengde.

(i) Et punkt $a \in A$ kalles et *indre punkt*[a] *i* A dersom det finnes en omegn $N_\delta(a) \subset A$, for en $\delta > 0$.

(ii) Vi sier at A er *åpen* (i X) dersom ethvert punkt i A er et indre punkt (med andre ord: for hver $a \in A$ finnes det en omegn $N_\delta(a) \subset A$, for en $\delta > 0$).

[a] *Interior point* på engelsk.



Et indre punkt P av en delmengde $A \subset \mathbb{R}^2$

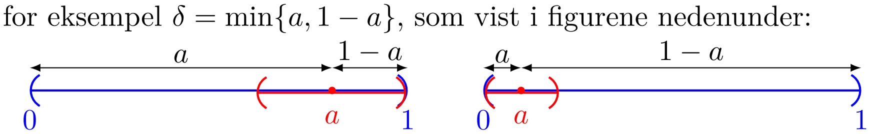
Merk at følgende delmengder av X er åpne per definisjon:

- Mengden X selv, siden enhver omegn ligger automatisk i X (per definisjon).
- Den tomme mengden $\emptyset$, siden den ikke inneholder punkter, slik at betingelsen (ii) automatisk er oppfylt.

Eksempel 10.3.2. (i) Mengden $(0,1)$ er åpen i $\mathbb{R}$, for hvis $a \in (0,1)$, finnes en $\delta > 0$ slik at

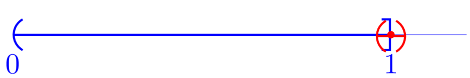
$$(a-\delta, a+\delta) \subset (0,1);$$

ta for eksempel $\delta = \min\{a, 1-a\}$, som vist i figurene nedenunder:



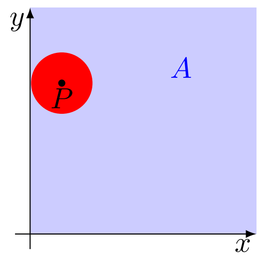
Omegner $(a-\delta, a+\delta)$ til a i rødt med $\delta = 1-a$ til venstre og $\delta = a$ til høyre

(ii) Mengden $(0,1]$ er *ikke* åpen i $\mathbb{R}$, siden enhver omegn $(1-\delta, 1+\delta)$ om 1 vil inneholde punkter utenfor $(0,1]$, nemlig punktene i $(1, 1+\delta)$, som vist i figuren:



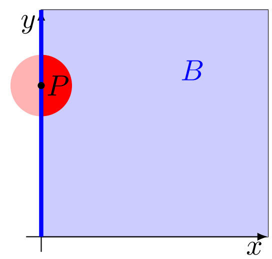
Omegnen $(1-\delta, 1+\delta)$ ligger ikke innenfor $(0,1]$, uansett hvor liten δ er.

EKSEMPEL 10.3.3. (i) Mengden $A = \{(x, y) \in \mathbb{R}^2 \mid x > 0,\ y > 0\}$ er åpen i $\mathbb{R}^2$. Dette er første kvadrant i planet uten aksene, og vi kan slå en sirkel rundt ethvert punkt i A som ligger helt innenfor A, og det indre av sirkelen vil da være en omegn om punktet som ligger i A, som vist i figuren:



EN OMEGN OM P SOM LIGGER INNENFOR A

(ii) Mengden $B = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0,\ y > 0\}$ er *ikke* åpen i $\mathbb{R}^2$, siden det indre av enhver sirkel rundt et punkt P på y-aksen vil inneholde punkter med $x < 0$, det vil si punkter utenfor B, som vist i figuren:



INGEN OMEGN OM P PÅ y-AKSEN KAN LIGGE INNENFOR B

Dessuten har vi også:

Setning 10.3.4

En omegn i et metrisk rom er åpen.

BEVIS

En omegn i et metrisk rom X er på formen $N_r(a)$ for en $a \in X$ og en $r \in \mathbb{R}^+$. Vi må vise at ethvert punkt $x \in N_r(a)$ er et indre punkt, det vil si at det finnes en $\delta \in \mathbb{R}^+$ slik at $N_\delta(x) \subset N_r(a)$. Inklusjonen betyr at enhver $z \in N_\delta(x)$ ligger også i $N_r(a)$, med andre ord at

$$d(z, x) < \delta \Longrightarrow d(z, a) < r.$$

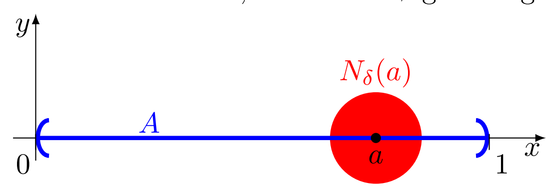
Det overlates til Oppgave 10.15 å vise at $\delta = r - d(x, a)$ duger. $\square$

Vi må dessuten være klar over at egenskapen av å være åpen avhenger av hvilket rom vi betrakter A som underrom av. Dersom $Y \subset X$ er to metriske rom, med Y et underrom av X, og vi har $A \subset Y$, så er A en delmengde både av Y og av X. Egenskapen "åpen" avhenger av om vi betrakter A som en delmengde av Y eller X, som følgende eksempel viser:

EKSEMPEL 10.3.5. (i) Betrakt $A = (0,1) \subset Y = \mathbb{R} \subset X = \mathbb{R}^2$. For den siste inklusjonen identifiserer vi $\mathbb{R}$ med delmengden

$$\{(x,0) \mid x \in \mathbb{R}\} = \text{“}x\text{-aksen”} \subset \mathbb{R}^2,$$

som nevnt i Eksempel 10.1.11. Da er A åpen som delmengde av $Y = \mathbb{R}$, som vist i Eksempel 10.3.2. Imidlertid er A ikke åpen som delmengde av $\mathbb{R}^2$: hvis $a \in A$, så er en δ-omegn $N_\delta(a)$ om a i $\mathbb{R}^2$ det indre av en sirkel av radius δ om a, og en slik kan ikke være inneholdt i A, for den vil nødvendigvis inneholde punkter utenfor x-aksen, som vist i følgende figur:



ILLUSTRASJON TIL EKSEMPEL 10.3.5(I)

(ii) Betrakt $A = [0,1) \subset Y = [0,\infty) = \mathbb{R}^{\geq 0} \subset X = \mathbb{R}$. Da er A åpen som delmengde av Y: en δ-omegn om $a \in A$ i Y er nemlig mengden av alle $x \in Y$ slik at $|x - a| < \delta$, og dette kan skrives som

$$(a-\delta, a+\delta) \cap Y = (a-\delta, a+\delta) \cap [0,\infty) = \begin{cases} (a-\delta, a+\delta), & \text{hvis } \delta < a, \\ [0, a+\delta), & \text{hvis } \delta \geq a \end{cases}$$

(lag en tegning!). En omegn om 0 i Y vil altså se ut som $[0,\delta)$. For en vilkårlig $a \in A$, kan vi ta $\delta = 1 - a$, og da vil $(a-\delta, a+\delta) \cap Y \subset A$.

Men A er ikke åpen som delmengde av $X = \mathbb{R}$, siden enhver omegn om 0 i $\mathbb{R}$ vil inneholde negative tall og derfor ikke være inneholdt i A.

Siste eksempel viser at vi alltid må være klar over hvilken “stor mengde” vi refererer til når vi sier at en delmengde er åpen. Tilbake til den generelle situasjonen $A \subset Y \subset X$ og $a \in A$, der vi betegner metrikkene på X og Y med samme bokstav d, så er en δ-omegn om a i Y

$$N_\delta^Y(a) = \{x \in Y \mid d(a,x) < \delta\},$$

mens en δ-omegn om a i X er

$$N_\delta^X(a) = \{x \in X \mid d(a,x) < \delta\}.$$

Det er lett å se at

$$N_\delta^Y(a) = N_\delta^X(a) \cap Y. \tag{290}$$

Generelt har vi følgende resultat som lenker åpenhet i Y og i X:

Setning 10.3.6

La Y være et underrom av X og $A \subset Y$ en delmengde. Da er A åpen i Y hvis og bare hvis $A = Y \cap G$ for en åpen delmengde G av X.

Bevis
Vi viser først “hvis”-delen (“$\Longleftarrow$”).

Anta at $A = Y \cap G$ for en åpen delmengde G av X. Vi vil vise at ethvert punkt $a \in A$ er et indre punkt i A . Siden $a \in G$ og G er åpen i X, finnes en $\delta \in \mathbb{R}^+$ slik at $N_\delta^X(a) \subset G$. Ved (290) har vi

$$N_\delta^Y(a) = N_\delta^X(a) \cap Y \subset G \cap Y = A,$$

som viser at a er et indre punkt i A.

Vi viser dernest “bare hvis”-delen (“$\Longrightarrow$”).

Anta at A er åpen i Y. Dette betyr at for enhver $a \in A$, finnes $\delta_a \in \mathbb{R}^+$ slik at $N_{\delta_a}^Y(a) \subset A$. Vi har $A \subset \cup_{a\in A} N_{\delta_a}^Y(a)$, og siden hver $N_{\delta_a}^Y(a) \subset A$, er

$$A = \cup_{a\in A} N_{\delta_a}^Y(a).$$

Dermed har vi

$$A = \cup_{a\in A} N_{\delta_a}^Y(a) \overset{(290)}{=} \cup_{a\in A} \Big(N_{\delta_a}^X(a) \cap Y\Big) \subset \cup_{a\in A} N_{\delta_a}^X(a).$$

Definerer vi $G = \cup_{a\in A} N_{\delta_a}^X(a)$, har vi at $A = Y \cap G$. (Vi overlater verifiseringen av dette til Oppgave 10.16.)

Til slutt viser vi at G er åpen i X: La $x \in G$. Da er $x \in N_{\delta_a}^X(a)$ for en $a \in A$. Siden $N_{\delta_a}^X(a)$ er åpen i X ved Setning 10.3.4, finnes en $r \in \mathbb{R}^+$ slik at $N_r^X(x) \subset N_{\delta_a}^X(a) \subset G$. Dermed er x et indre punkt i G. $\square$

Dette er et hendig kriterium for å sammenligne begrepet “åpen” i forhold til rommene X og Y, som neste eksempel viser:

Eksempel 10.3.7. Betrakt igjen $A = [0,1) \subset Y = [0,\infty) = \mathbb{R}^{\geq 0} \subset X = \mathbb{R}$ fra Eksempel 10.3.5. Siden $A = Y \cap (-1,\infty)$ og $(-1,\infty)$ er åpen i $X = \mathbb{R}$, er A åpen i Y ved Setning 10.3.6, som bekrefter det vi fant ut i Eksempel 10.3.5.

Lukkede mengder. Vi kan også definere hva det vil si at en mengde er lukket i et vilkårlig metrisk rom ved å generalisere definisjonen på en lukket delmengde av $\mathbb{R}$ gitt i Oppgave 7.28:

Definisjon 10.3.8: Lukket mengde

La X være et metrisk rom og $A \subset X$ en delmengde.

Vi sier at A er *lukket* dersom følgende holder: for enhver følge $\{x_n\}$ i A slik at x_n konvergerer mot $x \in X$, da er $x \in A$. (Vi uttrykker dette kriteriet som at A *inneholder alle sine følgegrenser*.)

Merk at følgende delmengder av X er lukket per definisjon:

- Mengden X selv, siden enhver følge som konvergerer i X konvergerer i X;
- Den tomme mengden $\emptyset$, siden den ikke inneholder punkter, ei heller noen følger.

Merknad 10.3.9. Det er viktig å være klar over at "lukket" ikke er ekvivalent med å "ikke være åpen". Som vi vil se i Eksemplene 10.3.10 og 10.3.11 kan en mengde være verken åpen eller lukket. Dette er grunnlaget for den matematiske vitsen: "Hva er forskjellen på en dør og en mengde? En dør er alltid enten åpen eller lukket, og aldri begge deler."

Eksempel 10.3.10. Betrakt mengden $B = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0,\ y > 0\}$ fra Eksempel 10.3.3(ii), hvor vi viste at B ikke er åpen i $\mathbb{R}^2$. Den er heller ikke lukket: for eksempel vil følgen fra Eksempel 10.2.10 ligge i B, men konvergere mot $(e, 0)$, som ligger utenfor B.

Som i tilfellet for åpne mengder avhenger begrepet lukkethet ikke bare av mengden A, men av det "store" rommet X vi betrakter A en delmengde av. Neste eksempel illustrerer dette:

Eksempel 10.3.11. Betrakt $A = (0, 1] \subset \mathbb{R}$. Da er A ikke lukket i $\mathbb{R}$, siden for eksempel følgen $\{\frac{1}{n}\}$ ligger i A og konvergerer mot $0 \in \mathbb{R} \setminus A$. (Siden vi viste i Eksemplene 10.3.5 og 10.3.7 at A heller ikke er åpen, viser dette igjen at en mengde trenger verken å være åpen eller lukket.)

Men A er lukket i $\mathbb{R}^+ = (0, \infty)$. For å vise dette må vi vise at for enhver følge $\{x_n\}$ i A som konvergerer *mot et punkt* $x \in \mathbb{R}^+$, så er $x \in A$. Siden $\{x_n\}$ ligger i A, har vi $0 < x_n \leq 1$ for alle $n \in \mathbb{Z}^+$. Siden grenser bevarer orden (Setning 7.4.14), må $0 \leq x \leq 1$. Siden vi vet at $x \in \mathbb{R}^+$, må derfor $x \in (0, 1]$. Altså er $x \in A$. Dette viser at A er lukket som delmengde av $\mathbb{R}^+$.

Generelt har vi følgende resultat som tilsvarer Setning 10.3.6:

Setning 10.3.12

La Y være et underrom av X og $A \subset Y$ en delmengde. Da er A lukket i Y hvis og bare hvis $A = Y \cap G$ for en lukket delmengde G av X.

Bevis

Overlates til Oppgave 10.20. □

Følgende resultat er veldig viktig og generaliserer resultatet fra Oppgave 7.28 for delmengder av $\mathbb{R}$:

Setning 10.3.13: Ekvivalent beskrivelse av lukkethet, I

La X være et metrisk rom og $A \subset X$ en delmengde. Da gjelder:

$$A \text{ er lukket} \iff X \setminus A \text{ er åpen}$$

Bevis

Vi viser først "bare hvis"-delen ("$\Longrightarrow$").

Anta at A er lukket. Vi vil vise at ethvert punkt $x \in X \setminus A$ er et indre punkt. For å gjøre det, vil vi vise at det finnes en $n \in \mathbb{Z}^+$ slik at $N_{\frac{1}{n}}(x) \subset (X \setminus A)$. Hvis dette ikke var sant, ville $N_{\frac{1}{n}}(x) \cap A \neq \emptyset$ for alle $n \in \mathbb{Z}^+$. Dette betyr at det for hver $n \in \mathbb{Z}^+$ eksisterer et punkt $x_n \in N_{\frac{1}{n}}(x) \cap A$. Vi får derfor en følge $\{x_n\}$ i A slik at $d(x_n, x) < \frac{1}{n}$ for alle $n \in \mathbb{Z}^+$. Dermed vil x_n konvergere mot x. Siden A er lukket, må $x \in A$, en selvmotsigelse, siden $x \in X \setminus A$.

Vi viser dernest "hvis"-delen ("$\Longleftarrow$").

Anta at $X \setminus A$ er åpen. Vi vil vise at A er lukket. La derfor $\{x_n\}$ være en følge i A som konvergerer mot $x \in X$. Vi vil vise at $x \in A$. Per definisjon av konvergens vil det for enhver $\epsilon \in \mathbb{R}^+$ eksistere en $N \in \mathbb{Z}^+$ slik at $d(x_n, x) < \epsilon$ for alle $n \geq N$. Dette betyr at $x_n \in N_\epsilon(x)$ for alle $n \geq N$. Altså er $N_\epsilon(x) \cap A \neq \emptyset$. Det finnes dermed ingen $\epsilon \in \mathbb{R}^+$ slik at $N_\epsilon(x) \subset (X \setminus A)$. Siden $X \setminus A$ er åpen, må $x \notin X \setminus A$. Altså er $x \in A$, som ønsket. $\square$

Merknad 10.3.14. Mengden $X \setminus A$ kalles *komplementet*[9] til A. Det er viktig å merke seg at komplementet til A er avhengig av "det store rommet" X som vi betrakter A som delmengde av.

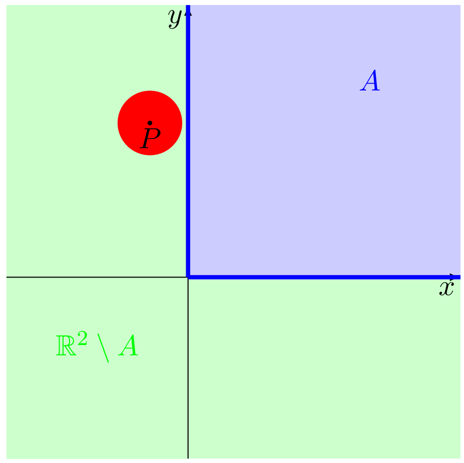
I mange lærebøker blir egenskapen i Setning 10.3.13 brukt som *definisjon* på lukkethet.

Eksempel 10.3.15. Mengden $[0, 1]$ er lukket i $\mathbb{R}$, fordi komplementet er $(-\infty, 0) \cup (1, \infty)$, som er åpen i $\mathbb{R}$. Av samme grunn er mengden $[0, \infty)$ lukket i $\mathbb{R}$, fordi komplementet er $(-\infty, 0)$.

Mengden $[0, 1)$ er ikke lukket i $\mathbb{R}$, fordi komplementet $(-\infty, 0) \cup [1, \infty)$ ikke er åpent i $\mathbb{R}$: punktet 1 er nemlig ikke et indre punkt.

Eksempel 10.3.16. Mengden $A = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0,\ y \geq 0\}$ er lukket i $\mathbb{R}^2$, fordi komplementet $\mathbb{R}^2 \setminus A = \{(x, y) \in \mathbb{R}^2 \mid x < 0 \text{ eller } y < 0\}$ er åpent: Rundt ethvert punkt i $\mathbb{R}^2 \setminus A$ kan vi slå en sirkel som ligger helt innenfor $\mathbb{R}^2 \setminus A$:

[9] *Complement* på engelsk.



Det finnes alltid omegner om $P \in \mathbb{R}^2 \setminus A$ som ligger innenfor $\mathbb{R}^2 \setminus A$

I $\mathbb{R}$ er vi imidlertid kanskje mest vant til å tenke på et intervall som lukket dersom det inneholder sine endepunkter: slik ser vi f.eks. at $[0, 1]$ og $[0, \infty)$ er lukket, mens $(0, 1]$ og $(0, \infty)$ ikke er lukket. Det finnes en generalisering av begrepet "endepunkter" i alle metriske rom:

Definisjon 10.3.17: Randpunkt og rand

La X være et metrisk rom og $A \subset X$ en delmengde.

Et punkt $x \in X$ er et randpunkt[a] til A hvis enhver omegn om x innholder både punkter i A og utenfor A; med andre ord: for enhver $\delta \in \mathbb{R}^+$ har vi

$$N_\delta(x) \cap A \neq \emptyset \text{ og } N_\delta(x) \cap (X \setminus A) \neq \emptyset. \tag{291}$$

Mengden av alle randpunkter til A kalles randen[b] til A og betegnes gjerne med $\mathrm{bd}(A)$ eller ∂A.

[a] *Boundary point* på engelsk, mer sjeldent kalt *frontier point*.
[b] *Boundary* på engelsk, mer sjeldent kalt *frontier*.

Merk at et randpunkt til A kan være med i A eller ikke. Merk også at $\partial X = \emptyset$.

Følgende er klart fra definisjonen (jf. symmetrien mellom A og $X \setminus A$ i (291)):

Observasjon 10.3.18

$\partial A = \partial(X \setminus A)$.

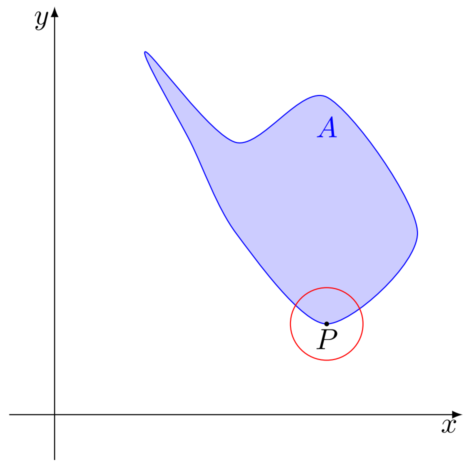
Det følger også fra definisjonen at et punkt $x \in A$ er x enten et indre punkt eller randpunkt til A, men ikke begge deler. Med andre ord er

$$A \cap \partial A = A \setminus \{\text{indre punkter i } A\}. \tag{292}$$

(Vi overlater til Oppgave 10.22 å vise dette.)

EKSEMPEL 10.3.19. Randen til hvert av de endelige intervallene (a, b), $[a, b)$, $(a, b]$ og $[a, b]$ er $\{a, b\}$. Randen til intervallene (a, ∞) og $[a, \infty)$ er $\{a\}$. Randen til intervallene $(-\infty, b)$ og $(-\infty, b]$ er $\{b\}$. Randen til $\mathbb{R}$ er den tomme mengden.

EKSEMPEL 10.3.20. La $A \subset \mathbb{R}^2$. Da er x et randpunkt til A dersom det indre av enhver sirkel vi slår med sentrum i x vil inneholde punkter både i og utenfor A, se figuren:



ET RANDPUNKT P AV EN DELMENGDE $A \subset \mathbb{R}^2$

Følgende resultat kan tas som ekvivalent definisjon på det å være lukket:

Setning 10.3.21: Ekvivalent beskrivelse av lukkethet, II

La X være et metrisk rom og $A \subset X$ en delmengde. Da er A lukket hvis og bare hvis $\partial A \subset A$.

BEVIS

Vi bruker Setning 10.3.13 og egenskap (292), sammen med Observasjon 10.3.18:

$$\begin{aligned} A \text{ lukket} &\overset{10.3.13}{\iff} X \setminus A \text{ åpen} \\ &\overset{\text{def}}{\iff} X \setminus A = \{\text{indre punkter i } X \setminus A\} \\ &\overset{(292)}{\iff} \partial(X \setminus A) \cap (X \setminus A) = \emptyset \\ &\overset{\text{Oppg. 2.6}}{\iff} \partial(X \setminus A) \subset A \\ &\overset{10.3.18}{\iff} \partial A \subset A. \end{aligned}$$

$\square$

Eksempel 10.3.22. Hver av delmengdene $(0,1)$, $[0,1)$, $(0,1]$ og $[0,1]$ i $\mathbb{R}$ har 0 og 1 som randpunkter. Ved Setning 10.3.21 er bare $[0,1]$ lukket, siden bare denne inneholder begge randpunktene.

Eksempel 10.3.23. Delmengden

$$A = \{(x,y) \in \mathbb{R}^2 \mid x \geq 0,\ y \geq 0\}$$

av $\mathbb{R}^2$ har rand

$$\partial A = \{(0,y) \mid y \geq 0\} \cup \{(x,0) \mid x \geq 0\},$$

som er unionen av de to ikke-negative delene av koordinataksene. Siden $\partial A \subset A$, er A lukket, ved Setning 10.3.21.

Eksempel 10.3.24. Delmengden

$$R = \{(x,y) \in \mathbb{R}^2 \mid y \geq 0 \text{ og } x^2 + y^2 < 1\}$$

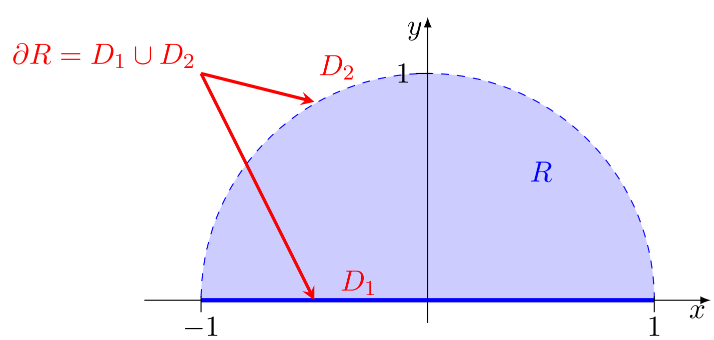
av $\mathbb{R}^2$ har rand $\partial R = D_1 \cup D_2$, hvor

$$D_1 = \{(x,y) \in \mathbb{R}^2 \mid y = 0, -1 \leq x \leq 1\}$$

og

$$D_2 = \{(x,y) \in \mathbb{R}^2 \mid y > 0, x^2 + y^2 = 1\},$$

som vist i figuren nedenunder. Siden $D_2 \not\subset R$, er R ikke lukket, ved Setning 10.3.21.



R er ikke lukket fordi $\partial R = D_1 \cup D_2$ og $D_2 \not\subset R$

Union og snitt av åpne og lukkede mengder. Vi definerte union og snitt av to mengder i §2. Det er opplagt hvordan man generaliserer definisjonene til union og snitt av endelig mange mengder. Men man kan også snakke om union og snitt av vilkårlig mange (også uendelig og overtellbart mange) mengder: La $\mathcal{A}$ være en mengde slik at hvert element $\alpha \in \mathcal{A}$

representerer en mengde A_α. Da er mengden $\{A_\alpha \mid \alpha \in \mathcal{A}\}$, noen ganger også uttrykt som $\{A_\alpha\}_{\alpha\in\mathcal{A}}$, en *mengde bestående av mengder*. Vi kaller en slik mengde gjerne for en *samling*[10] eller *familie*[11] (av mengder). Noen ganger legger vi til at den er *parametrisert av* $\mathcal{A}$.

Vi definerer *unionen av alle mengdene* A_α som

$$\bigcup_{\alpha\in\mathcal{A}} A_\alpha = \{x \mid x \in A_\alpha \ \text{ for en } \ \alpha \in \mathcal{A}\}$$

(noen ganger forenklet til $\bigcup_\alpha A_\alpha$) og *snittet av alle mengdene* A_α som

$$\bigcap_{\alpha\in\mathcal{A}} A_\alpha = \{x \mid x \in A_\alpha \ \text{ for alle } \ \alpha \in \mathcal{A}\}$$

(noen ganger forenklet til $\bigcap_\alpha A_\alpha$) Med andre ord er $\cup_\alpha A_\alpha$ alle elementene som er med i minst én av $A_\alpha$ene, mens $\cap_\alpha A_\alpha$ er alle elementene som er med i alle $A_\alpha$ene.

Som oftest er alle A_α delmengder av en "stor" mengde X, slik at vi kan snakke om at mengdene A_α er åpne eller lukkede. (Vi husker at disse begrepene avhenger av hvilket rom vi betrakter mengdene som delmengder av.)

Eksempel 10.3.25. (i) La $\mathcal{A} = \mathbb{R}^+$ og $A_\alpha = (-\alpha, \alpha) \subset \mathbb{R}$ for hver $\alpha \in \mathbb{R}^+$. Da er

$$\bigcup_{\alpha\in\mathcal{A}} A_\alpha = \mathbb{R}$$

og

$$\bigcap_{\alpha\in\mathcal{A}} A_\alpha = \{0\}.$$

Vi ser altså at et snitt av uendelig mange åpne mengder ikke trenger være åpent.

(ii) La $\mathcal{A} = \mathbb{Z}^+$ og $A_\alpha = [-1, 1 - \frac{1}{\alpha}]$. Da er

$$\bigcap_{\alpha\in\mathcal{A}} A_\alpha = [-1, 0]$$

og

$$\bigcup_{\alpha\in\mathcal{A}} A_\alpha = [-1, 1).$$

Vi ser altså at en union av uendelig mange lukkede mengder ikke trenger være lukket.

Vi har følgende sentrale resultat som garanterer en "pen" oppførsel av union og snitt av åpne og lukkede mengder:

[10] *Collection* på engelsk.

[11] *Family* på engelsk.

Setning 10.3.26

(i) For enhver familie $\{A_\alpha\}$ av åpne mengder er $\bigcup_\alpha A_\alpha$ åpen.
(ii) For enhver endelig familie $\{A_1, \ldots, A_n\}$ av åpne mengder er $A_1 \cap \cdots \cap A_n$ åpen.
(iii) For enhver familie $\{A_\alpha\}$ av lukkede mengder er $\bigcap_\alpha A_\alpha$ lukket.
(iv) For enhver endelig familie $\{A_1, \ldots, A_n\}$ av lukkede mengder er $A_1 \cup \cdots \cup A_n$ lukket.

Eksempel 10.3.25 viser at (ii) og (iv) ikke kan utvides til å gjelde uendelig mange mengder.

BEVIS FOR SETNING 10.3.26
Vi viser (i) og (ii) og overlater (iii) og (iv) til Oppgave 10.25.

(i) La $A = \bigcup_\alpha A_\alpha$ og la $x \in A$. Da er $x \in A_{\alpha_0}$ for en α_0. Siden A_{α_0} er åpen, finnes en $\delta > 0$ slik at $N_\delta(x) \subset A_{\alpha_0}$. Siden $A_{\alpha_0} \subset A$, er $N_\delta(x) \subset A$, slik at x er et indre punkt også i A. Dette viser at A er åpen.

(ii) La $B = A_1 \cap \cdots \cap A_n$ og la $x \in B$. Da er $x \in A_i$ for alle i. Siden hver A_i er åpen, finnes $\delta_i > 0$ slik at $N_{\delta_i}(x) \subset A_i$ for alle $i \in \{1, \ldots, n\}$. La $\delta = \min\{\delta_1, \ldots, \delta_n\}$. Da er $N_\delta(x) \subset N_{\delta_i}(x)$ for alle i, og dermed $N_\delta(x) \subset A_i$ for alle i. Dermed er $N_\delta(x) \subset B$, som viser at x er et indre punkt i B. Dette viser at B er åpen. $\square$

10.4. Kontinuitet i metriske rom

Det er nesten opplagt hvordan man definerer kontinuitet for funksjoner mellom metriske rom: Vi generaliserer Definisjon 9.1.1 ved å bare bytte ut den spesielle metrikken i $\mathbb{R}$ med en vilkårlig metrikk:

Definisjon 10.4.1: Kontinuitet

La X og Y være metriske rom med metrikker d_X og d_Y, henholdsvis, og $f : X \to Y$ en funksjon og $a \in X$.

Funksjonen f er *kontinuerlig* i a dersom det til ethvert reelt tall $\epsilon > 0$ finnes et reelt tall $\delta > 0$ slik at

$$d_Y(f(x), f(a)) < \epsilon \text{ når } d_X(x, a) < \delta. \tag{293}$$

Dersom f ikke er kontinuerlig i a, sier vi at f er *diskontinuerlig* i a, eller at f *har en diskontinuitet* i a.

Vi sier at f er *kontinuerlig* dersom f er kontinuerlig i alle $a \in X$.

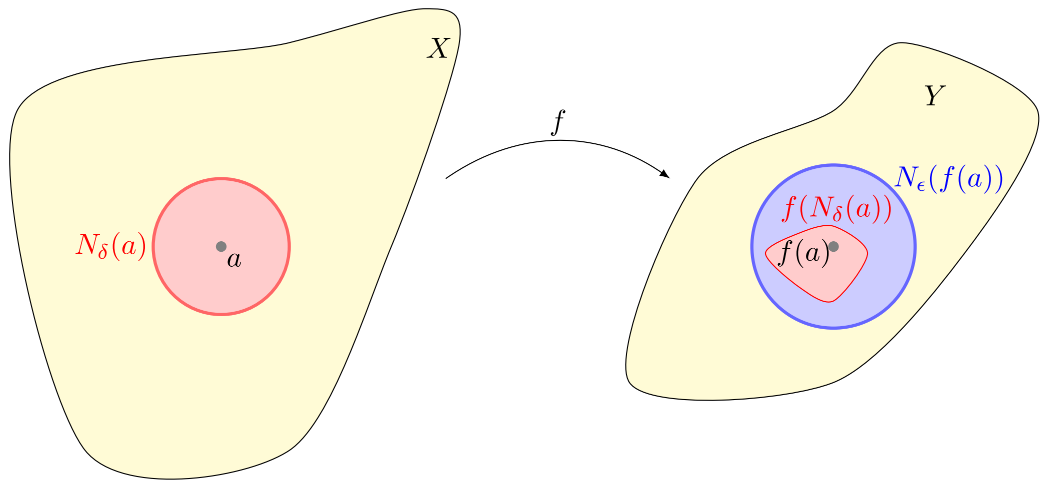
Betingelsen (293) kan skrives som

$$f(x) \in N_\epsilon(f(a)) \text{ når } x \in N_\delta(a)$$

og dermed som

$$f\left(N_\delta(a)\right) \subset N_\epsilon(f(a)), \tag{294}$$

som vist i følgende figur:



Kontinuitet av $f : X \to Y$ i punktet a.

Intuitivt betyr kontinuitet fremdeles at vi kan sørge for at funksjonsverdiene $f(x)$ er så nær vi vil $f(a)$, bare x er nær nok a.

Merknad 10.4.2. Vi har sett og ofte brukt at vi for en funksjon $f : X \to Y$ også kan skrive $f : X \to V(f)$, der $V(f) \subset Y$ er verdimengden til f, selv om vi strengt tatt burde betegne funksjonen med et annet symbol. Dette er heldigvis uproblematisk med tanke på kontinuitet: $f : X \to Y$ er kontinuerlig hvis og bare hvis $f : X \to V(f)$ er kontinuerlig, noe som kan lett sjekkes ved definisjonen.

Helt generelt har vi at hvis X, Y, Z er metriske rom, $Y \subset Z$ er et underrom og $f : X \to Y$ er en funksjon, da er f kontinuerlig hvis og bare hvis $i \circ f : X \to Z$ er kontinuerlig, hvor $i : Y \to Z$ er inklusjonsavbildningen (jf. Oppgave 10.27). Vi sier at *kontinuitet er bevart ved utvidelse og restriksjon av kodomenet.*

Tilsvarende har vi at hvis $f : X \to Y$ er en kontinuerlig funksjon mellom metriske rom og $Z \subset X$ er et underrom, da er også restriksjonen $f|_Z : Z \to Y$ kontinuerlig (jf. Oppgave 10.26). Vi sier at *kontinuitet er bevart ved restriksjon av definisjonsmengden/domenet.*

Eksempel 10.4.3 (Konstante funksjoner er kontinuerlige). La $f : X \to Y$ være en funksjon mellom metriske rom slik at $f(x) = y_0$ for alle $x \in X$, for en fiksert $y_0 \in Y$. Vi sier da at f er *konstant*. Vis at f er kontinuerlig.

Løsning. Siden $d_Y(f(x), f(a)) = d_Y(y_0, y_0) = 0$ uansett $x, a \in X$ vil enhver $\delta > 0$ fungere for enhver $\epsilon > 0$. $\square$

Vi har som før en del ekvivalente beskrivelser av kontinuitet. Den ene er igjen følge-kriteriet:

Setning 10.4.4: Følge-kriteriet for kontinuitet

La $f : X \to Y$ være en funksjon mellom to metriske rom. Funksjonen f er kontinuerlig i $a \in X$ hvis og bare hvis $\lim_{n\to\infty} f(x_n) = f(a)$ for alle følger $\{x_n\}$ i X som konvergerer mot a.

Bevis
I prinsippet likt beviset for Setning 9.4.1. □

Den andre er åpne-mengder-kriteriet fra Oppgave 9.5:

Setning 10.4.5: Åpne-mengder-kriteriet for kontinuitet

La $f : X \to Y$ være en funksjon mellom to metriske rom.

Funksjonen f er kontinuerlig i $a \in X$ hvis og bare hvis det for enhver åpen mengde $V \subset Y$ slik at $f(a) \in V$ finnes en åpen mengde $U \subset X$ slik at $a \in U \subset f^{-1}(V)$.

Funksjonen f er kontinuerlig hvis og bare hvis $f^{-1}(V) \subset X$ er åpen for alle åpne mengder $V \subset Y$.

Bevis
I prinsippet likt beviset utført i Oppgave 9.5. Vi overlater detaljene til Oppgave 10.28. □

Vi har fremdeles et resultat som sier at *sammensetningen av to kontinuerlige funksjoner er kontinuerlig* (jf. Setning 9.5.2).

Setning 10.4.6: Sammensetning av kontinuerlige funksjoner

Anta $g : X \to Y$ og $f : Y \to Z$ er funksjoner mellom metriske rom slik at g er kontinuerlig i punktet $a \in X$ og f er kontinuerlig i punktet $g(a) \in Y$. Da er den sammensatte funksjonen $f \circ g : X \to Z$ kontinuerlig i a.

Bevis
I prinsippet helt likt beviset for Setning 9.5.2, ved å bruke *følge-kriteriet* (Setning 10.4.4). Man kan også bruke *åpne-mengder-kriteriet* (Setning 10.4.5), jf. Oppgave 10.29. □

Funksjoner med verdimengde i euklidske rom. I tilfellet med en funksjon $\mathbf{f} : X \to \mathbb{R}^n$, der X er et metrisk rom med metrikk d_X og $\mathbb{R}^n$ er som vanlig utstyrt med den euklidske metrikken $d_{\mathbb{R}^n}$, har vi noen flere resultater til rådighet, som generaliserer resultater vi har sett tidligere.

Vi merker først at en slik funksjon kan beskrives ved hjelp av koordinatfunksjonene: vi kan nemlig skrive, for hver $x \in X$,

$$\mathbf{f}(x) = (f_1(x), \dots, f_n(x)) \in \mathbb{R}^n$$

og hver $x \mapsto f_i(x)$ definerer en funksjon $f_i : X \to \mathbb{R}$, for $i \in \{1, \dots, n\}$. Følgende resultat er veldig hendig og viser at kontinuiteten til $\mathbf{f}$ er ensbetydende med kontinuiteten til hver f_i:

Setning 10.4.7

La X være et metrisk rom og

$$\begin{aligned} \mathbf{f} : X &\longrightarrow \mathbb{R}^n \\ x &\mapsto (f_1(x), \dots, f_n(x)) \end{aligned}$$

en funksjon.

Da er $\mathbf{f}$ kontinuerlig i et punkt $a \in X$ hvis og bare hvis hver funksjon $f_i : X \to \mathbb{R}$ er kontinuerlig i a.

Bevis

Overlates til Oppgave 10.34. □

Vi kan derfor redusere oss til tilfellet $n = 1$ og studere kontinuiteten til funksjoner $\mathbf{f} : X \to \mathbb{R}$. I dette tilfellet har vi en ordrett generalisering av Setning 9.5.1:

Setning 10.4.8: Kombinasjoner av kontinuerlige funksjoner

Anta X er et metrisk rom og $f : X \to \mathbb{R}$ og $g : X \to \mathbb{R}$ er kontinuerlige i punktet a. La $c \in \mathbb{R}$. Da er også funksjonene $f + g$, $f - g$, cf og $f \cdot g$ kontinuerlige i a. Hvis i tillegg $g(a) \neq 0$, da er også $\frac{f}{g}$ kontinuerlig i a.

Bevis

Likt beviset for Setning 9.5.1, ved å bruke Setning 10.4.4. □

Eksempel 10.4.9. For hver $i \in \{1, \dots, n\}$, definerer vi *den i-te koordinatfunksjonen* $\phi_i : \mathbb{R}^n \to \mathbb{R}$ ved $\phi_i((x_1, \dots, x_n)) = x_i$. Da er ϕ_i kontinuerlig: for hvis $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$ og $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, da er

$$|\phi_i(\mathbf{x}) - \phi_i(\mathbf{a})| = |x_i - a_i| \leq \sqrt{\sum_{j=1}^{n} |x_j - a_j|^2} = d_{\mathbb{R}^n}(\mathbf{x}, \mathbf{a}),$$

slik at det holder å ta $\delta = \epsilon$ i definisjonen av kontinuitet. (For øvrig viser dette at ϕ_i er en *Lipschitz-funksjon*, jf. Definisjon 9.9.18.)

Ved å anvende Setning 10.4.8 gjentagende ganger får vi at alle funksjoner på formen

$$\begin{array}{rcl} \mathbf{f} : \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ \mathbf{x} = (x_1, \ldots, x_n) & \mapsto & \sum c_{k_1 \ldots k_n} x_1^{k_1} \cdots x_n^{k_n}, \end{array}$$

med alle $c_{k_1 \ldots k_n} \in \mathbb{R}$, er kontinuerlige. Uttrykket

$$P(x_1, \ldots, x_n) = \sum c_{k_1 \ldots k_n} x_1^{k_1} \cdots x_n^{k_n}$$

kalles et *(reellt) polynom av n variable.* Setning 10.4.8 gir også at funksjoner definert ved *rasjonale funksjoner*, det vi si på formen $\mathbf{x} \mapsto \frac{P(\mathbf{x})}{Q(\mathbf{x})}$, hvor P og Q er polynomer, er kontinuerlige på sin definisjonsmengde, som er alle punkter i $\mathbb{R}^n$ hvor Q ikke er 0.

La oss se på noen eksempler på bruk av disse resultatene.

Eksempel 10.4.10. Begrunn at funksjonen $f : \mathbb{R}^3 \setminus \{0,0,0\} \to \mathbb{R}^2$ gitt ved

$$f(x, y, z) = \left(e^{xy^2}, \frac{x}{\sqrt{x^2 + y^2 + z^2}} \right)$$

er kontinuerlig.

Løsning. Ved Setning 10.4.7 må vi vise at hver av funksjonene

$$f_1, f_2 : \mathbb{R}^3 \setminus \{0,0,0\} \to \mathbb{R}$$

gitt ved

$$f_1(x, y, z) = e^{xy^2}, \quad f_2(x, y, z) = \frac{x}{\sqrt{x^2 + y^2 + z^2}}$$

er kontinuerlige.

Funksjonen f_1 er sammensetningen av funksjonene

$$\begin{array}{rcl} \mathbb{R} & \longrightarrow & \mathbb{R} \\ t & \mapsto & e^t, \end{array}$$

som er kontinuerlig ved Setning 9.5.11, og

$$\begin{array}{rcl} \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ (x, y, z) & \mapsto & xy^2, \end{array}$$

som er kontinuerlig, siden den er gitt ved et polynom (Eksempel 10.4.9). Dermed er f_1 kontinuerlig ved Setning 10.4.6.

Funksjonen

(295)
$$\begin{array}{rcl} \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ (x, y, z) & \mapsto & \sqrt{x^2 + y^2 + z^2} \end{array}$$

er igjen kontinuerlig ved Setning 10.4.6, fordi den er sammensetningen av funksjonene

$$\begin{array}{ccc} [0,\infty) & \longrightarrow & \mathbb{R} \\ t & \mapsto & \sqrt{t}, \end{array}$$

som er kontinuerlig ved Følgesetning 9.5.9, og

$$\begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & [0,\infty) \\ (x,y,z) & \mapsto & x^2+y^2+z^2, \end{array}$$

som igjen er kontinuerlig siden den er gitt ved et polynom. Likeledes er funksjonen

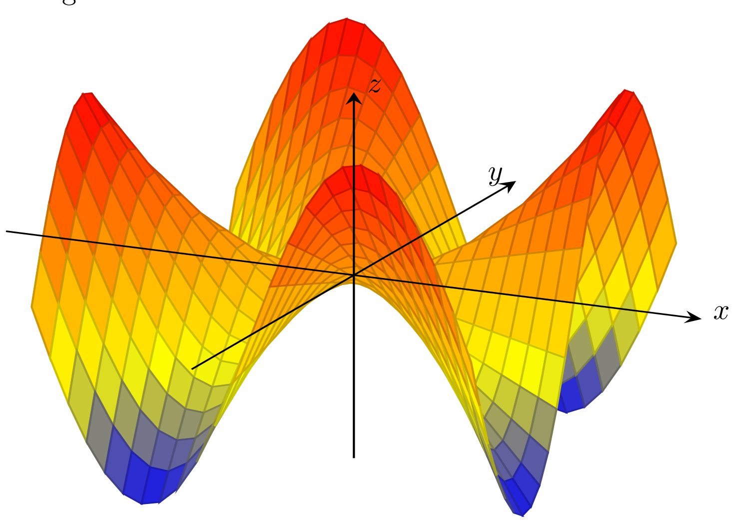
$$\begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ (x,y,z) & \mapsto & x \end{array} \tag{296}$$

kontinuerlig. Funksjonen f_2 er gitt som en kvotient av de to kontinuerlige funksjonene (295) og (296), og er derfor kontinuerlig (på sin definisjonsmengde) ved Setning 10.4.8. $\square$

Eksempel 10.4.11. Avgjør om funksjonen

$$\begin{array}{ccc} f:\mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \mapsto & \begin{cases} \frac{2xy(x^2-y^2)}{x^2+y^2}, & \text{hvis } (x,y)\neq(0,0) \\ 0, & \text{hvis } (x,y)=(0,0) \end{cases} \end{array}$$

er kontinuerlig.



Grafen til funksjonen i Eksempel 10.4.11

LØSNING. Utenom $(0,0)$ er funksjonen kontinuerlig siden den er gitt ved en rasjonal funksjon (Eksempel 10.4.9). Vi må derfor undersøke kontinuiteten i origo. Husk at med de metrikkene vi bruker er, for $(x,y) \in \mathbb{R}^2$:

$$|(x,y) - (0,0)| = \sqrt{x^2 + y^2}$$

og

$$|f(x,y) - f(0,0)| = |f(x,y)|\,.$$

Vi viser kontinuiteten på to måter, ved å bruke ϵ–δ definisjonen 10.4.1, og ved å bruke følge-kriteriet 10.4.4.

Bevis ved ϵ–δ definisjonen: For å vise at f er kontinuerlig i $(0,0)$ må vi finne, for en vilkårlig $\epsilon > 0$, en $\delta > 0$ slik at

$$|f(x,y)| < \epsilon \text{ når } \sqrt{x^2+y^2} < \delta. \tag{297}$$

Vi har, for $(x,y) \neq (0,0)$:

$$\begin{aligned}
\Big|f(x,y)\Big| &= \Big|\frac{2xy(x^2-y^2)}{x^2+y^2}\Big| = \Big|\frac{2x^3y}{x^2+y^2} - \frac{2xy^3}{x^2+y^2}\Big| \\
&\leq \Big|\frac{2x^3y}{x^2+y^2}\Big| + \Big|\frac{2xy^3}{x^2+y^2}\Big| \\
&\leq \Big|\frac{2x^3y}{x^2}\Big| + \Big|\frac{2xy^3}{y^2}\Big| \\
&= |2xy| + |2xy| = |4xy| = 4|x||y|.
\end{aligned}$$

Den utledede ulikheten $\Big|f(x,y)\Big| \leq 4|x||y|$ gjelder også for $(x,y) = (0,0)$, siden $f(0,0) = 0$.

Vi ser at $4|x||y| < \epsilon$ hvis $|x| < \frac{1}{2}\sqrt{\epsilon}$ og $|y| < \frac{1}{2}\sqrt{\epsilon}$, og dette siste er oppfylt hvis og bare hvis $x^2 < \frac{1}{4}\epsilon$ og $y^2 < \frac{1}{4}\epsilon$, som er igjen oppfylt hvis for eksempel $x^2 + y^2 < \frac{1}{4}\epsilon$. Vi lar derfor $\delta = \frac{1}{2}\sqrt{\epsilon}$.

Vi oppsummerer: hvis $\sqrt{x^2+y^2} < \delta = \frac{1}{2}\sqrt{\epsilon}$, er

$$|x| = \sqrt{x^2} \leq \sqrt{x^2+y^2} < \frac{1}{2}\sqrt{\epsilon},$$

og likeledes $|y| < \frac{1}{2}\sqrt{\epsilon}$. Da vil

$$\Big|f(x,y)\Big| \leq 4|x||y| < 4 \cdot \frac{1}{2}\sqrt{\epsilon} \cdot \frac{1}{2}\sqrt{\epsilon} = \epsilon,$$

og vi har vist (297). Altså er f kontinuerlig i $(0,0)$, og dermed overalt.

Bevis ved følge-kriteriet: For å vise at f er kontinuerlig i $(0,0)$, må vi vise at for en vilkårlig følge $\{(x_n, y_n)\}$ i $\mathbb{R}^2$ som konvergerer mot $(0,0)$, vil følgen $\{f(x_n, y_n)\}$ konvergere mot $f(0,0) = 0$.

Ved å bruke polare koordinater, kan vi skrive

$$(x_n, y_n) = (r_n \cos\theta_n, r_n \sin\theta_n),$$

med $r_n = \sqrt{x_n^2 + y_n^2} \in \mathbb{R}^{\geq 0}$ og $\theta_n \in [0, 2\pi)$. At $(x_n, y_n) \to (0,0)$ er ekvivalent med at

$$|(x_n, y_n) - (0,0)| = \sqrt{x_n^2 + y_n^2} = r_n \to 0.$$

For alle $(x_n, y_n) \neq (0,0)$, er $r_n > 0$, og

$$\begin{aligned} f(x_n, y_n) &= f(r_n \cos\theta_n, r_n \sin\theta_n) \\ &= \frac{2(r_n\cos\theta_n)(r_n\sin\theta_n)(r_n^2\cos^2\theta_n - r_n^2\sin^2\theta_n)}{r_n^2\cos^2\theta_n + r_n^2\sin^2\theta_n} \\ &= \frac{2r_n^4\cos\theta_n\sin\theta_n(\cos^2\theta_n - \sin^2\theta_n)}{r_n^2} \\ &= 2r_n^2\cos\theta_n\sin\theta_n(\cos^2\theta_n - \sin^2\theta_n). \end{aligned}$$

For $(x_n, y_n) = (0,0)$, er $r_n = 0$ og $f(x_n, y_n) = 0$, som kan uttrykkes ved samme formel. Altså har vi at

$$f(x_n, y_n) = 2r_n^2\cos\theta_n\sin\theta_n(\cos^2\theta_n - \sin^2\theta_n)$$

for alle $n \in \mathbb{Z}^+$. Vi har

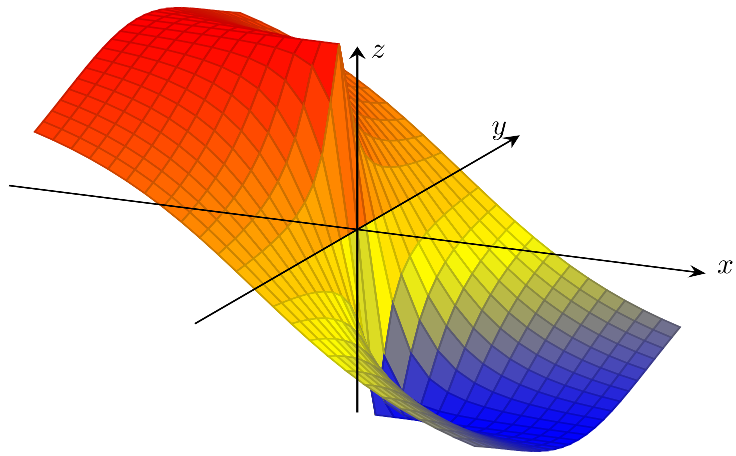
$$\begin{aligned} |f(x_n, y_n)| &= 2r_n^2\,|\cos\theta_n|\,|\sin\theta_n|\,\left|\cos^2\theta_n - \sin^2\theta_n\right| \\ &\leq 2r_n^2\,|\cos\theta_n|\,|\sin\theta_n|\,\left|\cos^2\theta_n + \sin^2\theta_n\right| \\ &\leq 2r_n^2 \cdot 1 \cdot 1 \cdot 1 = 2r_n^2, \end{aligned}$$

og siden $\lim\limits_{n\to\infty} r_n = 0$, vil også $\lim\limits_{n\to\infty} f(x_n, y_n) = 0$ (ved *skvisesetningen* 7.4.13), som er det vi ønsket å vise. $\square$

Eksempel 10.4.12. Vis at det ikke finnes noen $c \in \mathbb{R}$ som gjør funksjonen

$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x,y) &\mapsto \begin{cases} -\frac{x}{\sqrt{x^2+y^2}}, & \text{hvis } (x,y) \neq (0,0) \\ c, & \text{hvis } (x,y) = (0,0) \end{cases} \end{aligned}$$

kontinuerlig i $(0,0)$.



GRAFEN TIL FUNKSJONEN I EKSEMPEL 10.4.12

LØSNING. Vi argumenterer ved selvmotsigelse og antar at en slik c som gjør f kontinuerlig i origo eksisterer. Vi bruker følge-kriteriet i Setning 10.4.4. Vi velger en følge $\{(x_n, y_n)\}$ i $\mathbb{R}^2$ som konvergerer mot $(0,0)$ langs den rette linjen $y = kx$, for en fiksert $k \in \mathbb{R}$. Følgen $\{(x_n, y_n)\} = \left\{\left(\frac{1}{n}, \frac{k}{n}\right)\right\}$ oppfyller dette ved Setning 10.2.9, siden både $\frac{1}{n} \to 0$ og $\frac{k}{n} \to 0$. Vi har

$$f(x_n, y_n) = -\frac{\frac{1}{n}}{\sqrt{\left(\frac{1}{n}\right)^2 + \left(\frac{k}{n}\right)^2}} = -\frac{1}{\sqrt{1+k^2}},$$

som er en konstant uavhengig av n. Dermed vil $f(x_n, y_n) \to -\frac{1}{\sqrt{1+k^2}}$. Setning 10.4.4 gir dermed at $c = -\frac{1}{\sqrt{1+k^2}}$. Men dette gjelder for vilkårlig k, slik at c ikke kan eksistere. □

Løsningen viser en vanlig prosedyre: For å vise at en funksjon *ikke* er kontinuerlig i et bestemt punkt a, holder det ved følge-kriteriet i Setning 10.4.4 å finne en følge $\{(x_n, y_n)\}$ som konvergerer mot a slik at følgen $\{f(x_n, y_n)\}$ *ikke* konvergerer mot $f(a)$. I løsningen vi nettopp så fant vi sogar uendelig mange slike følger. For å finne slike følger når definisjonsmengden er $\mathbb{R}^2$ kan det i mange tilfeller være lurt å studere følger som konvergerer mot a *langs en bestemt kurve i planet*. De enkleste kurvene er rette linjer, som fungerte bra i dette tilfellet. Det er imidlertid ikke alltid man har så mye flaks at rette linjer fungerer, og da kan sjekke langs parabler (dersom a er origo vil det si $y = ky^2$ for varierende k), tredjegradskurver (dersom a er origo vil det si $y = ky^3$ for varierende k), osv. (For et eksempel på et slikt resonnement, se Oppgave 10.33.) Sjekker vi langs mange kurver og f av følgene alltid har grense $f(a)$, da kan det imidlertid være et tegn på at funksjonen faktisk *er* kontinuerlig i a.

Vi viser også hvordan vi kan uttrykke følger ved polare koordinater for å løse problemet:

Alternativ løsning. Vi argumenterer ved selvmotsigelse og antar at en slik c som gjør f kontinuerlig i origo eksisterer. Ved følge-kriteriet i Setning 10.4.4 har vi at for en vilkårlig følge $\{(x_n, y_n)\}$ i $\mathbb{R}^2$ slik at $(x_n, y_n) \to (0,0)$, vil $f(x_n, y_n) \to f(0,0) = c$. Vi kan skrive

$$(x_n, y_n) = (r_n \cos\theta_n, r_n \sin\theta_n),$$

med $r_n = \sqrt{x_n^2 + y_n^2} \in \mathbb{R}^{\geq 0}$ og $\theta_n \in [0, 2\pi)$. At $(x_n, y_n) \to 0$ er ekvivalent med at

$$|(x_n, y_n) - (0,0)| = \sqrt{x_n^2 + y_n^2} = r_n \to 0.$$

Hvis $(x_n, y_n) \neq (0,0)$, er

$$f(x_n, y_n) = -\frac{x_n}{\sqrt{x_n^2 + y_n^2}} = -\frac{r_n \cos\theta_n}{r_n} = -\cos\theta_n.$$

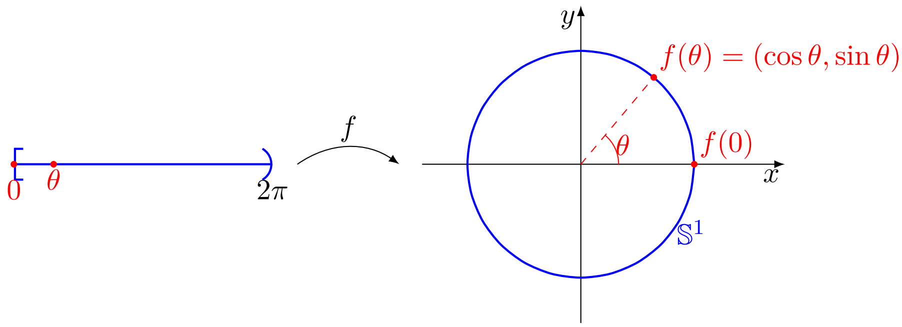
Grensen til denne følgen avhenger helt av grensen til følgen $\{\theta_n\}$ og kan derfor ikke være lik c for alle følger $\{(x_n, y_n)\}$. Hvis for eksempel $r_n = \frac{1}{n}$ og $\theta_n = 0$, som tilsvarer $(x_n, y_n) = \left(\frac{1}{n}, 0\right)$, vil $\lim\limits_{n\to\infty} f(x_n, y_n) = \lim\limits_{n\to\infty} \theta_n = 0$. Hvis for eksempel $r_n = \frac{1}{n}$ og $\theta_n = \frac{\pi}{2}$, som tilsvarer $(x_n, y_n) = \left(0, \frac{1}{n}\right)$, vil $\lim\limits_{n\to\infty} f(x_n, y_n) = \lim\limits_{n\to\infty} \theta_n = \dfrac{\pi}{2}$. Dette viser at f ikke er kontinuerlig i origo, uansett c. $\square$

Eksempel på en kontinuerlig bijektiv funksjon med diskontinuerlig invers. Gitt en injektiv kontinuerlig funksjon $f : X \to Y$ mellom metriske rom, er det naturlig å spørre seg om den inverse funksjonen $f^{-1} : f(X) \to X$ er kontinuerlig. Dersom X og Y er delmengder av $\mathbb{R}$, så vi i Følgesetning 9.6.14 at spørsmålet har et positivt svar når X er et intervall, mens vi så i Merknad 9.5.5 at svaret kan være negativt når X ikke er et intervall. Det er derfor naturlig å spørre seg om svaret alltid kan være positivt for vilkårlig Y dersom X er et intervall i $\mathbb{R}$. Følgende berømte eksempel viser at dette er for optimistisk:

Eksempel 10.4.13. La $\mathbb{S}^1$ være *enhetssirkelen*, med andre ord delmengden av $\mathbb{R}^2$ bestående av alle punkter i avstand 1 fra origo, betraktet som et underrom av $\mathbb{R}^2$. Se på den bijektive funksjonen

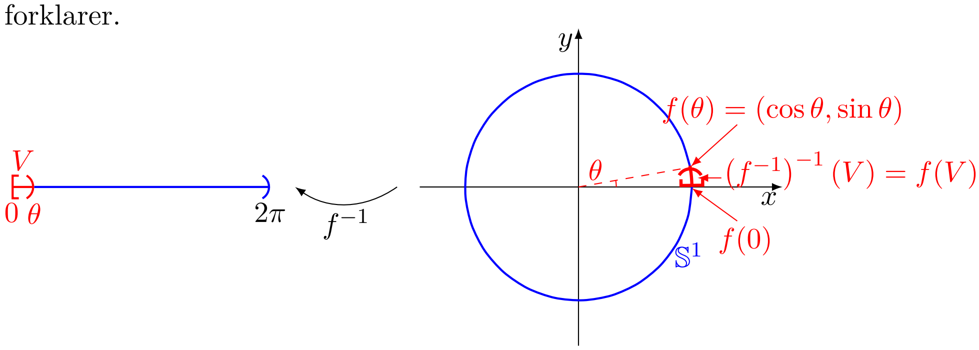
$$f : [0, 2\pi) \to \mathbb{S}^1, \quad f(\theta) = (\cos\theta, \sin\theta),$$

betraktet i Eksempel 2.3.13 visualisert i følgende figur:



Denne funksjonen er kontinuerlig ved Følgesetning 9.5.10 og Setning 10.4.7.

Den inverse funksjonen f^{-1} sender punktet $(\cos\theta, \sin\theta)$ til θ som vist i figuren nedenunder. Denne funksjonen er ikke kontinuerlig i 0, som vi nå forklarer.



La $\theta \in (0, 2\pi)$ og betrakt delmengden $V = [0, \theta)$ av $[0, 2\pi)$, merket i rødt i figuren ovenfor. Mengden V er åpen i $[0, 2\pi)$ ved Setning 10.3.6, siden vi for eksempel kan skrive $V = [0, 2\pi) \cap (-1, \theta)$ og $(-1, \theta)$ er åpen i $\mathbb{R}$. Inversbildet $\left(f^{-1}\right)^{-1}(V)$ til V ved f^{-1} i $\mathbb{S}^1$ er delen av sirkelbuen "fra og med $f(0) = (1, 0)$" og "til men ikke med $f(\theta)$", merket igjen i rødt i figuren ovenfor. (Merk at $\left(f^{-1}\right)^{-1}(V) = f(V)$.) Denne mengden er *ikke* åpen i $\mathbb{S}^1$. For hvis den var det, ville vi igjen ved Setning 10.3.6 kunne skrive

$$f(V) = W \cap \mathbb{S}^1 \tag{298}$$

for en åpen delmengde $W \subset \mathbb{R}^2$ som må inneholde $f(0)$. Men enhver slik åpen $W \subset \mathbb{R}^2$ må nødvendigvis inneholde punkter på $\mathbb{S}^1$ *under* x-aksen, slik at (298) er umulig.

Merknad 10.4.14. 9.5.5 En kontinuerlig bijektiv funksjon mellom metriske rom (eller topologiske rom, jf. §10.10) som har en kontinuerlig invers kalles en *bikontinuerlig funksjon* eller *homeomorfi*. Eksemplet ovenfor viser at ikke alle kontinuerlige bijektive funksjoner er bikontinuerlige. Se Oppgave 10.50 for en tilstrekkelig betingelse på definisjonsmengden som garanterer at en kontinuerlig bijektiv funksjon er bikontinuerlig.

10.5. Sammenhengende mengder; generalisering av skjæringssetning*

Det er naturlig å spørre seg om de viktige resultatene *skjæringssetningen* (Teorem 9.6.1) og *ekstremalverdisetningen* (Teorem 9.7.2) for kontinuerlige funksjoner har generaliseringer til vilkårlige metriske rom. Det kan ikke finnes helt tilsvarende formuleringer siden et vilkårlig metrisk rom ikke er ordnet. Vi vil likevel se at det finnes visse generaliseringer. Vi starter med *skjæringssetningen* i denne seksjonen.

Vi husker konsekvensen av *skjæringssetningen* gitt i Følgesetning 9.6.2, som sier at bildet av et intervall ved en kontinuerlig funksjon $f : D(f) \to \mathbb{R}$, med $D(f) \subset \mathbb{R}$, er igjen et intervall. Den naturlige generaliseringen av begrepet "intervall" til vilkårlige metriske rom viser seg å være begrepet "sammenhengende":

Definisjon 10.5.1: Sammenhengende og ikke-sammenhengende

Et metrisk rom X sies å være *ikke-sammenhengende*[a] dersom vi kan skrive $X = A \cup B$, der A og B er ikke-tomme åpne disjunkte delmengder av X.

Ellers sies X å være *sammenhengende*[b].

Ekvivalent er X sammenhengdende dersom de eneste delmengdene av X som er både åpne og lukket i X er X og $\emptyset$.

[a] *Disconnected* på engelsk.
[b] *Connected* på engelsk.

Husk at det at to mengder A og B er *disjunkte* betyr at $A \cap B = \emptyset$ (jf. §2.1).

Vi overlater til Oppgave 10.41 å vise at de to definisjonene av "sammenhengende" er ekvivalente og konsentrerer oss her på noen eksempler. Det sentrale eksemplet er:

Eksempel 10.5.2. En delmengde X av $\mathbb{R}$ er sammenhengende hvis og bare hvis X er et intervall. Vi overlater det til Oppgave 10.42 å vise dette.

Eksempel 10.5.3. Det metriske rommet $X = [0, 1) \cup (1, 2]$ (med metrikken indusert av $\mathbb{R}$) er ikke-sammenhengende, siden $[0, 1) \cap (1, 2] = \emptyset$ og både $[0, 1)$ og $(1, 2]$ er åpne delmengder av X, ved Setning 10.3.6, siden $[0, 1) = X \cap (-1, 1)$ og $(1, 2] = X \cap (1, 3)$.

Merk at samme resonnement ikke fungerer på $X = [0, 2] = [0, 1] \cup (1, 2]$, siden $[0, 1]$ ikke er åpen i X.

EKSEMPEL 10.5.4. Mengden av rasjonale tall $\mathbb{Q}$ er ikke-sammenhengende. Dette følger fra Eksempel 10.5.2, men vi kan også vise det direkte: For eksempel kan vi skrive

$$\mathbb{Q} = \left[(-\infty, \sqrt{2}) \cap \mathbb{Q}\right] \cup \left[(\sqrt{2}, \infty) \cap \mathbb{Q}\right],$$

og disse to delmengdene er selvsagt disjunkte, og begge er åpne i $\mathbb{Q}$, ved Setning 10.3.6.

EKSEMPEL 10.5.5. Det euklidske rommet $\mathbb{R}^n$ er sammenhengende. La oss vise dette, siden beviset er ganske interessant og gir en fin anvendelse av *åpne-mengder-kriteriet for kontinuitet* (Setning 10.4.5).

Anta, for å få en selvmotsigelse, at $\mathbb{R}^n$ kan skrives som $\mathbb{R}^n = A \cup B$, med A og B åpne og ikke-tomme og $A \cap B = \emptyset$. Velg punkter $\mathbf{x} \in A$ og $\mathbf{y} \in B$. Betrakt funksjonen

$$\begin{aligned} s : [0,1] &\longrightarrow \mathbb{R}^n \\ t &\mapsto t\mathbf{y} + (1-t)\mathbf{x}, \end{aligned}$$

som er kontinuerlig ved Setning 10.4.8. Merk at $s(0) = \mathbf{x}$ og $s(1) = \mathbf{y}$ og at verdimengden er

$$\{\mathbf{z} \in \mathbb{R}^n \mid \mathbf{z} = t\mathbf{y} + (1-t)\mathbf{x} \text{ for en } t \in [0,1]\},$$

som ikke er annet enn det rette linjestykket mellom $\mathbf{x}$ og $\mathbf{y}$. Vi har

$$[0,1] = s^{-1}(\mathbb{R}^n) = s^{-1}(A \cup B) \overset{(*)}{=} s^{-1}(A) \cup s^{-1}(B)$$

(hvor (*) følger fra Oppgave 2.16(e)). Siden f er kontinuerlig, er $s^{-1}(A)$ og $s^{-1}(B)$ åpne i $[0,1]$ ved *åpne-mengder-kriteriet for kontinuitet* (Setning 10.4.5). Siden $s(0) = \mathbf{x} \in A$ og $s(1) = \mathbf{y} \in B$, er $0 \in s^{-1}(A)$ og $1 \in s^{-1}(B)$, slik at $s^{-1}(A) \neq \emptyset$ og $s^{-1}(B) \neq \emptyset$. Siden $A \cap B = \emptyset$, er $s^{-1}(A) \cap s^{-1}(B) = \emptyset$. Dette viser at $[0,1]$ er ikke-sammenhengende, som motsier Eksempel 10.5.2.

Se Oppgave 10.43 for en generalisering av dette eksemplet.

EKSEMPEL 10.5.6. Velg en linje $\ell \subset \mathbb{R}^2$. Da er $X = \mathbb{R}^2 \setminus \ell$ ikke-sammenhengende, siden den kan skrives som en union av to disjunkte halvplan, som begge er åpne i $\mathbb{R}^2$ og dermed også i X.

EKSEMPEL 10.5.7. La X være et metrisk rom, og A og B to disjunkte, ikke-tomme åpne delmengder av X. Da er A og B også åpne i $A \cup B$ ved Setning 10.3.6. Dette viser at $A \cup B$ er ikke-sammenhengende.

Den naturlige generaliseringen av *skjæringssetningen* er følgende generalisering av Følgesetning 9.6.2:

Setning 10.5.8

La $f : X \to Y$ være en kontinuerlig funksjon mellom metriske rom. Hvis X er sammenhengende, da er også $f(X) \subset Y$ sammenhengende.

Vi uttrykker dette gjerne som at *en kontinuerlig funksjon bevarer sammenhengende mengder.*

Bevis for Setning 10.5.8
Avbildningen $f : X \to f(X)$ er surjektiv og kontinuerlig ved Merknad 10.4.2. Vi vil vise at $f(X)$ er sammenhengende. Anta at det ikke er sant, slik at vi kan skrive $f(X) = A \cup B$ med A og B ikke-tomme, åpne i $f(X)$ og $A \cap B = \emptyset$. Da er $f^{-1}(A)$ og $f^{-1}(B)$ åpne i X, siden f er kontinuerlig (Setning 10.4.5), og begge er ikke-tomme, siden funksjonen $f : X \to f(X)$ er surjektiv. I tillegg er det lett å verifisere at $X = f^{-1}(A) \cup f^{-1}(B)$ og $f^{-1}(A) \cap f^{-1}(B) = \emptyset$. Men dette motsier at X er sammenhengende. $\square$

Som konsekvens får vi et resultat som passer med den intuitive forståelsen av kontinuitet vi har fra skolen:

Følgesetning 10.5.9

La $f : I \to \mathbb{R}$ være en funksjon mellom metriske rom, med $I \subset \mathbb{R}$ et intervall. Hvis f er kontinuerlig, da er grafen $\Gamma(f) \subset \mathbb{R}^2$ til f sammenhengende.

Bevis
Betrakt funksjonen $F : I \to \mathbb{R}^2$ gitt ved $F(x) = (x, f(x))$. Siden funksjonene $x \mapsto x$ og $x \mapsto f(x)$ er kontinuerlige, er F kontinuerlig ved Setning 10.4.7. Intervallet I er sammenhengende ved Eksempel 10.5.2. Verdimengden til F er per definisjon lik grafen $\Gamma(f)$, som dermed er sammenhengende ved Setning 10.5.8. $\square$

Det er imidlertid interessant å merke seg at det motsatte av Følgesetning 10.5.9 *ikke* gjelder: en diskontinuerlig reell funksjon definert på et intervall kan også ha en sammenhengende graf, som følgende berømte eksempel viser. Derfor er ikke “sammenhengsegenskapen til grafen til en funksjon” en presis nok beskrivelse av kontinuitet.

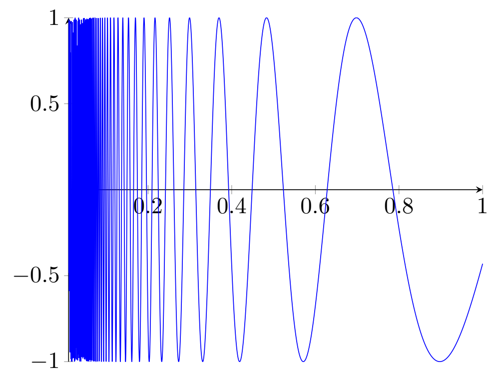
Eksempel 10.5.10. Betrakt funksjonen

$$\begin{aligned} f : [0,1] &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} \sin\left(\frac{1}{x}\right), & x \in (0,1] \\ 0, & x = 0, \end{cases} \end{aligned}$$

som er diskontinuerlig i 0 ved Eksempel 9.3.1. Grafen til f er lik

$$\Gamma(f) = \left\{ \left(x, \sin\left(\frac{1}{x}\right) \right) \mid x \in (0,1] \right\} \cup \left\{ (0,0) \right\} \subset \mathbb{R}^2,$$

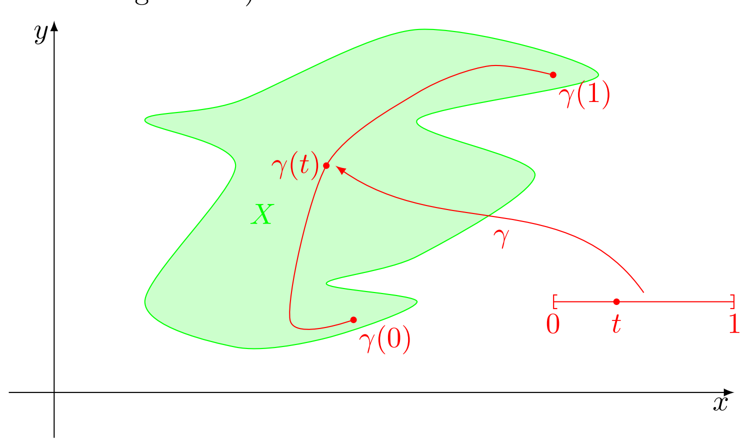
som vist i figuren nedenunder:



Topologenes sinuskurve i $\mathbb{R}^2$.

Vi overlater til Oppgave 10.45(a) å vise at $\Gamma(f)$ er sammenhengende. Mengden kalles *topologenes sinuskurve*[12] eller *Warsaw-kurven.*

Merknad 10.5.11. En nærliggende egenskap til å være sammenhengende er å være *kurvesammenhengende*[13]. Intuitivt er et metrisk rom X kurvesammenhengende dersom hvert par av punkter kan bindes sammen med en kontinuerlig kurve. Mer presist: Et metrisk rom X er *kurvesammenhengende* dersom det for ethvert par av punkter $x, y \in X$, finnes en kontinuerlig funksjon $\gamma : [0,1] \to X$ slik at $\gamma(0) = x$ og $\gamma(1) = y$. (Her er γ den lille greske bokstaven "gamma".)



En kurvesammenhengende delmengde X av $\mathbb{R}^2$.

[12] *The topologist's sine curve* på engelsk.
[13] *Path connected* på engelsk.

Ved å argumentere som i Eksempel 10.5.5 kan man vise at *et kurvesammenhengende metrisk rom er sammenhengende*, jf. Oppgave 10.43(b), som er et hendig kriterium å bruke for å vise at et metrisk rom er sammenhengende. Men det motsatte er ikke sant: *topologenes sinuskurve*[14] fra siste eksempel er *ikke* kurvesammenhengende, men sammenhengende. (At den ikke er kurvesammenhengende er ikke så lett å vise, så vi vil ikke gå inn på det her.)

En variant fremkommer hvis vi legger til intervallet $[-1,1]$ på y-aksen, slik at vi får mengden

$$\left\{\left(x, \sin\left(\frac{1}{x}\right)\right) \mid x \in (0,1]\right\} \cup \left\{(0,y) \mid y \in [-1,1]\right\}.$$

Denne er lukket og fremdeles sammenhengende (jf. Oppgave 10.45(b)-(c)), men ikke kurvesammenhengende. Mengden kalles *topologenes lukkede sinuskurve*[15].

10.6. Kompakte mengder; generalisering av ekstremalverdisetning*

Vi kan ikke snakke om maksimum og minimum til funksjoner i vilkårlige metriske rom, siden de generelt ikke er ordnede mengder. Husk imidlertid konsekvensen av *ekstremalverdisetningen* (og *skjæringssetningen*) gitt i Følgesetning 9.7.10. Den sier at dersom $[a,b] \subset \mathbb{R}$ er et lukket, begrenset intervall og $f : [a,b] \to \mathbb{R}$ er kontinuerlig, da er verdimengden $f([a,b])$ fremdeles et lukket, begrenset intervall. Denne versjonen har sin umiddelbare generalisering til euklidske rom:

Setning 10.6.1

La $f : X \to \mathbb{R}^m$ være en kontinuerlig funksjon, der $X \subset \mathbb{R}^n$ er en lukket, begrenset delmengde (med den euklidske metrikken). Da er $f(X)$ også lukket og begrenset.

BEVIS
Dette er en umiddelbar konsekvens av Setningene 10.6.8 og 10.6.11 som vi kommer til om litt. □

Vi sier gjerne at *en kontinuerlig funksjon mellom euklidske rom bevarer lukkede, begrensede mengder.*

I spesialtilfellet $m = 1$, betyr dette at $f : X \to \mathbb{R}$ oppnår minimal- og maksimalverdi.

Setningen er ikke sann dersom vi erstatter X med et vilkårlig lukket og begrenset metrisk rom. Et moteksempel gis i Eksempel 10.7.3 nedenunder, et annet er:

[14] *The topologist's sine curve* på engelsk.
[15] *The closed topologist's sine curve* på engelsk.

EKSEMPEL 10.6.2. La $X = (0, 1]$ med metrikken indusert fra $\mathbb{R}$. Da er X begrenset, siden $|x - y| \leq 1$ for alle $x, y \in X$. Samtidig er X lukket (i seg selv, per definisjon av lukket). Betrakt funksjonen $f : X \to \mathbb{R}$ gitt ved $f(x) = \frac{1}{x}$, som er kontinuerlig. Da er $f(X) = [1, \infty)$, som ikke er begrenset.

Merk at dette ikke motsier Setning 10.6.1: betrakter vi $X \subset \mathbb{R}$, da er X ikke lukket i $\mathbb{R}$, slik at setningen ikke kan anvendes.

Kjernepunktet i generaliseringen av *ekstremalverdisetningen* til vilkårlige metriske rom er å finne egenskaper på generelle metriske rom som kan erstatte egenskapen "lukket og begrenset", men som er ekvivalente med sistnevnte egenskap i euklidske rom. Flere arbeider i siste halvdel av 1800-tallet, av blant annet de tyske matematikerne Karl Weierstrass (1815–1897) og Eduard Heine (1821–1881), de franske matematikerne Émile Borel (1871–1956) og Henri Lebesgue (1875 –1941) og de italienske matematikerne Giulio Ascoli (1843–1896) og Cesare Arzelà (1847–1912) var nettopp opptatt av dette. En sentralt rolle spilte begrepet *delfølger* (Definisjon 7.7.2) og karakteriseringen av lukkede, begrensede intervaller som vi allerede har sett i Følgesetning 7.7.8 av *Bolzano-Weierstrass teoremet* 7.7.5: *Et intervall $I \subset \mathbb{R}$ er lukket og begrenset hvis og bare hvis enhver følge i I har en delfølge som konvergerer i I.* (Har du ikke lest §7.7, er det fornuftig å gjøre det nå.) Dette danner utgangspunktet for begrepet *kompakt mengde*, introdusert av Maurice Fréchet (1878–1973) i sin doktoravhandling fra 1906. Dette begrepet kan defineres på to ekvivalente måter (i virkelighetene enda flere), som begge er omtrent like vanlige i lærebøker:

Definisjon 10.6.3: Kompakt mengde

Et metrisk rom X sies å være *kompakt* dersom en av følgende ekvivalente betingelser er oppfylt:

(i) enhver følge i X har en delfølge som konvergerer i X;

(ii) for enhver samling $\{U_\alpha\}$ av åpne delmengder av X slik at $X = \cup_\alpha U_\alpha$, finnes endelig mange $U_{\alpha_1}, \ldots, U_{\alpha_n}$ i $\{U_\alpha\}$ slik at $X = U_{\alpha_1} \cup \cdots \cup U_{\alpha_n}$.

Vi overlater til Oppgave 10.52 å vise at betingelsene (i) og (ii) er ekvivalente, som er et ganske dypt resultat. Betingelse (ii) uttrykkes gjerne som at *enhver åpen overdekning av X har en endelig deloverdekning*.

MERKNAD 10.6.4. I motsetning til åpenhet og lukkethet er egenskapen kompakt kun avhengig av rommet X selv og ikke av hvilket rom vi betrakter X som delmengde av.

Eksempel 10.6.5. Ved karakteriseringen (i) i Definisjon 10.6.3 sier Følgesetning 7.7.8 nettopp at *et intervall $I \subset \mathbb{R}$ er kompakt hvis og bare hvis det er lukket og begrenset.*

For å bli litt bedre kjent med karakteriseringen (ii) i Definisjon 10.6.3, la oss se på et par eksempler der vi anvender denne.

Eksempel 10.6.6. Vis at intervallet $(0,1]$ ikke er kompakt ved å bruke karakteriseringen (ii) i Definisjon 10.6.3.

Bevis. Vi betrakter samlingen av åpne mengder $\left\{\left(\frac{1}{n},1\right]\right\}_{n\in\mathbb{Z}^+}$ i $(0,1]$. (At mengdene $\left(\frac{1}{n},1\right]$ er åpne i $(0,1]$ følger av Setning 10.3.6, siden $\left(\frac{1}{n},1\right] = \left(\frac{1}{n},2\right) \cap (0,1]$.) Da er $(0,1] = \cup_{n\in\mathbb{Z}^+}\left(\frac{1}{n},1\right]$. Men $(0,1]$ kan ikke være inneholdt i en union av noe endelig utvalg av mengdene, siden en slik endelig union vil være på formen $\left(\frac{1}{N},1\right]$ for en $N \in \mathbb{Z}^+$, fordi

$$\emptyset = (1,1] \subset \left(\frac{1}{2},1\right] \subset \left(\frac{1}{3},1\right] \subset \left(\frac{1}{4},1\right] \subset \cdots$$

Altså er ikke (ii) i Definisjon 10.6.3 oppfylt. □

Eksempel 10.6.7. Vis at $\mathbb{R}$ *ikke* er kompakt ved å bruke karakteriseringen (ii) i Definisjon 10.6.3.

Bevis. Vi kan skrive $\mathbb{R} = \bigcup_{n\in\mathbb{Z}}(n,n+2)$, men $\mathbb{R}$ er ikke inneholdt i en union av endelig mange av intervallene $(n,n+2)$. Altså er ikke (ii) i Definisjon 10.6.3 oppfylt. □

I et euklidsk rom har vi følgende berømte teorem oppkalt etter den tyske matematikeren Eduard Heine (1821–1881) og den franske matematikeren og politikeren Émile Borel (1871–1956):

Teorem 10.6.8: Heine-Borel teoremet

En delmengde av et euklidsk rom er kompakt hvis og bare hvis den er lukket og begrenset.

Bevis
Overlates til Oppgave 10.53. ("Bare hvis"-delen er sann i alle metriske rom og vises i neste resultat.) □

I et vilkårlig metrisk rom er kun "bare hvis"-delen sann:

Setning 10.6.9

La $K \subset X$ være en kompakt delmengde av et metrisk rom. Da er K lukket og begrenset.

Bevis

Vi utfører et kontrapositivt bevis og viser at dersom K ikke er lukket eller ikke er begrenset, da er K ikke kompakt.

Hvis K ikke er lukket, da finnes per definisjon en følge $\{x_n\}$ i K som konvergerer mot et punkt $x \in X \setminus K$. Siden enhver delfølge av $\{x_n\}$ konvergerer mot x ved Oppgave 10.47, har følgen $\{x_n\}$ ingen delfølge som konvergerer i K. Altså er K per definisjon ikke kompakt.

Hvis K ikke er begrenset, kan vi konstruere en følge $\{x_n\}$ i K på følgende måte: velg $x_1 \in K$. Siden K ikke er begrenset, finnes en $x_2 \in K$ slik at $d(x_1, x_2) \geq 1$. Når vi har valgt $x_1, \ldots, x_n$, velg av samme grunn en $x_{n+1} \in K$ slik at $d(x_{n+1}, x_n) \geq n$. Det er ikke vanskelig å se at $\{x_n\}$ ikke kan ha noen delfølge som er Cauchy. Dermed, ved Setning 10.2.7, kan $\{x_n\}$ heller ikke ha en delfølge som konvergerer i K. Altså er K per definisjon ikke kompakt. □

Merknad 10.6.10. Det motsatte av siste setning holder ikke i et vilkårlig metrisk rom: La $K = X = [a, b] \cap \mathbb{Q}$ for vilkårlige reelle tall $a < b$. Da er K lukket og begrenset. Men det finnes følger av rasjonale tall i $[a, b]$ som konvergerer mot et irrasjonalt tall (som for eksempel følgen i Eksempel 7.5.6), og slike følger har ingen konvergente delfølger ved Oppgave 7.27; dermed er ikke K kompakt.

Som et annet eksempel, la $K = X = (0, 1]$. Da er K lukket (i X) og begrenset, men K er ikke kompakt, som vist i Eksempel 10.6.5(b).

Se Eksempel 10.7.5 for et ytterligere eksempel.

Dette viser altså at kompakthet generelt er en sterkere betingelse enn bare å være lukket og begrenset.

Vi kan nå uttrykke generaliseringen av *ekstremalverdisetningen* til vilkårlige metriske rom, som en generalisering av Setning 10.6.1:

Setning 10.6.11

La $f : X \to Y$ være en kontinuerlig avbildning mellom metriske rom. Hvis X er kompakt, da er $f(X)$ kompakt.

Vi sier gjerne at *en kontinuerlig funksjon bevarer kompakte mengder*. Vi vil gi to bevis for setningen, som hvert bruker ett av kriteriene i Definisjon 10.6.3.

Bevis for Setning 10.6.11, variant I

Vi bruker kriteriet (i) i definisjonen av kompakthet.

La $\{y_n\}$ være en følge i $f(X)$. Vi vil vise at $\{y_n\}$ har en delfølge som konvergerer i $f(X)$.

Siden hver $y_n \in f(X)$, kan vi skrive $y_n = f(x_n)$ for $x_n \in X$. Siden X er kompakt, har $\{x_n\}$ en delfølge $\{x_{n_k}\}$ som konvergerer mot et punkt $x \in X$, det vil si $\lim_{k\to\infty} x_{n_k} = x$. Siden f er kontinuerlig, har vi

$$\lim_{k\to\infty} y_{n_k} = \lim_{k\to\infty} f(x_{n_k}) = f(\lim_{k\to\infty} x_{n_k}) = f(x) \in f(X),$$

og siden $\{y_{n_k}\}$ er en delfølge av $\{y_n\}$, er vi ferdig. □

Bevis for Setning 10.6.11, variant II

Vi bruker kriteriet (ii) i definisjonen av kompakthet.

La $\{U_\alpha\}$ være en samling åpne mengder i $f(X)$ slik at $f(X) = \cup_\alpha U_\alpha$. Vi vil vise at $f(X)$ kan uttrykkes som en union av endelig mange av $U_\alpha$ene.

Siden $f : X \to f(X)$ er kontinuerlig (ved Merknad 10.4.2), er hver $f^{-1}(U_\alpha)$ en åpen delmengde av X ved Setning 10.4.5. Vi har selvsagt at $X = \cup_\alpha f^{-1}(U_\alpha)$. Siden X er kompakt, finnes endelig mange av $f^{-1}(U_\alpha)$ene, si $f^{-1}(U_{\alpha_1}), \ldots, f^{-1}(U_{\alpha_n})$, slik at

$$X = f^{-1}(U_{\alpha_1}) \cup \cdots \cup f^{-1}(U_{\alpha_n}).$$

Det følger (sjekk!) at

$$f(X) = U_{\alpha_1} \cup \cdots \cup U_{\alpha_n}.$$

Dette viser at $f(X)$ er kompakt. □

Som en konsekvens av siste resultat, får vi et interessant kriterium som binder sammen kompakthet og kontinuitet som er parallell til Følgesetning 10.5.9, men sterkere, siden det er et "hvis og bare hvis"-utsagn:

Setning 10.6.12

La $f : I \to \mathbb{R}$ være en funksjon mellom metriske rom, med $I \subset \mathbb{R}$ et lukket begrenset intervall. Da er f kontinuerlig hvis og bare hvis grafen $\Gamma(f) \subset \mathbb{R}^2$ til f kompakt.

Bevis

Dette overlates til Oppgave 10.51. □

10.7. Flere eksempler på metriske rom

Vi har hittil fokusert oppmerksomheten på euklidske rom. Det er på tide å se på noen flere eksempler på metriske rom. Det skal vi gjøre i denne seksjonen og de to neste.

Den diskrete metrikken. En litt spesiell metrikk er følgende:

EKSEMPEL 10.7.1 (Diskret metrikk). La X være en ikke-tom mengde. Definér, for alle $x, y \in X$:

$$d(x,y) = \begin{cases} 1, & \text{hvis } x \neq y, \\ 0, & \text{hvis } x = y. \end{cases}$$

Dette er en metrikk på X (vi overlater dette til Oppgave 10.57), kalt den *diskrete metrikken*[16]. Med denne metrikken kalles X et *diskret metrisk rom*.

La $x \in X$. Vi merker at

$$N_\delta(x) = \begin{cases} \{x\}, & \text{hvis } \delta \leq 1, \\ X, & \text{hvis } \delta > 1, \end{cases}$$

altså består alle omegner enten kun av ett enkelt punkt eller av hele mengden X. Spesielt har vi for enhver delmengde $A \subset X$ og enhver $x \in A$, at $N_1(x) = \{x\} \subset A$, slik at x er et indre punkt og A er en åpen mengde. Altså er alle delmengder av et diskret metrisk rom åpne, og dermed også lukkede ved Setning 10.3.13.

Det er interessant, og kanskje overraskende, at funksjoner definert på diskrete rom automatisk er kontinuerlige:

EKSEMPEL 10.7.2. La X være et diskret metrisk rom , Y et vilkårlig metrisk rom og $f : X \to Y$ en funksjon. Vis at f er kontinuerlig.

LØSNING. La d_X og d_Y betegne metrikkene på X og Y, henholdsvis. Gitt en $a \in X$ og en $\epsilon > 0$, la $\delta \leq 1$. Hvis $x \in X$ slik at $d_X(x,a) < \delta$, er $x = a$ per definisjon av den diskrete metrikken, slik at

$$d_Y(f(x), f(a)) = d_Y(f(a), f(a)) = 0 < \epsilon.$$

(Alternativt kan vi bruke åpne-mengder-kriteriet fra Setning 10.4.5 og det faktum at alle delmengder av X er åpne.) $\square$

Vi overlater til Oppgave 10.58 å studere konvergens av følger i et diskret metrisk rom og spørsmålet om et slikt rom kan være komplett eller kompakt.

Den diskrete metrikken gir atter et eksempel på en lukket, begrenset mengde som ikke avbildes på en lukket begrenset mengde ved en kontinuerlig funksjon, som igjen viser at Setning 10.6.1 ikke holder i et vilkårlig metrisk rom:

[16] *Discrete metric* på engelsk.

Eksempel 10.7.3. La $f : \mathbb{Z} \to \mathbb{R}$ gitt ved $f(n) = n$ for alle $n \in \mathbb{Z}^+$. Utstyr $\mathbb{Z}$ med den diskrete metrikken og $\mathbb{R}$ med den euklidske. Da er f kontinuerlig ved Eksempel 10.7.2. Siden $d(m, n) \leq 1$ for alle $m, n \in \mathbb{Z}$, er $\mathbb{Z}$ begrenset, og $\mathbb{Z}$ er lukket (i seg selv, per definisjon av lukkethet). Men $f(\mathbb{Z}) = \mathbb{Z}$ er ikke begrenset som delmengde av $\mathbb{R}$ med den euklidske metrikken.

Vi får også et nytt enkelt eksempel på en kontinuerlig, bijektiv funksjon som ikke har en kontinuerlig invers (jf. §10.4):

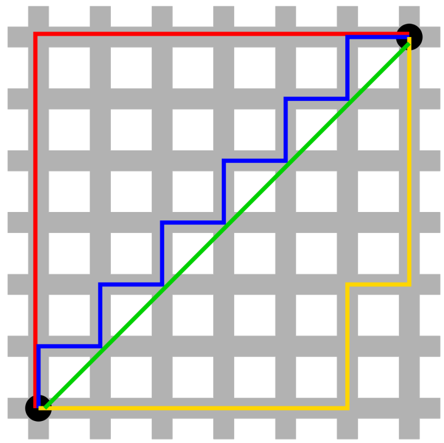
Eksempel 10.7.4. La $I \subset \mathbb{R}$ være et intervall med den induserte euklidske metrikken. La X være lik I utstyrt med den diskrete metrikken og $f : X \to I$ være *identitetsavbildningen*, det vil si $f(x) = x$ for alle $x \in X$. Da er f kontinuerlig ved Eksempel 10.7.2, men den inverse funksjonen $g = f^{-1} : I \to X$ er det ikke: for enhver $x \in X$ er $\{x\}$ åpen i X (jf. Eksempel 10.7.1), mens $g^{-1}(\{x\}) = \{x\}$ er *ikke* åpen i I, og dermed er g ikke kontinuerlig ved Setning 10.4.5.

Vi får også et nytt eksempel på at det motsatte av Setning 10.6.9, og dermed *Heine-Borel teoremet* 10.6.8, ikke holder i et vilkårlig metrisk rom:

Eksempel 10.7.5. Betrakt $\mathbb{Z}$ med de diskrete metrikken d. Da er $\mathbb{Z}$ lukket og begrenset, siden $d(m, n) \leq 1$ for alle $m, n \in \mathbb{Z}$. Men følgen $\{n\}$ har ingen konvergent delfølge, og $\mathbb{Z} = \cup_{n \in \mathbb{Z}} \{n\}$ er en åpen overdekning uten noen endelig deloverdekning, derfor er ikke $\mathbb{Z}$ kompakt.

Andre metrikker på $\mathbb{R}^n$. Som nevnt tidligere, kan en mengde ha flere forskjellige metrikker, deriblant den diskrete. Et godt eksempel er $\mathbb{R}^n$, som i tillegg til den euklidske har mange andre metrikker. Vi skal se på et par andre metrikker på $\mathbb{R}^n$.

Eksempel 10.7.6 ((Manhattan) Taxicab metrikken). Betrakt planet $\mathbb{R}^2$. Kjører man taxi i for eksempel Manhattan i New York, så er ikke den korteste avstanden mellom to punkter lik avstanden til "diagonalen" mellom punktene merket i grønt i figuren nedenunder, rett og slett fordi man ikke kan kjøre gjennom bygninger:



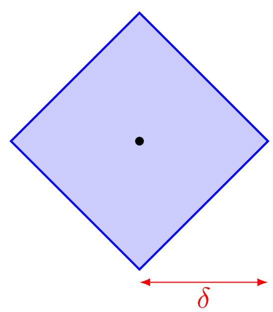
Avstanden mellom to punkter i planet i Taxicab metrikken (Kilde: Wikipedia, public domain)

Man må istedenfor kjøre en av strekningene merket i rødt, blått eller gult på figuren. Det er ikke vanskelig å se at avstanden mellom to punkter i planet $\mathbf{x} = (x_1, x_2)$ og $\mathbf{y} = (y_1, y_2)$ målt på denne måten er

$$\rho(\mathbf{x}, \mathbf{y}) = |x_1 - y_1| + |x_2 - y_2|.$$

Funksjonen $\rho : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definerer en metrikk på $\mathbb{R}^2$.

En omegn om et punkt ser slik ut:



En δ-omegn i Taxicab metrikken i $\mathbb{R}^2$

Metrikken kan generaliseres til en metrikk på $\mathbb{R}^n$ for alle $n \in \mathbb{Z}^+$: for $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ og $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$, la

$$\rho(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i|.$$

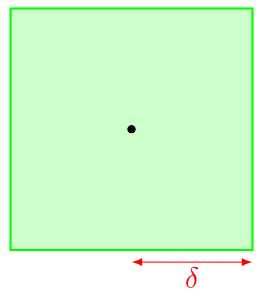
Da definerer $\rho : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ en metrikk på $\mathbb{R}^n$ (vises i Oppgave 10.59), kalt *(Manhattan) Taxicab metrikken.*

Eksempel 10.7.7 (Chebyshev-metrikken). Betrakt igjen planet $\mathbb{R}^2$. En annen mulighet for å måle avstanden mellom to punkter $\mathbf{x} = (x_1, x_2)$ og $\mathbf{y} = (y_1, y_2)$ er som

$$d_\infty(\mathbf{x}, \mathbf{y}) = \max\{|x_1 - y_1|, |x_2 - y_2|\}.$$

Denne funksjonen d_∞ definerer en metrikk på $\mathbb{R}^2$, kalt *Chebyshev-metrikken* etter den russiske matematikeren Pafnuty Lvovich Chebyshev (1821–1894), eller *sjakkbrett-metrikken*, siden avstanden mellom to punkter på sjakkbrettet målt på denne måten er det minste antallet trekk en konge trenger for å bevege seg mellom punktene.

En δ-omegn om et punkt består av det indre av et kvadrat med sidelengder 2δ om punktet:



En δ-omegn i Chebyshev metrikken i $\mathbb{R}^2$

Metrikken kan generaliseres til en metrikk på $\mathbb{R}^n$ for alle $n \in \mathbb{Z}^+$: for $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ og $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$, la

$$d_\infty(\mathbf{x}, \mathbf{y}) = \max_{1 \leq i \leq n} \{|x_i - y_i|\}.$$

Da definerer d_∞ en metrikk på $\mathbb{R}^n$ (vises i Oppgave 10.60).

Merknad 10.7.8. Den euklidske metrikken og metrikkene i Eksemplene 10.7.6 og 10.7.7 er helt like i tilfellene $n = 1$, det vil si på $\mathbb{R}$.

For $n \geq 2$ er de ikke like: vi har for eksempel sett at δ-omegner har forskjellig geometrisk form. På den annen side er det lett å se, nettopp ved å argumentere med den geometriske formen til omegnene, at enhver omegn i den ene metrikken inneholder en omegn av hver av de to andre metrikkene. Vi sier at metrikkene er *ekvivalente*. Genererelt er to metrikker på samme mengde *ekvivalente* hvis enhver omegn i den ene metrikken inneholder en omegn i den andre metrikken, og omvendt. I praksis betyr dette at alle spørsmål som avgjøres ved hjelp av omegner (som konvergens av følger og kontinuitet av funksjoner) har samme svar i de to metrikkene, og vi kan derfor bytte mellom metrikkene som det måtte passe oss. For eksempel vil ved Setning 10.2.12 alle Cauchy-følger i $\mathbb{R}^n$ konvergere uansett hvilke av de tre metrikkene (den euklidske, Manhattan Taxicab eller Chebyshev) vi bruker. Altså er også $\mathbb{R}^n$ utstyrt med Taxicab eller Chebyshev-metrikken komplett (jf. Definisjon 10.2.8).

Vi skal se på et par viktige eksempler til på metriske rom i de neste to seksjonene. Stoffet i resten av kapitlet er nokså krevende, så ikke bekymre deg om du ikke forstår alt.

10.8. Rom av følger*

Det finnes flere eksempler på metriske rom der elementene/punktene er følger. Slike rom kalles *rom av følger*[17]. Disse vil dere studere mer i kurset *MAT232–Funksjonalanalyse*. Vi skal se på ett slikt eksempel, og noen underrom.

EKSEMPEL 10.8.1. La ℓ_∞ være mengden av alle *begrensede* følger i $\mathbb{R}$. Vi betegner en slik følge i dette tilfellet gjerne med $\{x(k)\}_{k=1}^{\infty}$ eller bare $\{x(k)\}$. For $\mathbf{x} = \{x(k)\}$ og $\mathbf{y} = \{y(k)\}$ i ℓ_∞ definerer vi

$$d_\infty(\mathbf{x}, \mathbf{y}) = \sup\{|x(k) - y(k)|\}.$$

(Siden følgene er begrenset, er dette et reellt tall.) Dette definerer en metrikk på ℓ_∞ (vises i Oppgave 10.62), kalt *supremum-metrikken*. Dette er en generalisering av metrikken i Eksempel 10.7.7 til et rom med "tellbart mange koordinater".

EKSEMPEL 10.8.2. Rommet ℓ_∞ har flere interessante underrom.

Siden enhver konvergent følge av reelle tall er begrenset (Setning 7.4.10), er mengden c av alle *konvergente* følger i $\mathbb{R}$ en delmengde av ℓ_∞, og dermed et underrom med den induserte metrikken.

Et underrom av c er mengden c_0 av alle følger i $\mathbb{R}$ som konvergerer mot 0. Et underrom av c_0 er rommet c_{00} av alle følger $\{x(k)\}_{k=1}^{\infty}$ i $\mathbb{R}$ slik at kun endelig mange $x(k) \neq 0$ (ekvivalent: det finnes en $N \in \mathbb{Z}^+$ slik at $x(k) = 0$ for alle $k \geq N$). Vi har altså en kjede av underrom

$$c_{00} \subset c_0 \subset c \subset \ell_\infty.$$

Merk at for $\mathbf{x} = \{x(k)\}$ og $\mathbf{y} = \{y(k)\}$ i c_{00} er kun endelig mange $x(k) \neq 0$ og $y(k) \neq 0$, og dermed er metrikken på c_{00} gitt ved

$$d_\infty(\mathbf{x}, \mathbf{y}) = \max\{|x(k) - y(k)|\}.$$

Vi kan nå gi ytterligere et eksempel på et komplett og et ikke-komplett metrisk rom. Når vi jobber med følger i rommet ℓ_∞ og underrommene i siste eksempel, må vi holde tungen rett i munnen, for elementene i rommene er følger, så en følge i disse rommene er en følge hvor leddene selv er følger!

Setning 10.8.3

Rommet ℓ_∞ (jf. Eksempel 10.8.1) er komplett.

[17] *Sequence spaces* på engelsk.

Bevis
Anta at $\{\mathbf{x}_n\}$ er en Cauchy-følge i ℓ_∞. Vi vil vise at den konvergerer. Husk at hver $\mathbf{x}_n$ er i seg selv en følge. Vi skriver

$$\mathbf{x}_n = \{\mathbf{x}_n(1), \mathbf{x}_n(2), \ldots, \mathbf{x}_n(k), \ldots\} \quad \text{for} \quad n \in \mathbb{Z}^+.$$

La $\epsilon > 0$ være gitt. Per definisjon finnes en $N \in \mathbb{Z}^+$ slik at

$$d_\infty(\mathbf{x}_n, \mathbf{x}_m) < \epsilon \quad \text{for alle} \quad n, m \geq N. \tag{299}$$

Per definisjon av metrikken er (299) ekvivalent med

$$\sup_k |\mathbf{x}_n(k) - \mathbf{x}_m(k)| < \epsilon \quad \text{for alle} \quad n, m \geq N. \tag{300}$$

Dette impliserer

$$|\mathbf{x}_n(k) - \mathbf{x}_m(k)| < \epsilon \quad \text{for alle} \quad n, m \geq N \quad \text{og} \quad k \in \mathbb{Z}^+. \tag{301}$$

Derfor er følgen $\{\mathbf{x}_n(k)\}$ Cauchy i $\mathbb{R}$ for alle $k \in \mathbb{Z}^+$. Siden $\mathbb{R}$ er komplett (Setning 7.6.1), konvergerer $\{\mathbf{x}_n(k)\}$ i $\mathbb{R}$, si til $x(k) \in \mathbb{R}$, for hver $k \in \mathbb{Z}^+$. Vi definerer

$$\mathbf{x} = \{x(1), x(2), x(3), \ldots, x(k), \ldots\}$$

og vil vise at

$$\mathbf{x} \in \ell_\infty, \tag{302}$$

$$\mathbf{x}_n \to \mathbf{x} \quad \text{i} \quad \ell_\infty, \tag{303}$$

som vil vise setningen.

For å vise (302), legg merke til at $\{\mathbf{x}_n\}$ er begrenset, siden den er Cauchy (Setning 10.2.6). Dermed, hvis vi lar $\mathbf{0}$ betegne *null-følgen* $\{0, 0, 0, \ldots\}$, finnes en $C \in \mathbb{R}$ slik at

$$d_\infty(\mathbf{x}_n, \mathbf{0}) \leq C \quad \text{for alle} \quad n \in \mathbb{Z}^+.$$

Per definisjon av metrikken, er dette ekvivalent til $\sup_k |\mathbf{x}_n(k)| \leq C$ for alle $n \in \mathbb{Z}^+$, som igjen er ekvivalent til

$$|\mathbf{x}_n(k)| \leq C \quad \text{for alle} \quad k \in \mathbb{Z}^+, n \in \mathbb{Z}^+. \tag{304}$$

Fikserer vi k og lar $n \to \infty$ i (304) (og bruker at grenser bevarer orden ved Setning 7.4.14) får vi

$$\lim_{n \to \infty} |\mathbf{x}_n(k)| \leq C \quad \text{for alle} \quad k \in \mathbb{Z}^+.$$

Siden absoluttverdifunksjonen er kontinuerlig, gir Setning 9.4.1 at

$$|x(k)| \leq C \quad \text{for alle} \quad k \in \mathbb{Z}^+.$$

Dermed er $\mathbf{x} = \{x(1), x(2), \ldots\}$ en begrenset følge i $\mathbb{R}$, som betyr at (302) er sant.

For å vise (303), lar vi $m \to \infty$ i (301). Siden grenser bevarer orden (Setning 7.4.14), får vi

$$\lim_{m \to \infty} |\mathbf{x}_n(k) - \mathbf{x}_m(k)| \leq \epsilon \quad \text{for alle} \quad n \geq N \quad \text{og} \quad k \in \mathbb{Z}^+.$$

Siden absoluttverdifunksjonen er kontinuerlig, gir Setning 9.4.1 at

$$|\mathbf{x}_n(k) - x(k)| \le \epsilon \text{ for alle } n \ge N \text{ og } k \in \mathbb{Z}^+.$$

Dermed er

$$d_\infty(\mathbf{x}_n, \mathbf{x}) = \sup_k |\mathbf{x}_n(k) - x(k)| \le \epsilon \text{ for alle } n \ge N,$$

som viser (303). □

Setning 10.8.4

Rommet c_{00} (jf. Eksempel 10.8.2) er ikke komplett.

Bevis

Vi definerer en følge $\{\mathbf{x}_n\}$ i c_{00} ved

$$\mathbf{x}_n = \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots, \frac{1}{n}, 0, 0, 0, \dots\right\},$$

med andre ord

$$\mathbf{x}_n(k) = \begin{cases} \frac{1}{k}, & \text{hvis } 1 \le k \le n, \\ 0, & \text{hvis } k > n. \end{cases}$$

For $m > n$ er dermed

$$\mathbf{x}_m(k) - \mathbf{x}_n(k) = \begin{cases} 0, & \text{hvis } 1 \le k \le n, \\ \frac{1}{k}, & \text{hvis } n < k \le m, \\ 0, & \text{hvis } k > m, \end{cases}$$

slik at

$$d_\infty(\mathbf{x}_m, \mathbf{x}_n) = \sup_k |\mathbf{x}_m(k) - \mathbf{x}_n(k)| = \frac{1}{n+1}.$$

Dermed er $\{\mathbf{x}_n\}$ en Cauchy-følge. Vi vil vise at denne ikke har noen grense i c_{00}. Vi argumenterer ved selvmotsigelse og antar at $\{\mathbf{x}_n\}$ har en grense $\mathbf{x} = \{x(k)\} \in c_{00}$. Siden $\mathbf{x} \in c_{00}$, har vi per definisjon at $x(k) = 0$ for alle $k \ge N$, for en eller annen $N \in \mathbb{Z}^+$. For $m \ge N$ er derfor $\mathbf{x}_m(N) = \frac{1}{N}$, mens $x(N) = 0$. Dermed er

$$d_\infty(\mathbf{x}_m, \mathbf{x}) \ge |\mathbf{x}_m(N) - x(N)| = \frac{1}{N}$$

for alle $m \ge N$, som viser at $\lim_{m\to\infty} \mathbf{x}_m \ne \mathbf{x}$. □

Merknad 10.8.5. Vi overlater til Oppgave 10.65 å vise at underrommene c og c_0 er komplette. Faktisk går det også an å vise at c_0 er kompletteringen av c_{00} (jf. Merknad 10.2.15).

Vi avslutter seksjonen med et eksempel:

EKSEMPEL 10.8.6. Fiksér $\mathbf{x} = \{x(k)\}$ og $\mathbf{y} = \{y(k)\}$ i ℓ_∞. For enhver $t \in [0,1]$ er følgen $t\mathbf{x}+(1-t)\mathbf{y} = \{tx(k)+(1-t)y(k)\}$ begrenset og definerer dermed et element i ℓ_∞.

Vis at funksjonen $F : [0,1] \to \ell_\infty$ definert ved $f(t) = t\mathbf{x} + (1-t)\mathbf{y}$ er kontinuerlig.

LØSNING. Vi skal vise at F er kontinuerlig i et vilkårlig punkt $a \in [0,1]$. Vi vil derfor finne, til enhver $\epsilon \in \mathbb{R}^+$, en $\delta \in \mathbb{R}^+$ slik at

$$d_\infty(F(t), F(a)) < \epsilon \ \text{ når } \ |t-a| < \delta. \tag{305}$$

Vi studerer som vanlig uttrykket til venstre:

$$\begin{aligned} d_\infty(F(t), F(a)) &= \sup_k |(tx(k) + (1-t)y(k)) - (ax(k) + (1-a)y(k))| \\ &= \sup_k |(t-a)\,(x(k) - y(k))| \\ &= \sup_k |t-a|\,|x(k) - y(k)| \\ &= |t-a| \sup_k |x(k) - y(k)| \\ &= |t-a|\, d_\infty(\mathbf{x}, \mathbf{y}). \end{aligned}$$

Vi ser at dette er 0 hvis $\mathbf{x} = \mathbf{y}$; hvis $\mathbf{x} \neq \mathbf{y}$ er dette $< \epsilon$ hvis og bare hvis $|t-a| < \frac{\epsilon}{d_\infty(\mathbf{x},\mathbf{y})}$. Vi kan dermed la $\delta = \frac{\epsilon}{d_\infty(\mathbf{x},\mathbf{y})}$ for å oppfylle (305), hvis $\mathbf{x} \neq \mathbf{y}$, og δ være vilkårlig, hvis $\mathbf{x} = \mathbf{y}$. $\square$

10.9. Funksjonsrom*

Viktige mengder i matematikk er mengder hvor elementene er funksjoner med en felles definisjonsmengde. Utstyrt med en metrikk kalles slike rom gjerne for *funksjonsrom*[18]. Disse vil dere studere mer i kursene *MAT211–Reell Analyse* og *MAT232–Funksjonalanalyse*. Spesielt er disse rommene viktige i studiet av løsninger på differensialligninger, siden disse kan uttrykkes som punkter i funksjonsrom som oppfyller visse betingelser.

Vi skal se på et viktig slikt eksempel.

EKSEMPEL 10.9.1. La $\mathcal{C}[a,b]$ være mengden av alle kontinuerlige funksjoner $[a,b] \to \mathbb{R}$. For $f, g \in \mathcal{C}[a,b]$, er funksjonen

$$\begin{array}{rcl} [a,b] & \longrightarrow & \mathbb{R} \\ x & \mapsto & |f(x) - g(x)| \end{array}$$

kontinuerlig ved Setningene 9.5.1 og 9.5.2, sammen med Eksempel 9.2.3. Den oppnår derfor en maksimumsverdi, ved *ekstremalverdisetningen* 9.7.2.

[18] *Function spaces* på engelsk.

Vi har dermed en veldefinert funksjon

$$\begin{aligned} d_\infty : \mathcal{C}[a,b] \times \mathcal{C}[a,b] &\longrightarrow \mathbb{R} \\ (f,g) &\mapsto \sup_{x\in[a,b]}\{|f(x)-g(x)|\} \\ &\quad = \max_{x\in[a,b]}\{|f(x)-g(x)|\}. \end{aligned}$$

Da definerer d_∞ en metrikk på $\mathcal{C}[a,b]$ (vises i Oppgave 10.67).

Vi kan også bytte ut $[a,b]$ med en hvilken som helst delmengde X av $\mathbb{R}$ og definere $\mathcal{C}(X)$ til å være mengden av alle kontinuerlige *begrensede* funksjoner $X \to \mathbb{R}$. Da vil

$$\begin{aligned} d_\infty : \mathcal{C}(X) \times \mathcal{C}(X) &\longrightarrow \mathbb{R} \\ (f,g) &\mapsto \sup_{x\in[a,b]}\{|f(x)-g(x)|\}. \end{aligned}$$

definere en metrikk på $\mathcal{C}(X)$. Vi må kreve at funksjonene i $\mathcal{C}(X)$ er begrenset for å garantere at supremum finnes.

Tilsvarende definisjoner får vi ved å bytte ut $\mathbb{R}$ med $\mathbb{C}$.

Det neste resultatet, bevist av Weierstrass, er svært viktig og gir et nytt eksempel på et komplett metrisk rom:

Setning 10.9.2

Funksjonsrommet $\mathcal{C}[a,b]$ (jf. Eksempel 10.9.1) er komplett.

Bevis

La $\{f_n\}$ være en Cauchy-følge i $\mathcal{C}[a,b]$. Hvert element er en kontinuerlig funksjon $f_n : [a,b] \to \mathbb{R}$, slik at $f_n(x) \in \mathbb{R}$ for alle $n \in \mathbb{Z}^+$, $x \in [a,b]$. Vi vil vise at f_n konvergerer mot en $f \in \mathcal{C}[a,b]$.

Gitt $\epsilon > 0$. Cauchy-betingelsen sier at det finnes en $N \in \mathbb{Z}^+$ slik at

$$d_\infty(f_m, f_n) = \sup_{x\in[a,b]} |f_m(x) - f_n(x)| < \epsilon \text{ for alle } m,n \geq N.$$

Spesielt betyr dette at

$$|f_m(x) - f_n(x)| < \epsilon \text{ for alle } m,n \geq N, x \in [a,b],$$

hvilket innebærer at hver følge $\{f_n(x)\}$ er en Cauchy-følge av reelle tall, for hver fiksert $x \in [a,b]$. Siden $\mathbb{R}$ er komplett (Setning 7.6.1), konvergerer $\{f_n(x)\}$ mot et reellt tall, som selvsagt avhenger av x. Vi får dermed definert en funksjon

$$\begin{aligned} f : [a,b] &\longrightarrow \mathbb{R} \\ x &\mapsto \lim_{n\to\infty} f_n(x). \end{aligned}$$

For å fullføre beviset, vil vi vise at

$$f \in \mathcal{C}[a,b], \text{ det vil si: } f \text{ er kontinuerlig;} \tag{306}$$

$$f_n \to f \text{ i } \mathcal{C}[a,b]. \tag{307}$$

Vi viser først (306). La $x_0 \in [a,b]$ være et vilkårlig punkt. Vi vil vise at f er kontinuerlig i x_0. Vi vil altså finne, til enhver gitt $\epsilon > 0$, en $\delta > 0$ slik at

$$|f(x) - f(x_0)| < \epsilon \text{ når } |x - x_0| < \delta,\ x \in [a,b]. \tag{308}$$

Vi skriver, for $x \in [a,b]$:

$$\begin{aligned} |f(x) - f(x_0)| &= |f(x) - f_n(x) + f_n(x) - f_n(x_0) + f_n(x_0) - f(x_0)| \\ &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)|, \end{aligned} \tag{309}$$

for en hvilken som helst $n \in \mathbb{Z}^+$, hvor vi har brukt trekantulikheten (i $\mathbb{R}$). Nå vil vi gjøre hvert av leddene mindre enn $\frac{\epsilon}{3}$.

Ved Cauchy-egenskapen til $\{f_n\}$ kan vi velge $N \in \mathbb{Z}^+$ slik at

$$d_\infty(f_m, f_n) = \sup_{x \in [a,b]} |f_m(x) - f_n(x)| < \frac{\epsilon}{3} \text{ for alle } m, n \geq N.$$

Spesielt betyr dette at

$$|f_m(x) - f_n(x)| < \frac{\epsilon}{3} \text{ for alle } m, n \geq N, x \in [a,b]. \tag{310}$$

Husk at for fiksert x er $\{f_m(x)\}$ en følge av reelle tall. Fikserer vi n og lar $m \to \infty$ i (310), får vi, siden grenser bevarer orden (Setning 7.4.14), at

$$\lim_{m \to \infty} |f_m(x) - f_n(x)| \leq \frac{\epsilon}{3} \text{ for alle } n \geq N, \quad x \in [a,b].$$

Siden absoluttverdifunksjonen er kontinuerlig (Eksempel 9.2.3), gir Setning 9.4.1 at

$$|f(x) - f_n(x)| \leq \frac{\epsilon}{3} \text{ for alle } n \geq N, \quad x \in [a,b]. \tag{311}$$

Siden f_n er kontinuerlig (fordi den tilhører $\mathcal{C}[a,b]$), finnes det en $\delta_n > 0$ slik at

$$|f_n(x) - f_n(x_0)| < \frac{\epsilon}{3} \text{ når } |x - x_0| < \delta_n,\ x \in [a,b]. \tag{312}$$

Dermed har vi, ved (309) med en hvilken som helst $n \geq N$, og ved (311)-(312), at for alle $x \in [a,b]$ slik at $|x - x_0| < \delta_n$ er

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon, \end{aligned}$$

og vi har vist (308), og dermed (306).

Vi har faktisk så godt som vist (307) i samme slengen: ved (311) har vi

$$d_\infty(f, f_n) = \sup_{x \in [a,b]} |f(x) - f_n(x)| \leq \frac{\epsilon}{3} < \epsilon \text{ for alle } n \geq N,$$

som viser (307). □

I løpet av beviset for siste setning, møtte vi på to typer konvergens:

- For fiksert $x \in [a,b]$ er $\{f_n(x)\}$ en følge av reelle tall. At denne konvergerer mot $f(x)$ betyr per definisjon at for enhver $\epsilon > 0$, finnes en $N \in \mathbb{Z}^+$ (som avhenger av ϵ *og* av x) slik at

(313) $$|f_n(x) - f(x)| < \epsilon \text{ for alle } n \geq N \text{ (for fiksert } x).$$

- At følgen $\{f_n\}$ *konvergerer som en følge i* $\mathcal{C}[a,b]$ mot en funksjon $f \in \mathcal{C}[a,b]$ betyr per definisjon av konvergens med hensyn på metrikken i $\mathcal{C}[a,b]$ at for enhver $\epsilon > 0$, finnes en $N \in \mathbb{Z}^+$ (som avhenger kun av ϵ) slik at

(314) $$\sup_{x\in[a,b]} |f_n(x) - f(x)| < \epsilon \text{ for alle } n \geq N.$$

Dette er ekvivalent med

(315) $$|f_n(x) - f(x)| < \epsilon \text{ for alle } n \geq N \text{ og } x \in [a,b].$$

Betingelsen (313) er en svakere betingelse enn (315). Vi skiller mellom disse to typene konvergens og kaller den første for *punktvis konvergens* og den andre for *uniform konvergens*. Vi kan snakke om slike typer konvergens uavhengig av om funksjonene i følgen og grensefunksjonen er kontinuerlige. Vi oppsummerer og generaliserer litt:

Definisjon 10.9.3: Punktvis og uniform konvergens

La X være et metrisk rom og la $f_n : X \to \mathbb{K}$, med $\mathbb{K} = \mathbb{R}$ eller $\mathbb{C}$, $n \in \mathbb{Z}^+$, være en følge av funksjoner.

Vi sier at $\{f_n\}$ *konvergerer punktvis*[a] eller er *punktvis konvergent*[b] dersom alle følgene $\{f_n(x)\}$ konvergerer i $\mathbb{K}$ for alle $x \in \mathbb{K}$. Vi kan da definere *grensefunksjonen* $f : X \to \mathbb{K}$ ved

$$f(x) = \lim_{n\to\infty} f_n(x).$$

At $\{f_n\}$ konvergerer punktvis mot f betyr altså at for enhver $\epsilon > 0$ og $x \in X$, finnes en $N \in \mathbb{Z}^+$ (som avhenger av ϵ og av x) slik at

$$|f_n(x) - f(x)| < \epsilon \text{ for alle } n \geq N.$$

Vi sier at $\{f_n\}$ *konvergerer uniformt*[c] eller er *uniformt konvergent*[d] mot f dersom det for enhver $\epsilon > 0$, finnes en $N \in \mathbb{Z}^+$ (som avhenger av ϵ) slik at

$$|f_n(x) - f(x)| < \epsilon \text{ for alle } n \geq N \text{ og alle } x \in X.$$

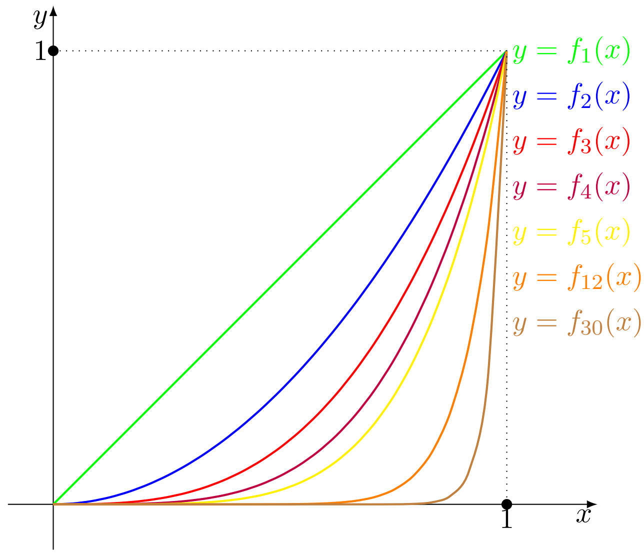
[a] *Converges pointwise* på engelsk.
[b] *Pointwise convergent* på engelsk.
[c] *Converges uniformly* på engelsk.
[d] *Uniformly convergent* på engelsk.

Forskjellen er altså at for uniform konvergens kreves det at N skal være uavhengig av $x \in X$, mens for punktvis konvergens kan den godt være avhengig av $x \in X$.

Eksempel 10.9.4. La $f_n : [0,1] \to \mathbb{R}$ være gitt ved $f_n(x) = x^n$. Da er $\lim_{n\to\infty} f_n(x) = 0$ for alle $x \in [0,1)$ ved Oppgave 7.21(a), mens $\lim_{n\to\infty} f_n(x) = 1$

for $x = 1$. Funksjonsfølgen konvergerer altså punktvis mot grensefunksjonen $f : [0,1] \to \mathbb{R}$ gitt ved

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



Grafene til noen av funksjonene f_n fra Eksempel 10.9.4

Men konvergensen er ikke uniform: Hvis den var det, ville det for enhver $\epsilon > 0$ eksistere en $N \in \mathbb{Z}^+$ slik at

$$|x^n - f(x)| < \epsilon \ \text{ for alle } \ n \geq N \ \text{ og } \ x \in [0,1].$$

Spesielt må dette være sant for alle $x \in [0,1)$, hvor $f(x) = 0$, slik at vi har:

$$|x^n| < \epsilon \ \text{ for alle } \ n \geq N \ \text{ og } \ x \in [0,1),$$

som er ekvivalent med

$$x < \sqrt[n]{\epsilon} \ \text{ for alle } \ n \geq N \ \text{ og } \ x \in [0,1).$$

Men dette kan ikke holde hvis $\epsilon < 1$, siden da er også $\sqrt[n]{\epsilon} < 1$ for alle $n \in \mathbb{Z}^+$, slik at vi får motsigelsen:

$$x < \sqrt[N]{\epsilon} < 1 \ \text{ for alle } \ x \in [0,1).$$

Vi merker oss også at grensefunksjonen f er diskontinuerlig.

Det interessante med funksjonsfølger er at vi kan bruke konvergente funksjonsfølger til å definere nye funksjoner, grensefunksjonene, som ikke kan uttrykkes ved hjelp av de operasjonene vi er vant med. Et berømt eksempel er *Weierstrass-funksjonen* vi traff på i §1.4. Når vi skriver

$$f(x) = \cos(\pi x) + \frac{3}{4}\cos(9\pi x) + \frac{9}{16}\cos(81\pi x) + \frac{27}{64}\cos(729\pi x) + \cdots,$$

eller

$$f(x) = \sum_{k=0}^{\infty} \left(\frac{3}{4}\right)^k \cos(9^k \pi x)$$

med notasjonen for *rekker* fra slutten på §7.4, mener vi at $f : \mathbb{R} \to \mathbb{R}$ er grensefunksjonen til følgen $\{f_n : \mathbb{R} \to \mathbb{R}\}$ hvor

$$f_n(x) = \sum_{k=0}^{n} \left(\frac{3}{4}\right)^k \cos(9^k \pi x).$$

Funksjonen f kan ikke uttrykkes på en enklere måte enn dette.

Merk at følgen $\{f_n(x)\}$ er en *rekke*, assosiert til følgen $\left\{\left(\frac{3}{4}\right)^n \cos(9^n \pi x)\right\}$ (jf. slutten på §7.4). En slik funksjonsfølge kalles derfor også en *funksjons-rekke*[19].

Siden kontinuitet er et såpass viktig begrep, er det essensielt å finne resultater som garanterer at grensefunksjonen f er kontinuerlig når alle f_n er kontinuerlige. Punktvis konvergens er ikke nok, som Eksempel 10.9.4 viste.

Imidlertid garanterer uniform konvergens at grensefunksjonen er kontinuerlig hvis funksjonene i følgen er det:

Setning 10.9.5

La X være et metrisk rom og la $f_n : X \to \mathbb{R}$ (eller $\mathbb{C}$), $n \in \mathbb{Z}^+$, være en følge av kontinuerlige funksjoner. Dersom $\{f_n\}$ konvergerer uniformt mot funksjonen f, da er f kontinuerlig.

Bevis

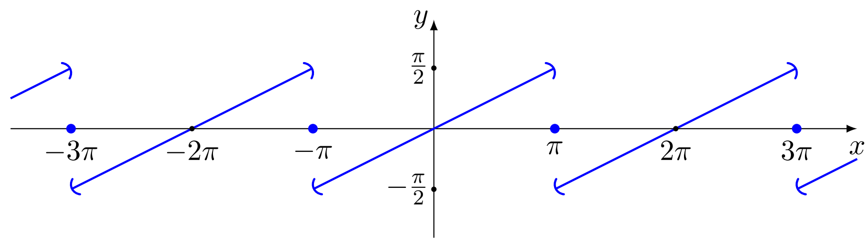
Beviset er en modifisering av deler av beviset for Setning 10.9.2, ved å bytte ut $[a, b]$ med X, og overlates til Oppgave 10.73. $\square$

Merknad 10.9.6. Den franske matematikeren og fysikeren Augustin-Louis Cauchy (1789–1857) publiserte i sin berømte bok *Cours d'Analyse de l'École Royale Polytechnique* fra 1821 en setning som hevdet det samme som Setning 10.9.5 under den svakere betingelsen *punktvis konvergens* istedenfor *uniform konvergens*. Eksempel 10.9.4 viser at Cauchys resultat ikke var riktig, og viser oss samtidig at selv de store kan feile! Det var faktisk Niels Henrik Abel (1802–1829) som i en avhandling publisert i 1826 først fant et moteksempel til Cauchys setning: Det var funsjonsfølgen $\{f_n : \mathbb{R} \to \mathbb{R}\}$ med

$$\begin{aligned} f_n(x) &= \sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} + \cdots + (-1)^{n-1} \frac{\sin nx}{n} \\ &= \sum_{k=1}^{n} (-1)^{k-1} \frac{\sin kx}{k}, \end{aligned}$$

som konvergerer mot den diskontinuerlige funksjonen med følgende graf:

[19] *Series of functions* på engelsk.



Grafen til grensefunksjonen til følgen $f_n(x)$.

Abel var ikke i stand til å finne feilen i Cauchys bevis. Det var først i 1847 at den tyske matematikeren Philipp Ludwig von Seidel (1821–1896) og den irske matematikeren og fysikeren George Stokes (1819–1903) rettet feilen ved hjelp av *uniform konvergens*. Selve konseptet *uniform konvergens* dukket først opp, uten en formell definisjon, i en artikkel fra 1838 av den tyske matematikeren Christoph Gudermann (1798–1852). Blant studentene i Gudermanns forelesninger var Karl Weierstrass (1815–1897), som så dybden i det nye begrepet og tok det systematisk i bruk i sine forelesninger. Han introduserte navnet *uniform konvergent* ("gleichmäßig konvergent" på tysk) i en artikkel fra 1841, publisert først i 1894.

Merknad 10.9.7. Et viktig eksempel på grensefunksjoner er grensen f til funksjonfølger $\{f_n\}$ av typen $f_n(x) = \sum_{k=0}^{n} c_k(x-a)^k$, hvor $a \in \mathbb{R}$ er fiksert, og $c_k \in \mathbb{R}$ for $k \in \mathbb{N}$. Med andre ord er f en funksjonsrekke. Funksjonsrekker av denne typen kalles *potensrekker*[20]. Disse vil dere møte på blant annet i kurset *MAT112-Grunnkurs i Matematikk II*, jf. [**AE**, §9.5].

Man kan vise at en potensrekke enten

(i) konvergerer kun for $x = a$;
(ii) konvergerer for alle x;
(iii) konvergerer hvis og bare hvis $x \in (a-R, a+R)$, pluss muligens ett eller begge av endepunktene, for en $R > 0$.

Et berømt teorem av Abel fra den ovennevnte avhandlingen publisert i 1826 garanterer at grensefunksjonen er kontinuerlig overalt hvor den er definert.

Et konkret eksempel har vi sett i Oppgave 7.21(c), hvor vi viste at

$$\sum_{n=0}^{\infty} x^n = \lim_{n\to\infty} \sum_{k=0}^{n} x^k = \frac{1}{1-x}, \quad x \in (-1,1).$$

Dette er en potensrekke med $a = 0$ og $c_k = 1$ for alle k. Andre berømte eksempler på konvergente potensrekker med $a = 0$ er, jf. igjen [**AE**, §9.5]:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad x \in \mathbb{R}, \tag{316}$$

[20] *Power series* på engelsk.

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}, \quad x \in \mathbb{R}, \tag{317}$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}, \quad x \in \mathbb{R}. \tag{318}$$

I noen lærebøker tas uttrykkene (316)-(318) som definisjon på funksjonene. I tilfellet med sinus- og cosinusfunksjonen har dette fordelen at vi slipper å måtte definere disse ved hjelp av begrepet buelengde av en sirkelsektor.

10.10. Topologiske rom*

Det er interessant å merke seg at en del av egenskapene vi har sett for metriske rom kan formuleres ved hjelp av omegner og åpne mengder, uten "ϵ-argumenter" eller bruk av selve metrikken. Dette er utgangspunktet for begrepet *topologi* på en mengde X, som er grunnlaget for de mange retningene innenfor *geometri*, slik som *differensialgeometri* og *algebraisk geometri*. En introduksjon til dette gis i kursene *MAT242–Topologi* og *MAT243–Mangfoldigheter*. Vi skal gi en kort oversikt her.

Prinsippet er å definere *aksiomatisk* hva det vil si at en delmengde av X er åpen. En omegn om et punkt $x \in X$ blir deretter definert som en åpen mengde som inneholder x. Deretter kan vi definere sentrale begreper som konvergens av følger og kontinuitet (samt sammenhengende og kompakt) kun ved hjelp av de åpne mengdene og omegnene.

Hvordan definerer vi så *åpenhet* av delmengder av en mengde X aksiomatisk? Jo, vi tar utgangspunkt i det vi kjenner til fra metriske rom: vi vet at X og $\emptyset$ er åpne delmengder av X (jf. teksten rett etter Definisjon 10.3.1); dessuten vet vi at åpne mengder oppfyller betingelsene om union og snitt i Setning 10.3.26(i)-(ii). Dette er alt vi trenger for følgende aksiomatiske definisjon:

Definisjon 10.10.1: Topologi og topologisk rom

La X være en mengde. En *topologi*[a] på X er en samling $\mathcal{T}$ av delmengder av X som har følgende egenskaper:

- X og $\emptyset$ er med i $\mathcal{T}$.
- En vilkårlig union av mengder i $\mathcal{T}$ er igjen med i $\mathcal{T}$.
- Et vilkårlig endelig snitt av mengder i $\mathcal{T}$ er igjen med i $\mathcal{T}$.

Vi kaller elementene i $\mathcal{T}$ for *åpne mengder*.

En mengde X med en topologi $\mathcal{T}$ på X kalles et *topologisk rom*[b].

[a] *Topology* på engelsk.
[b] *Topological space* på engelsk.

Som før kaller vi elementene i X for *punkter*.

Vi kan altså si at et topologisk rom er en mengde X sammen med en samling delmengder, kalt åpne, slik at X og $\emptyset$ begge er åpne, og slik at vilkårlige unioner og endelige snitt av åpne mengder er åpne. Ved Setning 10.3.26(i)-(ii) er dermed alle metriske rom topologiske rom. I slike tilfeller sier vi at *topologien er indusert av metrikken på rommet* og kaller den for *den metriske topologien.*

Et topologisk rom er strengt tatt et *par* $(X, \mathcal{T})$ hvor X er en mengde og $\mathcal{T}$ er en topologi på X. Vi sier imidlertid ofte kun at "X er et topologisk rom" når topologien er underforstått.

Grunnleggende definisjoner. Før vi ser på noen eksempler på topologiske rom, tar vi med noen viktige definisjoner.

Definisjon 10.10.2: Omegn

La X være et topologisk rom og $x \in X$. En *omegn om x* er en åpen mengde som inneholder x.

Merk at begrepet er noe videre enn en *omegn* i et metrisk rom som vi har definert tidligere, som er en åpen mengde av en spesiell type. Det er i praksis ingen forskjell hvilken definisjon vi opererer med for metriske rom, jf. Merknad 2.1.24.

Merknad 10.10.3. Noen matematiske tekster definerer en *omegn om x* til å være en mengde A som er slik at det finnes en åpen mengde V slik at $x \in V \subset A \subset X$. En *omegn* i vår forstand som i Definisjon 10.10.2 blir da kalt en *åpen omegn.* Det er viktig å være klar over hvilken definisjon teksten opererer med.

Neste definisjon er begrepet *lukket mengde.* I et metrisk rom X så vi at det finnes (minst) tre ekvivalente beskrivelser av at en delmengde $A \subset X$ er lukket:

(i) A inneholder alle sine følgegrenser (Definisjon 10.3.8);
(ii) $X \setminus A$ er åpen i X (Setning 10.3.13);
(iii) A inneholder randen ∂A (Setning 10.3.21).

Som vi skal se nedenunder, gjelder ikke lenger ekvivalensen (i) $\Leftrightarrow$ (ii) i et generelt topologisk rom (kun implikasjonen (i) $\Leftarrow$ (ii) er sann), mens ekvivalensen (ii) $\Leftrightarrow$ (iii) gjelder fremdeles. Vi bruker derfor (ii) som utgangspunkt for definisjon av lukkethet i topologiske rom:

Definisjon 10.10.4: Lukket mengde

La X være et topologisk rom. En delmengde $A \subset X$ sies å være *lukket* dersom komplementet $X \setminus A$ er åpent.

Definisjon 10.3.17 og Setning 10.3.21 overføres ordrett til topologiske rom:

Definisjon 10.10.5: Randpunkt og rand

La X være et metrisk rom og $A \subset X$ en delmengde.

Et punkt $x \in X$ er et randpunkt til A hvis enhver omegn om x innholder både punkter i A og utenfor A.

Mengden av alle randpunkter til A kalles randen til A og betegnes gjerne med $\mathrm{bd}(A)$ eller ∂A.

Setning 10.10.6: Ekvivalent beskrivelse av lukkethet

La X være et metrisk rom og $A \subset X$ en delmengde. Da er A lukket hvis og bare hvis $\partial A \subset A$.

Bevis

Dette følger beviset for Setning 10.3.21 så å si ordrett og overlates til Oppgave 10.74. □

Vi pleier å skille ut topologiske rom som har en viktig tilleggsegenskap:

Definisjon 10.10.7: Hausdorffrom

En topologi sies å være *Hausdorff* og det tilhørende topologiske rommet et *Hausdorffrom* dersom det for ethvert par av punkter $x, y \in X$, finnes en omegn U_x om x og en omegn U_y om y slik at $U_x \cap U_y = \emptyset$.

Grunnen til at vi skiller ut Hausdorffrom er at denne ekstra betingelsen er nødvendig for at en del resultater vi kjenner fra metriske rom generaliseres. Se for eksempel Setning 10.10.20 (kontra Eksempel 10.10.19), Setning 10.10.41 (kontra Eksempel 10.10.42), samt Oppgavene 10.80 og 10.92. Som vi vil se nedenunder, er ikke alle topologiske rom Hausdorff (jf. Eksemplene 10.10.11 og 10.10.14-10.10.16).

Merknad 10.10.8. Et metrisk rom er Hausdorff, siden vi kan la $\delta = \frac{1}{2}d(x, y)$, og da er $N_\delta(x) \cap N_\delta(y) = \emptyset$.

Underrom. Hvis X er et topologisk rom, med topologi $\mathcal{T}$, og $Y \subset X$ er en delmengde, kan man vise (jf. Oppgave 10.75) at samlingen av delmengder av Y gitt ved

$$\mathcal{T}_Y = \{U \subset Y \mid U = Y \cap V \text{ for en } V \in \mathcal{T}\} \tag{319}$$

definerer en topologi på Y. Med andre ord:

$$U \text{ er åpen i } Y \iff U = Y \cap V \text{ for en åpen delmengde } V \text{ i } X, \tag{320}$$

som generaliserer Setning 10.3.6 for metriske rom.

Definisjon 10.10.9: Underromstopologien

Topologien definert ved (319), ekvivalent (320), kalles *den induserte topologien*[a], *den relative topologien*[b] eller *underromstopologien*[c]. Hvis $Y \subset X$ har denne topologien, sier vi at Y er et *underrom* av X.

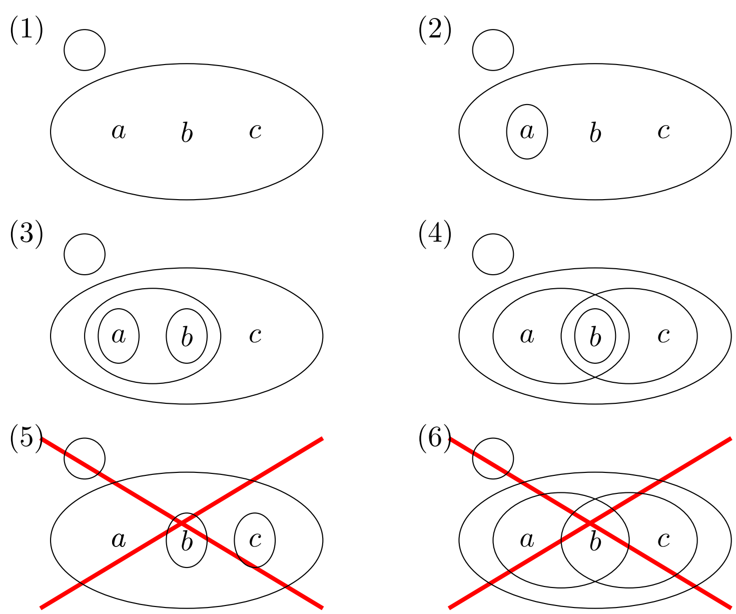
[a] *Induced topology* på engelsk.
[b] *Relative topology* på engelsk.
[c] *Subspace topology* på engelsk.

Dette er kompatibelt med begrepene for metriske rom vi har introdusert i Definisjon 10.1.9.

Eksempler. La oss se på noen eksempler på topologiske rom.

EKSEMPEL 10.10.10 (Diskret topologi). Hvis X er en mengde, er samlingen $\mathcal{T}$ av *alle* delmengdene av X automatisk en topologi på X. Dette er samme topologi vi får indusert av den diskrete metrikken på X (jf. Eksempel 10.7.1), og topologien er derfor Hausdorff. Denne topologien kalles den *diskrete topologien* på X.

EKSEMPEL 10.10.11. La oss se på en endelig mengde bestående av tre punkter, si $X = \{a, b, c\}$ og hvordan vi kan definere forskjellige samlinger av delmengder av X. I figuren nedenunder lar vi alle innringede mengder forme en samling av delmengder. (Ringen rundt intet punkt symbolerer den tomme mengden.)



Vi ser at mengdene X og $\emptyset$ er med i samlingen i alle tilfeller, slik at betingelsen (i) i Definisjon 10.10.1 er oppfylt. Siden mengden er endelig, må vi for å verifisere (ii) og (iii) i samme definisjon kun passe på at unioner og snitt av delmengdene i samlingen fremdeles er med i delmengden. Vi kan lett verifisere at dette er oppfylt i tilfellene (1)-(4). I tilfelle (1) består samlingen kun av mengdene X og $\emptyset$, og en slik topologi kalles gjerne *indiskret* eller *triviell*, men den er temmelig uinteressant. Vi merker også at i alle tilfellene (1)-(4) oppfyller topologien ikke Hausdorff-betingelsen (Definisjon 10.10.7), siden det for eksempel aldri finnes to disjunkte mengder i samlingen som inneholder b og c. Ved Merknad 10.10.8 er ingen av disse topologiene indusert av noen metrikk på X.

Samlingen (5) definerer ikke en topologi: Delmengdene $\{b\}$ og $\{c\}$ er med i samlingen, men unionen av dem $\{b, c\}$ er ikke med.

Samlingen (6) definerer heller ikke en topologi: Delmengdene $\{a, b\}$ og $\{b, c\}$ er med i samlingen, mens snittet $\{b\}$ er ikke med.

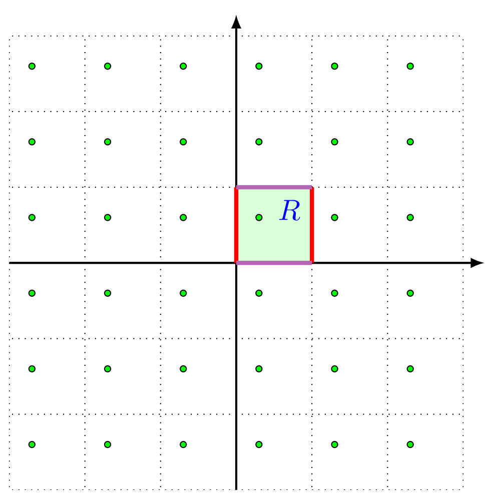
Eksempel 10.10.12 (Torusen). Betrakt følgende ekvivalensrelasjon på planet $\mathbb{R}^2$:

$$(x, y) \sim (x+1, y) \sim (x, y+1), \quad \text{for alle} \;\; x, y \in \mathbb{R}^2.$$

Mer presist, for alle $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$, lar vi

$$(x_1, y_1) \sim (x_2, y_2) \Longleftrightarrow x_1 - x_2 \in \mathbb{Z}, \; y_1 - y_2 \in \mathbb{Z}. \tag{321}$$

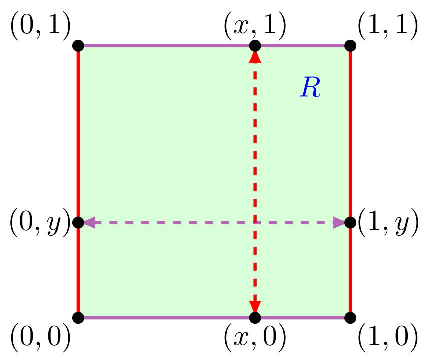
Vi overlater til Oppgave 10.77 å vise at dette virkelig er en ekvivalensrelasjon. Dette betyr i praksis at vi deler opp $\mathbb{R}^2$ i et gitter bestående av kvadrater av sidelengde 1 og betrakter alle punktene i samme posisjon i hvert kvadrat som ekvivalente. For eksempel danner alle punktene markert i grønt i figuren nedenunder en ekvivalensklasse:



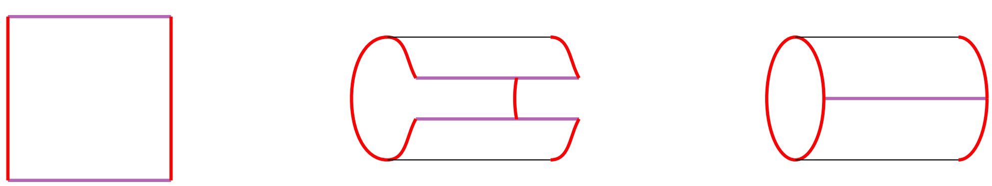
La $T = \mathbb{R}^2/\sim$, mengden av alle ekvivalensklassene. Da har vi som vanlig en naturlig avbildning

$$f : \mathbb{R}^2 \longrightarrow T$$

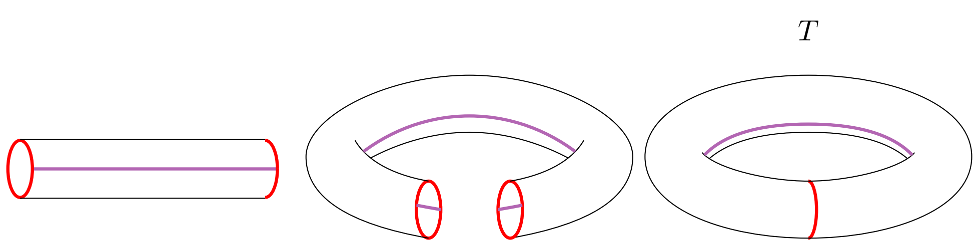
som sender et punkt på sin ekvivalensklasse. Altså blir for eksempel alle punktene i grønt i figuren ovenfor avbildet på samme punkt i T. For å visualisere mengden T bedre kan vi restrisere oss til kvadratet R med kanter $(0,0), (1,0), (1,1), (0,1)$ markert i figuren ovenfor. Restriksjonen $f|_R$ av f til R identifiserer de to horisontale sidene (markert i fiolett) ved at $f(x,0) = f(x,1)$ for alle $x \in [0,1]$ og identifiserer de to vertikale sidene (markert i rødt) ved at $f(0,y) = f(1,y)$ for alle $y \in [0,1]$, som vist i følgende figur:



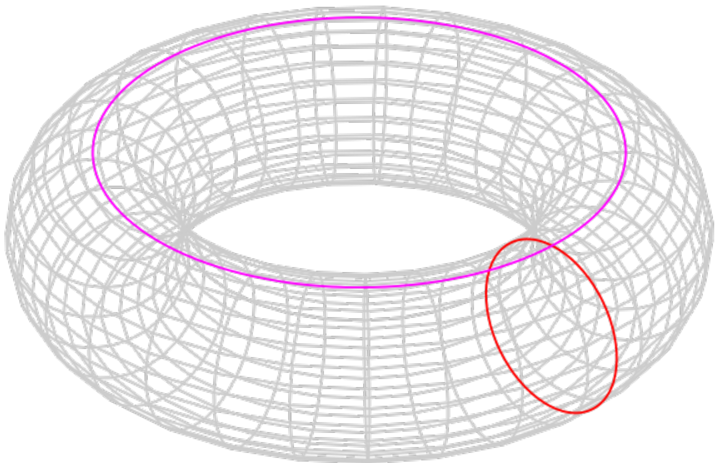
I det indre av R er $f|_R$ injektiv. At $f|_R$ identifiserer de to horisontale sidene (i fiolett) av R kan visualiseres i følgende figur:



At $f|_R$ identifiserer de to vertikale sidene (i rødt) på R kan videre visualiseres i følgende figur:



Den resulterende figuren visualiserer T, og vi ser at den ser ut som en smultring. I matematikk kalles dette objektet en *torus* og vi har allerede stiftet bekjentskap med det da vi snakket om Poincarés formodning i §1.4:



EN TORUS (KILDE: WIKIPEDIA, PUBLIC DOMAIN)

Vi kan definere en topologi $\mathcal{T}$ på torusen T ved å bruke avbildningen f:

$$U \in \mathcal{T} \iff f^{-1}(U) \text{ er åpen i } \mathbb{R}^2. \tag{322}$$

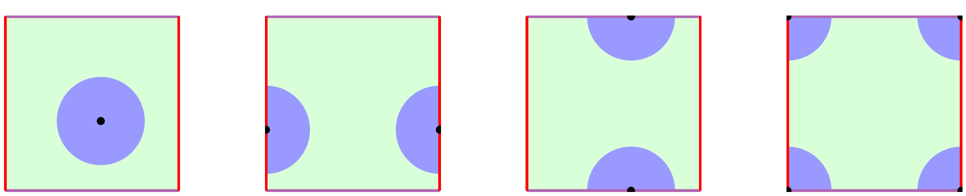
Med andre ord:

$$U \text{ er åpen i } T \iff f^{-1}(U) \text{ er åpen i } \mathbb{R}^2.$$

Vi overlater til Oppgave 10.77 å vise at dette virkelig definerer en topologi på T, samt at funksjonen $f : \mathbb{R}^2 \to T$ er kontinuerlig med denne topologien.

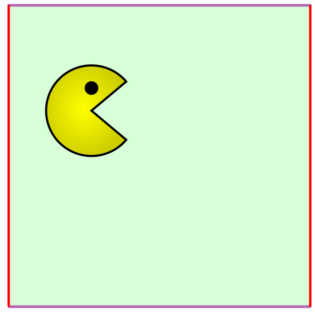
Avbildningen $f|_R : R \to T$ gjør oss i stand til å bruke rektanglet R som modell på torusen T. Vi må bare huske på at punktene på de horisontale sidene av R er identifisert på T, og likeledes punktene på de vertikale sidene. Figuren nedenfor viser inversbildet i R av typiske omegner av punkter på T, når vi betegner ekvivalensklasser av $(x, y) \in \mathbb{R}^2$ som vanlig med $[x, y]$:

(I) Omegn om $[x, y]$, $0 < x < 1$, $0 < y < 1$.
(II) Omegn om $[0, y] = [1, y]$, $0 < y < 1$.
(III) Omegn om $[x, 0] = [x, 1]$, $0 < x < 1$.
(IV) Omegn om $[0, 0] = [1, 0] = [1, 1] = [0, 1]$.

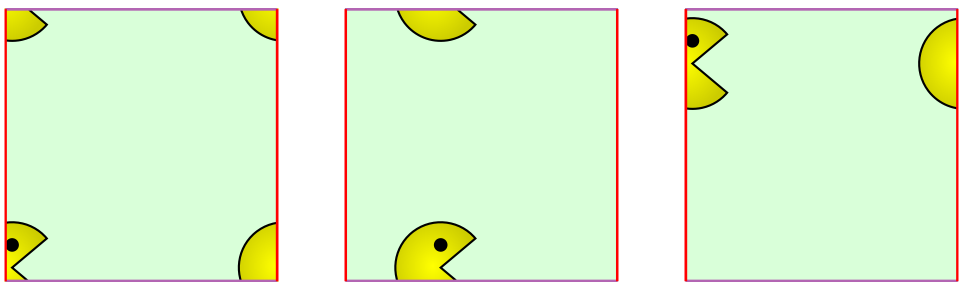


(I) (II) (III) (IV)

Torusen blir også kalt for "pacmans univers". Pacman er et arkadespill fra 1980, som regnes som en klassiker, og har blitt et symbol på 1980-årenes populærkultur. I spillet styrer man en gul sirkel med munn, som beveger seg på en flate og spiser dotter og premier som den kommer over, samtidig som den forsøker å ikke bli spist opp av fire spøkelser som også beveger seg rundt på flaten. På dataskjermen virker pacman å leve på et kvadrat eller rektangel:



Men pacmans verden er mer interessant enn som så, for hvis vi beveger den utenfor en av sidene på kvadratet, dukker den opp igjen på motsatt side, som vist i figurene nedenfor:



Dette betyr at pacman egentlig lever på torusen.

MERKNAD 10.10.13. Topologien definert på torusen i Eksempel 10.10.12 kan generaliseres: La $(X, \mathcal{T})$ være et topologisk rom slik at X er utstyrt med en ekvivalensrelasjon $\sim$ og la $\overline{X} = X/\sim$, mengden av alle ekvivalensklassene. La $f : X \to \overline{X}$ være den sedvanlige avbildningen som avbilder x på sin ekvivalensklasse. Da er samlingen $\overline{\mathcal{T}}$ av delmengder av $\overline{X}$ definert ved

$$U \in \overline{\mathcal{T}} \iff f^{-1}(U) \in \mathcal{T}, \tag{323}$$

eller ekvivalent

$$U \text{ er åpen i } \overline{X} \iff f^{-1}(U) \text{ er åpen i } X,$$

en topologi på $\overline{X}$ (jf. Oppgave 10.79). Denne topologien kalles *kvotienttopologien*[21].

Kvotienttopologien gjør oss for eksempel i stand til å gjøre det *projektive n-rommet* $\mathbb{P}^n_{\mathbb{R}}$ over $\mathbb{R}$ (jf. Eksempel 2.4.17) til et topologisk rom, via den euklidske topologien (eller andre) på $\mathbb{R}^{n+1} \setminus (0, \dots, 0)$.

[21] *Quotient topology* på engels.

Eksempel 10.10.14 (Endelig komplement topologien). La X være en mengde. La $\mathcal{T}$ være samlingen av delmengder $U \subset X$ slik at $X - U$ er lik X, eller er endelig eller tom. Vi overlater til Oppgave 10.76 å verifisere at dette definerer en topologi på X, kalt *endelig komplement topologien*[22], som ikke er Hausdorff dersom X er uendelig.

Eksempel 10.10.15 (Tellbart komplement topologien). La X være en uendelig mengde. La $\mathcal{T}$ være samlingen av delmengder $U \subset X$ slik at $X - U$ er enten lik X, eller er tom, endelig eller tellbar (jf. Definisjon 7.9.3). Vi overlater til Oppgave 10.81(a) å verifisere at dette definerer en topologi på X, kalt *tellbart komplement topologien*[23], som ikke er Hausdorff dersom X er overtellbar (jf. Definisjon 7.9.3).

Eksempel 10.10.16 (Zariski-topologien). La K være en kropp (tenk gjerne på $\mathbb{R}$ eller $\mathbb{C}$) og betrakt det kartesiske produktet K^n, som er mengden av alle ordnede n-tupler $(a_1, \ldots, a_n)$ med $a_i \in K$ for alle $i \in \{1, \ldots, n\}$. For enhver mengde S bestående av polynomer av n variable med koeffisienter i K, kan vi definere delmengden

$$V(S) = \{(x_1, \ldots, x_n) \in K^n \mid f(x_1, \ldots, x_n) = 0 \text{ for alle } f \in S\}.$$

Merk at $V(\emptyset) = K^n$ og $V(\{1\}) = \emptyset$.

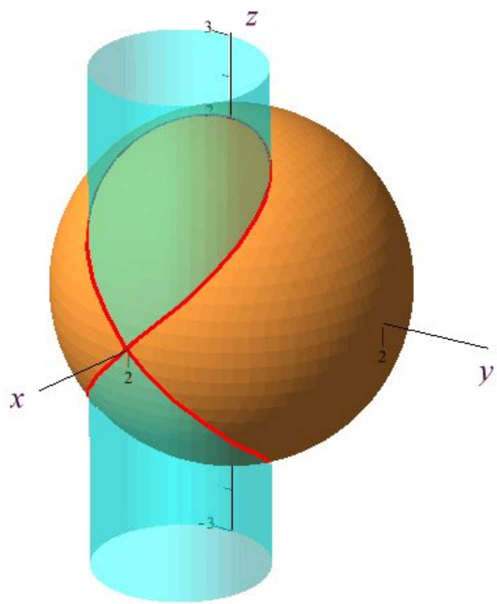
For eksempel, hvis $K = \mathbb{R}$, $n = 2$ og $S = \{y - x^2\}$, er $V(S)$ lik mengden av nullpunkter til polynomet $y - x^2$, som er parabelen gitt ved ligningen $y = x^2$. Hvis $K = \mathbb{R}$, $n = 2$ og $S = \{x^2 + y^2 + z^2 - 1, x^2 + y^2 - x\}$, da er $V(S)$ lik mengden av alle punkter som samtidig er nullpunkter til polynomene $x^2 + y^2 + z^2 - 1$ og $x^2 + y^2 - x$. Nullpunktsmengden $V(\{x^2 + y^2 + z^2 - 1\})$ til det første polynomet beskriver en kuleflate i $\mathbb{R}^3$, mens nullpunktsmengden $V(\{x^2 + y^2 - x\})$ til det andre polynomet beskriver en sylinder i $\mathbb{R}^3$. Mengden $V(S) \subset \mathbb{R}^3$ er snittet mellom disse to geometriske objektene, med andre ord har vi:

$$\begin{aligned} V(S) &= V(\{x^2 + y^2 + z^2 - 1, x^2 + y^2 - x\}) \\ &= V(\{x^2 + y^2 + z^2 - 1\}) \cap V(\{x^2 + y^2 - x\}), \end{aligned}$$

som er kurven i rødt formet som et 8-tall beskrevet i figuren nedenunder.

[22] *Finite complement topology* på engelsk.

[23] *Countable complement topology* eller *cocountable topology* på engelsk.



$V(\{x^2 + y^2 + z^2 - 1, x^2 + y^2 - x\})$ I RØDT.
(KILDE: WIKIPEDIA, CREATIVE COMMONS ATTRIBUTION-SHARE ALIKE 3.0 UNPORTED LICENSE.)

Nå definerer vi en samling $\mathcal{T}$ av delmengder av K^n bestående av alle delmengder på formen $K^n \setminus V(S)$, for S en samling polynomer. Merk at $K^n \setminus V(\emptyset) = \emptyset$, mens $K^n \setminus V(\{1\}) = K^n$. Det går an å vise at dette definerer en topologi på K^n, kalt *Zariski-topologien*, etter den russiskfødte amerikanske matematikeren Oscar Zariski (1899–1986). Denne topologien spiller en sentral rolle i fagfeltet *algebraisk geometri*, som grovt sagt kan sies å studere egenskaper av geometriske objekter definert ved hjelp av polynomer.

Hvis $n = 1$, er S en samling av polynomer av én variabel. Siden disse har endelig mange nullpunkter og vi for hver endelig mengde $\{a_1, \ldots, a_k\} \subset K$ har at polynomet $f(x) = (x - a_1) \cdots (x - a_k)$ har nøyaktig $\{a_1, \ldots, a_k\}$ som nullpunktsmengde, ser vi at Zariski-topologien i dette tilfellet er lik endelig komplement topologien (Eksempel 10.10.14) på K. Spesielt er Zariski-topologien ikke Hausdorff dersom K er en uendelig kropp. (Dette gjelder også for $n > 1$).

Følger i topologiske rom. Vi kan overføre flere av begrepene vi har i metriske rom til topologiske rom. Vi starter med konvergens av følger:

Definisjon 10.10.17: Konvergens av følge

La X være et topologisk rom. Følgen $\{x_n\}$ i X *konvergerer* mot $x \in X$ dersom det for enhver omegn V om x finnes en $N \in \mathbb{Z}^+$ slik at $x_n \in V$ for alle $n \geq N$.

EKSEMPEL 10.10.18. La X være en uendelig mengde utstyrt med *tellbart komplement topologien* (jf. Eksempel 10.10.15), det vil si at de åpne mengdene er delmengder $U \subset X$ slik at $X - U$ er enten lik X, eller er tom, endelig eller tellbar. Da vil en følge $\{x_n\}$ i X konvergere mot $x \in X$ hvis og bare hvis det finnes en $N \in \mathbb{Z}^+$ slik at $x_n = x$ for alle $n \geq N$, jf. Oppgave 10.81(b).

I motsetning til i metriske rom (jf. Setning 10.2.2), kan en følge i et topologisk rom konvergere mot forskjellige punkter, som neste eksempel viser:

EKSEMPEL 10.10.19. Betrakt mengden $X = \{a, b, c\}$ med åpne mengder $\emptyset, \{a\}, \{b\}, \{a, b\}, X$ (som er topologien gitt i figur (3) i Eksempel 10.10.11). Da vil følgen $\{c, c, c, c, c, \ldots\}$ konvergere både mot a, b og c, siden den eneste omegnen om c er mengden X selv, og denne inneholder også a og b.

Imidlertid garanterer Hausdorff-betingelsen (Definisjon 10.10.7) at grenser er entydige:

Setning 10.10.20: Entydighet av grenser

Dersom en følge $\{x_n\}$ i et Hausdorff rom konvergerer mot x og x', da er $x = x'$.

I slike tilfeller skriver vi som før $\lim_{n\to\infty} x_n = x$ eller "$x_n \to x$ når $n \to \infty$" og kaller x for *grensen* til følgen.

BEVIS FOR SETNING 10.10.20
Anta, for å få en selvmotsigelse, at $x \neq x'$. Siden X er Hausdorff, finnes det en omegn V_x om x og en omegn $V_{x'}$ om x' slik at $V_x \cap V_{x'} = \emptyset$. Per definisjon av konvergens finnes en $N \in \mathbb{Z}^+$ slik at $x_n \in V_x$ for alle $n \geq N$ og en $N' \in \mathbb{Z}^+$ slik at $x_n \in V_{x'}$ for alle $n \geq N'$. Men da vil $x_n \in V_x \cap V_{x'}$ for alle $n \geq \max\{N, N'\}$, som motsier at $V_x \cap V_{x'} = \emptyset$. $\square$

Som vi nevnte tidligere er ikke egenskapen *lukkethet* i et topologisk rom ekvivalent med egenskapen av å inneholde alle sine følgegrenser, i motsetnig til i metriske rom, hvor dette er én av flere ekvivalente definisjoner (jf. Definisjon 10.3.8). Neste setning og det påfølgende eksemplet viser at i et topologisk rom gjelder alltid den ene implikasjonen mellom egenskapene, men ikke alltid den andre:

Setning 10.10.21

La X være et topologisk rom og $A \subset X$ en lukket delmengde. Hvis $\{x_n\}$ er en følge i A slik at x_n konvergerer mot $x \in X$, da er $x \in A$. (Med andre ord: A inneholder sine følgegrenser).

Bevis
Vi argumenterer ved selvmotsigelse og antar at A er lukket og at $\{x_n\}$ er en følge i A slik at x_n konvergerer mot $x \in X \setminus A$. Siden $X \setminus A$ er åpen (per definisjon av at A er lukket), eksisterer det en omegn V om x slik at $V \subset X \setminus A$, det vil si, $V \cap A = \emptyset$ (jf. Oppgave 2.6). Dette medfører at $x_n \notin V$ for alle $n \in \mathbb{Z}^+$, men dette motsier at $x_n \to x$. $\square$

Eksempel 10.10.22. Egenskapen av å inneholde alle sine følgegrenser impliserer ikke nødvendigvis lukkethet: Betrakt mengden $X = \{a, b, c\}$ med åpne mengder $\emptyset, \{a\}, X$ (som er topologien gitt i figur (2) i Eksempel 10.10.11), og delmengden $A = \{a\}$. Den eneste følgen i A er den konstante følgen $\{a, a, a, \ldots\}$, som konvergerer mot a, men ikke mot b eller c, siden A selv er åpen og innholder verken b eller c. Derfor inneholder A alle sine følgegrenser. men A er ikke lukket, siden $X \setminus A = \{b, c\}$, som ikke er åpen.

Merknad 10.10.23. Rommet X i siste eksempel er ikke-Hausdorff. Det finnes også ekempler på Hausdorffrom som inneholder alle sine følgegrenser, men som ikke er lukket, men disse er litt mer kompliserte, så vi skal ikke gå innpå disse her.

Kontinuitet i topologiske rom. Definisjonen av kontinuitet i topologiske rom oppnås ved å kopiere "åpne-mengder kriteriet" fra Setning 10.4.5:

Definisjon 10.10.24: Kontinuitet

La $f : X \to Y$ være en funksjon mellom to topologiske rom.

Vi sier at f er *kontinuerlig* i $a \in X$ hvis det for enhver omegn V om $f(a)$ finnes en omegn U om a slik at $U \subset f^{-1}(V)$ (ekvivalent: $f(U) \subset V$).

Vi sier at f er *kontinuerlig* hvis $f^{-1}(V) \subset X$ er åpen for alle åpne mengder $V \subset Y$.

Merknad 10.10.25. Det er ikke vanskelig ut fra forrige definisjon å vise at f er kontinuerlig hvis og bare hvis f er kontinuerlig i alle $a \in X$.

Vår intuitive forståelse av kontinuitet kan ikke lenger formuleres ved hjelp av avstander, men kan formuleres som at f er kontinuerlig i a dersom vi for enhver omegn om $f(a)$ kan sørge for at alle funksjonsverdier $f(x)$ er i denne omegnen for alle x i en passende omegn om a.

MERKNAD 10.10.26. Det vi sa i Merknad 10.4.2 om at kontinuitet er bevart ved utvidelse og restriksjon av kodomenet og ved restriksjon av domenet gjelder fremdeles (jf. Oppgave 10.85).

Eksempel 10.4.3 generaliseres til topologiske rom:

EKSEMPEL 10.10.27 (Konstante funksjoner er kontinuerlige). La X og Y være topologiske rom og $f : X \to Y$ en funksjon slik at $f(x) = y_0$ for alle $x \in X$, for en fiksert $y_0 \in Y$. Vis at f er kontinuerlig.

LØSNING. For enhver åpen mengde $V \subset Y$ er

$$f^{-1}(V) = \begin{cases} X, & \text{hvis } x \in V, \\ \emptyset, & \text{hvis } x \notin V, \end{cases}$$

som begge er åpne delmengder av X per definisjon av en topologi. □

EKSEMPEL 10.10.28. Betrakt mengden $X = \{a, b, c\}$. Man kan lett verifisere at

$$\mathcal{T} = \{\emptyset, \{a\}, \{b\}, \{a, b\}, \{a, c\}, X\}$$

er en topologi på X. La $f : X \to \mathbb{R}$ være gitt ved $f(a) = 1, f(b) = 2, f(c) = 3$ og betrakt $\mathbb{R}$ som et topologisk rom via sin euklidske metrikk.

Vis at f er diskontinuerlig i b og c, men kontinuerlig i a.

LØSNING. Funksjonen f er diskontinuerlig i b siden for eksempel intervallet $(1, 4)$ er en åpen delmengde av $\mathbb{R}$ som inneholder $2 = f(b)$, mens det inverse bildet $f^{-1}((1, 4)) = \{b, c\}$, som ikke er åpen i X.

Funksjonen f er diskontinuerlig i c siden for eksempel intervallet $(2, 4)$ er en åpen delmengde av $\mathbb{R}$ som inneholder $3 = f(c)$, mens det inverse bildet $f^{-1}((2, 4)) = \{c\}$, som ikke er åpen i X.

Funksjonen er imidlertid kontinuerlig i a, siden vi for enhver åpen delmengde $V \subset \mathbb{R}$ som inneholder $1 = f(a)$ har

$$f^{-1}(V) = \begin{cases} X, & \text{hvis } 2, 3 \in V, \\ \{a, b\} & \text{hvis } 2 \in V, 3 \notin V, \\ \{a, c\} & \text{hvis } 2 \notin V, 3 \in V, \\ \{a\} & \text{hvis } 2, 3 \notin V, \end{cases}$$

og alle disse er åpne delmengder av X. □

Eksempel 10.10.29. La $f : X \to Y$ være en funksjon mellom topologiske rom slik at X er utstyrt med den diskrete topologien (Eksempel 10.10.10). Siden alle delmengder av X er per definisjon åpne, er f automatisk kontinuerlig. Dette generaliserer Eksempel 10.7.2.

Vi har fremdeles et resultat som sier at *sammensetningen av to kontinuerlige funksjoner er kontinuerlig* (jf. Setning 10.4.6).

Setning 10.10.30: Sammensetning av kontinuerlige funksjoner

Anta $g : X \to Y$ og $f : Y \to Z$ er funksjoner mellom topologiske rom slik at g er kontinuerlig i punktet $a \in X$ og f er kontinuerlig i punktet $g(a) \in Y$. Da er den sammensatte funksjonen $f \circ g : X \to Z$ kontinuerlig i a.

Bevis
La $x \in X$ og sett $y = g(x) \in Y$ og $z = f(y) \in Z$. La V være en omegn om z i Z. Siden f er kontinuerlig, er $f^{-1}(V)$ åpen i Y. Siden g er kontinuerlig, er $g^{-1}(f^{-1}(V))$ åpen i X. Siden $g^{-1}(f^{-1}(V)) = (f \circ g)^{-1}(V)$ (ved Oppgave 2.22) og $x \in (f \circ g)^{-1}(V)$, er vi ferdig. □

Vi har dessverre ikke lenger et "følge-kriterium" for kontinuitet, som vi har i metriske rom (Setning 10.4.4). Den ene retningen av følge-kriteriet er imidlertid sant, en kontinuerlig funksjon bevarer grenser av følger:

Setning 10.10.31

La $f : X \to Y$ være en funksjon mellom to topologiske rom. Hvis f er kontinuerlig i $a \in X$, og $\{x_n\}$ konvergerer mot a i X, da konvergerer $\{f(x_n)\}$ mot $f(a)$.

Bevis
La V være en vilkårlig omegn om $f(a)$. Per definisjon av kontinuitet er $f^{-1}(V)$ åpen i X, og inneholder a. Siden $\{x_n\}$ konvergerer mot a, vil det per definisjon finnes en $N \in \mathbb{Z}^+$ slik at $x_n \in f^{-1}(V)$ for alle $n \geq N$. Dermed er $f(x_n) \in V$ for alle $n \geq N$. Per definisjon konvergerer derfor følgen $\{f(x_n)\}$ mot $f(a)$. □

Det er enkelt å vise at det motsatte av Setning 10.10.31 ikke alltid holder; en funksjon som bevarer grenser av følger trenger ikke være kontinuerlig:

Eksempel 10.10.32. La X være $\mathbb{R}$ utstyrt med *tellbart komplement topologien* (jf. Eksempel 10.10.15), det vil si at de åpne mengdene er delmengder $U \subset X$ slik at $X - U$ er enten lik X, eller er tom, endelig eller tellbar. La

Y være $\mathbb{R}$ utstyrt med den *diskrete topologien* (jf. Eksempel 10.10.10), det vil si at alle delmengder av X er åpne.

La $f : X \to Y$ være identitetsfunksjonen. Denne er ikke kontinuerlig i noe punkt $a \in X$, siden $\{a\}$ er åpen i Y, men $f^{-1}(\{a\}) = \{a\}$ ikke er åpen i X, siden komplementet $X - \{a\} = \mathbb{R} \setminus \{a\}$ er overtellbart.

La $\{x_n\}$ være en følge i X som konvergerer mot a. Da må alle $x_n = a$ for stor nok n ved Eksempel 10.10.18. Dermed er også alle $f(x_n) = f(a)$ for stor nok n, slik at $\{f(x_n)\}$ konvergerer mot $a = f(a)$.

Sammenhengende topologiske rom. Begrepet "sammenhengende" overføres til topologiske rom. Definisjonen 10.5.1 forblir den samme ved å erstatte "metrisk rom" med "topologisk rom":

Definisjon 10.10.33: Sammenhengende rom

Et topologisk rom X sies å være *ikke-sammenhengende*[a] dersom vi kan skrive $X = A \cup B$, der A og B er ikke-tomme åpne disjunkte delmengder av X.

Ellers sies X å være *sammenhengende*[b].

Ekvivalent er X sammenhengdende dersom de eneste delmengdene av X som er både åpne og lukket i X er X og $\emptyset$.

[a] *Disconnected* på engelsk.
[b] *Connected* på engelsk.

Igjen overlater vi til Oppgave 10.88 å vise at de to definisjonene av "sammenhengende" er ekvivalente.

Setning 10.5.8 gjelder fremdeles for topologiske rom; *en kontinuerlig funksjon bevarer sammenhengende mengder*:

Setning 10.10.34

La $f : X \to Y$ være en kontinuerlig funksjon mellom topologiske rom. Hvis X er sammenhengende, da er også $f(X) \subset Y$ sammenhengende.

Bevis
Likt beviset for Setning 10.5.8. □

Eksempel 10.10.35. Vis at torusen T (jf. Eksempel 10.10.12) er sammenhengende.

Løsning. Funksjonen $f : \mathbb{R}^2 \to T$ fra Eksempel 10.10.12 er kontinuerlig (ved Oppgave 10.77) og surjektiv, og $\mathbb{R}^2$ er sammenhengende ved Eksempel 10.5.5. Dermed er T sammenhengende ved Setning 10.10.34. □

EKSEMPEL 10.10.36. La $X = \{a, b, c\}$ med topologi $\mathcal{T} = \{\emptyset, \{a\}, \{b, c\}, X\}$. (Det er lett å verifisere at dette er en topologi.) Da er X ikke-sammenhengende, fordi $X = \{a\} \cup \{b, c\}$ og $\{a\}$ og $\{b, c\}$ er åpne, disjunkte delmengder.

MERKNAD 10.10.37. Alt det vi sa i Merknad 10.5.11 gjelder om vi erstatter "metrisk rom" med "topologisk rom".

Kompakte topologiske rom. Når det gjelder definisjonen av "kompakthet", må vi være litt mer forsiktige. Betingelsene (i) og (ii) i Definisjon 10.6.3 er nemlig ikke lenger ekvivalente for topologiske rom. Det finnes både eksempler på topologiske rom som oppfyller (i), men ikke (ii), og omvendt, uten at vi skal gå nærmere inn på dette. Betingelsen som har vist seg å være mest interessant, er (ii), og dette blir utgangspunktet for definisjonen i et topologisk rom:

Definisjon 10.10.38: Kompakt mengde

La X være et topologisk rom.

Vi sier at X er *kompakt* dersom det for enhver samling $\{U_\alpha\}$ av åpne delmengder av X slik at $X = \cup_\alpha U_\alpha$, finnes endelig mange $U_{\alpha_1}, \dots, U_{\alpha_n}$ i $\{U_\alpha\}$ slik at $X = U_{\alpha_1} \cup \cdots \cup U_{\alpha_n}$

Med denne definisjonen overføres Setning 10.6.11 til topologiske rom; *en kontinuerlig funksjon bevarer kompakte mengder*:

Setning 10.10.39

La $f : X \to Y$ være en kontinuerlig avbildning mellom topologiske rom. Hvis X er kompakt, da er $f(X)$ kompakt.

BEVIS
Likt variant II av beviset for Setning 10.6.11. □

EKSEMPEL 10.10.40. Vis at torusen T (jf. Eksempel 10.10.12) er kompakt.

LØSNING. La R og $f : \mathbb{R}^2 \to T$ være som i Eksempel 10.10.12. Siden f er kontinuerlig (ved Oppgave 10.77), er også restriksjonen $f|_R : R \to T$ kontinuerlig (ved Merknad 10.10.26). Vi så i Eksempel 10.10.12 at $f|_R$ er surjektiv. Siden R er kompakt ved Heine-Borel teoremet 10.6.8, er T kompakt ved Setning 10.10.39. □

I et metrisk rom har vi sett at en kompakt delmengde må være lukket og begrenset (Setning 10.6.9). Begrensethet gir ikke mening for et topologisk rom, men lukkethet gjør. Legger vi til Hausdorff-betingelsen (Definisjon 10.10.7), er det fremdeles sant at en kompakt delmengde må være lukket:

Setning 10.10.41

La $K \subset X$ være en kompakt delmengde av et Hausdorffrom X. Da er K lukket.

BEVIS
Overlates til Oppgave 10.90. □

Følgende eksempel viser at Hausdorff-betingelsen er nødvendig:

EKSEMPEL 10.10.42. Betrakt mengden $X = \{a, b, c\}$ med åpne delmengder $\emptyset, \{a\}, \{b\}, \{a, b\}, X$ fra Eksemplene 10.10.11 (figur (3)) og 10.10.19. Siden vi kun har endelig mange åpne delmengder i X, er alle delmengder kompakte. (Mer generelt er alle endelige delmengder av topologiske rom kompakte.) Spesielt er delmengden $\{a\}$ kompakt. Men den er ikke lukket, siden komplementet $X - \{a\} = \{b, c\}$ ikke er åpent.

Oppgaver

Oppgaver til §10.1

OPPGAVE 10.1. Vis at de to betingelsene (i) og (ii) i Definisjon 10.1.4 virkelig er ekvivalente.

OPPGAVE 10.2. Begrunn at en delmengde R av $\mathbb{R}^2$ er begrenset hvis og bare hvis vi kan slå en sirkel i $\mathbb{R}^2$ som inneholder mengden R.

OPPGAVE 10.3. Avgjør om følgende funksjoner $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ definerer en metrikk:

(a) $d(x, y) = (x - y)^2$,
(b) $d(x, y) = \sqrt{|x - y|}$,
(c) $d(x, y) = |x^2 - y^2|$,
(d) $d(x, y) = |x - 5y|$.

OPPGAVE 10.4. Vis at dersom d er en metrikk på en mengde X, da er også funksjonen $\rho : X \times X \to \mathbb{R}$ definert ved $\rho(x, y) = \frac{d(x,y)}{1+d(x,y)}$ en metrikk på X.

OPPGAVE 10.5. Vis at d i (288) oppfyller betingelsene til en metrikk. For å vise trekantulikheten, bruk *Cauchy-Schwarz ulikheten* fra Oppgave 8.17.

OPPGAVE 10.6. La $\mathbb{K} = \mathbb{R}$ eller $\mathbb{C}$. Et $\mathbb{K}$-vektorrom V (jf. Definisjon 4.5.5) sies å være et *indreproduktrom*[24] dersom det finnes en funksjon

$$\begin{array}{ccc} V \times V & \longrightarrow & \mathbb{K} \\ (\mathbf{x}, \mathbf{y}) & \mapsto & \langle \mathbf{x}, \mathbf{y} \rangle \end{array}$$

(kalt *indreprodukt*) som oppfyller følgende for alle $c \in \mathbb{K}$, $\mathbf{x}, \mathbf{y} \in V$:

- $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$ (kompleks konjugert);
- $\langle c\mathbf{x}, \mathbf{y} \rangle = c\langle \mathbf{x}, \mathbf{y} \rangle$;
- $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$.

(a) Utled egenskapene $\langle \mathbf{x}, c\mathbf{y} \rangle = \bar{c}\langle \mathbf{x}, \mathbf{y} \rangle$ og $\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$ fra aksiomene.

(b) Vis at indreproduktet på $\mathbb{R}^n$ i Definisjon 10.1.7 gjør euklidsk rom $\mathbb{R}^n$ til et indreproduktrom.

(c) Vis at

$$\langle (z_1, \ldots, z_n), (w_1, \ldots, w_n) \rangle = \sum_{j=1}^{n} z_j \overline{w_j}$$

definerer et indreprodukt på $\mathbb{C}^n$ (som er et $\mathbb{R}$-vektorrom ved Oppgave 8.8).

OPPGAVE 10.7. La V være et *indreproduktrom* over $\mathbb{K} = \mathbb{R}$ eller $\mathbb{C}$, jf. Oppgave 10.6. Definér funksjonen

$$\begin{array}{ccc} V & \longrightarrow & \mathbb{R} \\ \mathbf{x} & \mapsto & \|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}. \end{array}$$

(a) Vis at funksjonen er veldefinert ved å vise at $\langle \mathbf{x}, \mathbf{x} \rangle \in \mathbb{R}^{\geq 0}$ for alle $\mathbf{x} \in V$.

(b) Vis at funksjonen

$$\begin{array}{ccc} V \times V & \longrightarrow & \mathbb{R} \\ (\mathbf{x}, \mathbf{y}) & \mapsto & \|\mathbf{x} - \mathbf{y}\| \end{array}$$

definerer en metrikk på V. (Dette generaliserer den euklidske metrikken til alle indreproduktrom, jf. Oppgave 10.6(b).)

(c) Hvilken metrikk på $\mathbb{C}^n$ blir indusert av indreproduktet fra Oppgave 10.6(c)?

Oppgaver til §10.2

OPPGAVE 10.8. Bevis Setning 10.2.9.

OPPGAVE 10.9. (a) Hvis at en følge $\{\mathbf{x}_k\} = \{(x_{k,1}, \ldots, x_{k,n})\}$ i det euklidske rommet $\mathbb{R}^n$ er Cauchy hvis og bare hvis hver følge $\{x_{k,j}\}$ er Cauchy i $\mathbb{R}$ (jf. Oppgave 8.29).

[24] *Inner product space* på engelsk.

(b) Bevis Setning 10.2.12. (Hint: ta utgangspunkt i beviset på Setning 8.5.12.)

Oppgave 10.10. Vis at A er tett i et metrisk rom X hvis og bare hvis ethvert punkt $x \in X$ er grensen til en følge $\{x_n\}$ i A.

Oppgave 10.11. Vis at en delmengde av et komplett metrisk rom er komplett hvis og bare hvis den er lukket.

Oppgave 10.12. Definisjon 9.8.7 av en *kontraksjon* generaliseres til et vilkårlig metrisk rom X: En funksjon $f : X \to X$ kalles en *kontraksjon* dersom det finnes en $K \in \mathbb{R}$ slik at $K < 1$ og $d\,(f(a), f(b)) \leq Kd(a, b)$ for alle $a, b \in X$.

Vis at *kontraksjonsteoremet* 9.8.11 fremdeles gjelder dersom vi erstatter "$X \subset \mathbb{C}$" med "X et *komplett* metrisk rom".

Oppgave 10.13. Vis følgende ekvivalente karakterisering av konvergens, som generaliserer resultatet fra Oppgave 7.16: *En følge $\{a_n\}$ i et metrisk rom konvergerer mot a hvis og bare hvis enhver omegn om a inneholder alle bortsett fra endelig mange a_n.*

Oppgave 10.14. Definisjon 9.10.1 generaliseres naturlig til metriske rom: La X være et metrisk rom og $A \subset X$ en delmengde. Et punkt $x \in X$ er et *opphopningspunkt* eller *grensepunkt* til A dersom vi for alle $\delta > 0$ har at $A \cap N_\delta(x)^* \neq \emptyset$.

Vis at Setning 9.10.2 generaliseres til metriske rom: $x \in X$ er et opphopningspunkt til A hvis og bare hvis det finnes en følge $\{x_n\}$ slik at $x_n \in A$, $x_n \neq x$ for alle n og $x_n \to x$.

Oppgaver til §10.3

Oppgave 10.15. Fullfør beviset for Setning 10.3.4 ved å følge indikasjonene i beviset. (Hint: bruk *trekantulikheten*.)

Oppgave 10.16. Verifiser at $A = G \cap Y$, som hevdet i beviset for Setning 10.3.6.

Oppgave 10.17. La X være unionen av koordinataksene i $\mathbb{R}^2$. Er X lukket eller åpen, eller ingen av delene, i $\mathbb{R}^2$?

Oppgave 10.18. Avgjør om mengden $S = \left\{ \frac{1}{n} \;\middle|\; n \in \mathbb{Z}^+ \right\}$ er lukket eller åpen, eller ingen av delene, i

(a) $\mathbb{R}$,
(b) $\mathbb{R}^+$.

Oppgave 10.19. Avgjør om mengden

$$X = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \;\text{ og }\; y > 0\}$$

er lukket eller åpen, eller ingen av delene, i

(a) $\mathbb{R}^2$,
(b) $\mathbb{R} \times \mathbb{R}^+ = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$,
(c) $\mathbb{S}^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$. (Dette er enhetssirkelen i $\mathbb{R}^2$.)

Oppgave 10.20. Bevis Setning 10.3.12.

Oppgave 10.21. La X være et metrisk rom og $x \in X$. Vis at $\{x\}$ er lukket i X.

Oppgave 10.22. Vis (292).

Oppgave 10.23. La X være en mengde, B en delmengde av X og $\{A_\alpha\}$ en samling delmengder av X. Vis følgende generaliseringer mengdelovene fra §2.1:

(a) $(\bigcup_\alpha A_\alpha) \cap B = \bigcup_\alpha (A_\alpha \cap B)$;
(b) $(\bigcap_\alpha A_\alpha) \cup B = \bigcap_\alpha (A_\alpha \cup B)$.

Oppgave 10.24. La $\{A_\alpha\}$ være en samling delmengder av en mengde X. Vis følgende generaliseringer av De Morgans lover fra §2.1:

(a) $X \setminus \bigcup_\alpha A_\alpha = \bigcap_\alpha (X \setminus A_\alpha)$;
(b) $X \setminus \bigcap_\alpha A_\alpha = \bigcup_\alpha (X \setminus A_\alpha)$.

Oppgave 10.25. Bevis del (iii) og (iv) av Setning 10.3.26, ved å bruke definisjon av lukket mengde eller (i) og (ii) i setningen pluss resultatet fra Oppgave 10.24 (eller begge deler).

Oppgaver til §10.4

Oppgave 10.26. Vis at hvis $f : X \to Y$ er en kontinuerlig funksjon mellom metriske rom og $Z \subset X$ er et underrom, da er også restriksjonen $f|_Z : Z \to Y$ kontinuerlig.

Oppgave 10.27. (a) La X være et underrom av det metriske rommet Y og $i : X \to Y$ inklusjonsavbildningen. Vis at i er kontinuerlig.

(b) La X, Y, Z være metriske rom, $Y \subset Z$ et underrom og $i : Y \to Z$ inklusjonsavbildningen. Vis at $f : X \to Y$ er kontinuerlig hvis og bare hvis $i \circ f : X \to Z$ er kontinuerlig.

Oppgave 10.28. Bevis Setning 10.4.5.

Oppgave 10.29. Bevis Setning 10.4.6 ved å bruke *åpne-mengder-kriteriet* (Setning 10.4.5).

Oppgave 10.30. Utstyr $\mathbb{R}^n$ med den euklidske metrikken.

(a) Vis at $\big||\mathbf{x}| - |\mathbf{y}|\big| \leq |\mathbf{x} - \mathbf{y}|$.
(b) Vis at *normfunksjonen*

$$\begin{array}{ccc} \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ \mathbf{x} & \mapsto & |\mathbf{x}| \end{array}$$

er kontinuerlig.

Oppgave 10.31. Begrunn at funksjonen

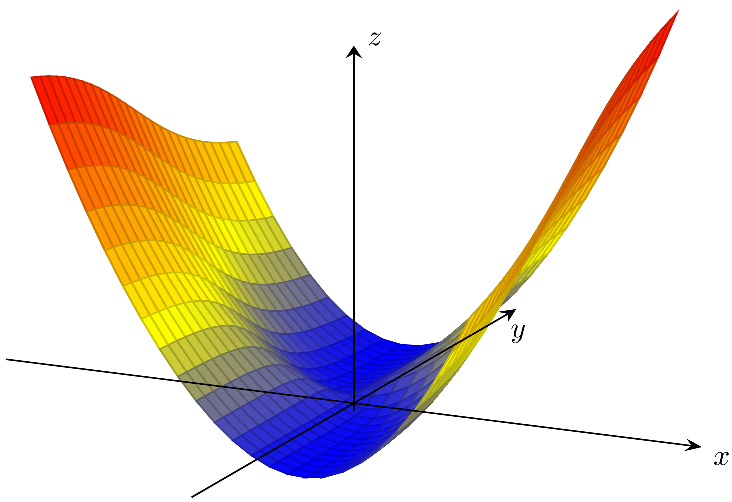
$$\begin{array}{rcl} f:\mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \mapsto & \begin{cases} \frac{\sin(xy^2)}{xy}, & \text{hvis } \ x \neq 0 \ \text{ og } \ y = 0, \\ y, & \text{hvis } \ x = 0 \ \text{ eller } \ y = 0 \end{cases} \end{array}$$

er kontinuerlig i $(0,0)$. (Hint: bruk resultatet fra Oppgave 9.4).

Oppgave 10.32. Finn $c \in \mathbb{R}$ som gjør funksjonen

$$\begin{array}{rcl} f:\mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \mapsto & \begin{cases} \frac{x^2+3xy+\sqrt{x^2+y^2}}{\sqrt{x^2+y^2}}, & \text{hvis } \ (x,y) \neq (0,0) \\ c, & \text{hvis } \ (x,y) = (0,0) \end{cases} \end{array}$$

kontinuerlig i $(0,0)$.

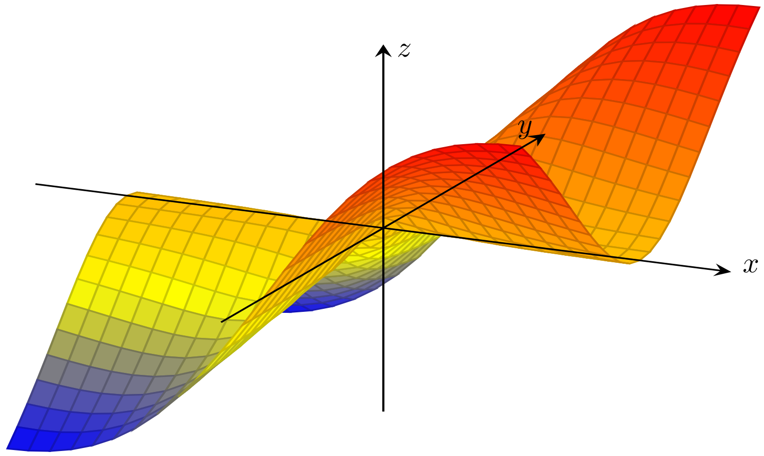


Grafen til funksjonen i Oppgave 10.32

Oppgave 10.33. Vis at det ikke finnes noen $c \in \mathbb{R}$ som gjør funksjonen

$$\begin{array}{rcl} f:\mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \mapsto & \begin{cases} \frac{xy^2}{x^2+y^4}, & \text{hvis } \ (x,y) \neq (0,0) \\ c, & \text{hvis } \ (x,y) = (0,0) \end{cases} \end{array}$$

kontinuerlig i $(0,0)$. (Hint: velg følger langs parabelen $x = ly^2$ for varierende l og argumentér som i Eksempel 10.4.12.)



GRAFEN TIL FUNKSJONEN I OPPGAVE 10.33

OPPGAVE 10.34. Bevis Setning 10.4.7. Start med å vise ulikhetene

$$|f_j(x) - f_j(y)| \leq |\mathbf{f}(x) - \mathbf{f}(y)| \leq \sqrt{\sum_{i=1}^{n} |f_i(x) - f_i(y)|}$$

for alle $j \in \{1, \dots, n\}$ og $x, y \in X$.

OPPGAVE 10.35. Vis *lukkede-mengder kriteriet for kontinuitet*: La $f : X \to Y$ være en funksjon mellom metriske rom. Vis at f er kontinuerlig hvis og bare hvis $f^{-1}(Z) \subset X$ er lukket for alle lukkede mengder $Z \subset Y$.

OPPGAVE 10.36. La X være et metrisk rom og $f : X \to \mathbb{C}$ en kontinuerlig funksjon. La

$$Z(f) = \{x \in X \mid f(x) = 0\}$$

være *nullpunktsmengden til* f.

Vis at $Z(f)$ er lukket i X.

OPPGAVE 10.37. Vis følgende generalisering av Oppgave 9.7: La $f, g : X \to Y$ være kontinuerlige funksjoner mellom metriske rom slik at $f(x) = g(x)$ for alle $x \in E$, hvor E er en tett delmengde av X. Da er $f(x) = g(x)$ for alle $x \in X$.

OPPGAVE 10.38. La (X, d) være et metrisk rom og $x_0 \in X$. Vis at funksjonen

$$\begin{array}{ccc} X & \longrightarrow & \mathbb{R}, \\ x & \mapsto & d(x, x_0) \end{array}$$

er kontinuerlig.

OPPGAVE 10.39. Definisjon 9.9.3 av uniform kontinuitet har sin opplagte generalisering til metriske rom: *En funksjon* $f : X \to Y$ *mellom to metriske* *rom med metrikker* d_X *og* d_Y*, henholdsvis, er* uniformt kontinuerlig *i a dersom det til ethvert reelt tall* $\epsilon > 0$ *finnes et reelt tall* $\delta > 0$ *slik at*

$$d_Y(f(x), f(y)) < \epsilon \ \textit{når} \ d_X(x, y) < \delta$$

for alle $x, y \in X$*.*

Vis de naturlige generaliseringene av Setningene 9.9.10, 9.9.12 og 9.9.28 til metriske rom.

Oppgave 10.40. Denne oppgaven generaliserer Setning 9.9.20 til vilkårlige metriske rom.

La X og Y være metriske rom med metrikker d_X og d_Y, henholdsvis. En funksjon $f : X \to Y$ kalles en *Lipschitz-funksjon* dersom det finnes en $K \in \mathbb{R}$ slik at

$$d_Y\left(f(x_1), f(x_2)\right) \leq K d_X(x_1, x_2)$$

for alle $x_1, x_2 \in X$. (Betingelsen kalles igjen en *Lipschitz-betingelse* med *Lipschitz-konstant* K.)

Vis at en Lipschitz-funksjon er uniformt kontinuerlig (jf. definisjonen gitt i Oppgave 10.39).

Oppgaver til §10.5–10.6

Oppgave 10.41. Vis at de to definisjonene på sammenhengende mengder i Definisjon 10.5.1 er ekvivalente.

Oppgave 10.42. Vis at en delmengde X av $\mathbb{R}$ er sammenhengende hvis og bare hvis X er et intervall, som hevdet i Eksempel 10.5.2. (Husk Definisjon 2.1.20.)

Oppgave 10.43. (a) En delmengde $X \subset \mathbb{R}^n$ er *konveks*[25] dersom for hvert par av punkter $\mathbf{x}, \mathbf{y} \in X$, så er linjestykket mellom $\mathbf{x}$ og $\mathbf{y}$ inneholdt i X.

Vis at konvekse delmengder av $\mathbb{R}^n$ er sammenhengende. (Hint: ta utgangspunkt i Eksempel 10.5.5.)

(b) Vis at et kurvesammenhengende metrisk rom er sammenhengende (jf. Merknad 10.5.11).

Oppgave 10.44. Betrakt delmengden $X \subset \mathbb{R}^2$ som er unionen av x-aksen og grafen til funksjonen $x \mapsto \frac{1}{x}$ på $(1, \infty)$. Mer presist: $X = A \cup B$ med

$$A = \left\{(x, y) \in \mathbb{R}^2 \ \mid \ y = 0\right\} \ \text{ og } \ B = \left\{(x, y) \in \mathbb{R}^2 \ \mid \ x > 0 \ \text{ og } \ y = \frac{1}{x}\right\}.$$

Avgjør om X er sammenhengende eller ikke.

[25] *Convex* på engelsk.

OPPGAVE 10.45. Betrakt *topologenes sinuskurve*

$$X = \left\{ \left(x, \sin\left(\frac{1}{x}\right)\right) \mid x \in (0,1] \right\} \cup \left\{ (0,0) \right\} \subset \mathbb{R}^2$$

fra Eksempel 10.5.10.

(a) Vis at X er sammenhengende. (Hint: Anta at $X = A \cup B$ motsier sammenhengsegenskapen og skriv $X = Y \cup \{(0,0)\}$. Anta at $(0,0) \in A$ og betrakt hver for seg tilfellene $A = \{(0,0)\}$ og $A \neq \{(0,0)\}$.)

(b) En variant fremkommer hvis vi legger til intervallet $[-1,1]$ på y-aksen, slik at vi får mengden

$$\overline{X} = \left\{ \left(x, \sin\left(\frac{1}{x}\right)\right) \mid x \in (0,1] \right\} \cup \left\{ (0,y) \mid y \in [-1,1] \right\}.$$

Vis at $\overline{X}$ er lukket. (Mengden $\overline{X}$ kalles *topologenes lukkede sinuskurve*[26], jf. Merknad 10.5.11.)

(c) Vis at $\overline{X}$ er sammenhengende.

OPPGAVE 10.46. Vis følgende ekvivalente karakterisering av konvergens av en delfølge, som generaliserer resultatet fra Oppgave 7.44 *En følge* $\{a_n\}$ *i et metrisk rom har en delfølge som konvergerer mot a hvis og bare hvis enhver omegn om a inneholder uendelig mange ledd i følgen* $\{a_n\}$*.*
(Sammenlign med resultatet fra Oppgave 10.13.)

OPPGAVE 10.47. La $\{x_n\}$ være en følge i et metrisk rom.
Vis at $\{x_n\}$ konvergerer mot x hvis og bare hvis enhver delfølge av $\{x_n\}$ konvergerer mot x. (Dette generaliserer Observasjon 7.7.3.)

OPPGAVE 10.48. La $S = \{0\} \cup \left\{ \frac{1}{n} \mid n \in \mathbb{Z}^+ \right\}$.

(a) Vis at S er kompakt ved å bruke begge kriteriene i Definisjon 10.6.3.

(b) Vis at S er kompakt ved å henvise til et viktig resultat.

OPPGAVE 10.49. La K være et kompakt metrisk rom og $A \subset K$ en lukket delmengde. Vis at A er kompakt. (Forsøk med begge kriteriene i Definisjon 10.6.3.)

OPPGAVE 10.50. La $f : X \to Y$ være en kontinuerlig bijektiv funksjon fra et kompakt metrisk rom X til et metrisk rom Y.

(a) Vis at den inverse funksjonen $f^{-1} : Y \to X$ er kontinuerlig, med andre ord at f er *bikontinuerlig* eller en *homeomorfi*, i språket fra Merknad 9.5.5. (Hint: Bruk resultatet fra Oppgave 10.49 og *åpnemengder-kriteriet for kontinuitet* (Setning 10.4.5).)

(b) Vurdér Eksempel 10.4.13 i lys av (a).

OPPGAVE 10.51. Bevis Setning 10.6.12. (Hint: for "bare hvis"-delen bruk Setning 10.6.11 og argumentér som i beviset for Følgesetning 10.5.9. For "hvis"-delen, bruk følgekriteriene for kontinuitet og kompakthet.

[26] *The closed topologist's sine curve* på engelsk.

Oppgave 10.52. Dette er en nokså krevende oppgave hvor vi skal vise at kriteriene (i) og (ii) i Definisjon 10.6.3 virkelig er ekvivalente.

(a) Vis først (ii) $\Rightarrow$ (i) med et kontrapositivt bevis.
Hint: La $\{x_n\}$ være en følge i X som ikke har noen konvergent delfølge. Argumentér at enhver $x \in X$ har en omegn V_x som inneholder kun endelig mange ledd av $\{x_n\}$ (jf. Oppgave 10.46) og betrakt familien $\{V_x\}_{x\in X}$.

I resten av oppgaven vil vi bevise (i) $\Rightarrow$ (ii). Anta derfor at X oppfyller (i).

(b) Vis at X har en endelig eller tellbar tett delmengde.
Hint: Fiksér $\delta > 0$ og konstruér en følge i X slik: velg $x_1 \in X$, og etter å ha valgt $x_1, \ldots, x_n \in X$, velg, hvis mulig,
$$x_{n+1} \in X \setminus N_\delta(x_1) \cup \cdots \cup N_\delta(x_n).$$
Vis at denne prosessen må ta slutt (her brukes antagelsen (i)), slik at X kan skrives som en union av endelig mange omegner av radius δ. La så $\delta = \frac{1}{n}$ for alle $n \in \mathbb{Z}^+$ og betrakt sentrene til de tilsvarende omegnene.

Kall delmengden fra (b) for Y.

For å vise at (ii) er oppfylt, la $\{U_\alpha\}_{\alpha\in\mathcal{A}}$ være en samling av åpne mengder slik at $X = \cup_\alpha U_\alpha$. Vi vil vise at X kan skrives som en union av *endelig* mange $U_\alpha$er.

(c) Vis først at X kan skrives som en union av *endelig eller tellbart* mange $U_\alpha$er.
Hint: For enhver $x \in X$, vis at det finnes en $\alpha_x \in \mathcal{A}$ og en $r_x > 0$ slik at $N_{2r_x}(x) \subset U_{\alpha_x}$ og dernest en $p \in Y$ slik at $x \in N_{r_x}(p) \subset U_{\alpha_x}$. Betrakt så samlingen $\{N_r(p)\}_{p\in Y, r\in\mathbb{Q}}$.

(d) Vis dernest at X kan skrives som en union av *endelig* mange $U_\alpha$er.
Hint: Fra (c) er vi enten ferdig eller vi kan skrive X som en tellbar union $X = U_1 \cup U_2 \cup U_3 \cup \cdots = \cup_{n=1}^\infty U_n$, med $U_n \in \{U_\alpha\}$ for alle n. Hvis $X \neq U_1 \cup \cdots \cup U_n$ for alle $n \in \mathbb{Z}^+$, vil $W_n = X \setminus (U_1 \cup \cdots \cup U_n)$ oppfylle $W_n \neq \emptyset$ og $W_{n+1} \subset W_n$ for alle $n \in \mathbb{Z}^+$. Velg $x_n \in W_n$ for alle $n \in \mathbb{Z}^+$ og oppnå en motsigelse ved egenskapen (i).

Oppgave 10.53. I denne oppgaven skal vi bevise *Heine-Borel Teoremet* 10.6.8. Ved Setning 10.6.9 gjenstår det kun å vise at hvis $X \subset \mathbb{R}^n$ er lukket og begrenset (hvor $\mathbb{R}^n$ har den euklidske metrikken), da er X kompakt.

(a) Vis først at enhver delmengde av $\mathbb{R}^n$ på formen $I_1 \times \cdots \times I_n$ med hver I_j et lukket begrenset intervall er kompakt. Hint: Bruk kriterium (i) for kompakthet i Definisjon 10.6.3 og anta at $\{\mathbf{x}_k\}_{k=1}^\infty$ er en følge i X. Skriv $\mathbf{x}_k = (x_{k,1}, x_{k,2,}, \ldots, x_{k,n})$. Ved *Bolzano-Weierstrass teoremet* 7.7.5 har $\{x_{k,1}\}$ en konvergent delfølge, si $\{x_{k_{l_1},1}\}_{l_1}^\infty$. Videre har $\{x_{k_{l_1},2}\}$ en konvergent delfølge. Fortsett prosedyren.

(b) Konkludér ved hjelp av resultatet fra Oppgave 10.49.

Oppgave 10.54. La $\{a_n\}$ være en Cauchy-følge i et metrisk rom. Vis at dersom $\{a_n\}$ har en delfølge som konvergerer mot a, da konvergerer også $\{a_n\}$ selv mot a. (Dette generaliserer resultatet fra Oppgave 7.46.)

Oppgave 10.55. Vis at et kompakt metrisk rom er komplett. (Hint: Bruk kriteriet (i) i Definisjon 10.6.3 og resultatet fra Oppgave 10.54.)

Oppgave 10.56. Vi viser til Oppgave 10.39 for definisjonen av uniform kontinuitet til en funksjon mellom metriske rom. Bevis *Heine-Cantor teoremet* i sin generelle form (som generaliserer Teorem 9.9.13): *En kontinuerlig funksjon $f : X \to Y$ mellom metriske rom, der X er kompakt, er uniformt kontinuerlig.*

(Hint: ta utgangspunkt i kriterium (i) for kompakthet i Definisjon 10.6.3 og beviset for Teorem 9.9.13 utført i Oppgave 9.58.)

Oppgaver til §10.7

Oppgave 10.57. Vis at d i Eksempel 10.7.1 oppfyller betingelsene til en metrikk.

Oppgave 10.58. La X være et diskret metrisk rom (jf. Eksempel 10.7.1).

(a) Hva vil det si at en følge i X konvergerer?
(b) Hva vil det si at en følge i X er Cauchy?
(c) Vurdér om X kan være komplett.
(d) Vurdér om X kan være kompakt.

Oppgave 10.59. (a) Vis at ρ i Eksempel 10.7.6 oppfyller betingelsene til en metrikk.

(b) Mer generelt: La (X_i, d_i) være metriske rom, $i = 1, \dots, n$, $n \in \mathbb{Z}^+$. Betrakt funksjonen $\rho : (X_1 \times \cdots \times X_n)^2 \to \mathbb{R}$ definert ved

$$\rho(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} d_i(x_i, y_i)$$

for $\mathbf{x} = (x_1, \dots, x_n), \mathbf{y} = (y_1, \dots, y_n) \in X_1 \times \cdots \times X_n$. Vis at ρ definerer en metrikk på $X_1 \times \cdots \times X_n$.

Oppgave 10.60. (a) Vis at d_∞ i Eksempel 10.7.7 oppfyller betingelsene til en metrikk.

(b) Mer generelt: La (X_i, d_i) være metriske rom, $i = 1, \dots, n$, $n \in \mathbb{Z}^+$. Betrakt funksjonen $d_\infty : (X_1 \times \cdots \times X_n)^2 \to \mathbb{R}$ definert ved

$$d_\infty(\mathbf{x}, \mathbf{y}) = \max_{i=1,\dots,n} \{d_i(x_i, y_i)\}$$

for $\mathbf{x} = (x_1, \dots, x_n), \mathbf{y} = (y_1, \dots, y_n) \in X_1 \times \cdots \times X_n$. Vis at d_∞ definerer en metrikk på $X_1 \times \cdots \times X_n$.

Oppgave 10.61 (SNCF-metrikken). La (X, d) være et metrisk rom og $P \in X$ et punkt.

(a) Vis at funksjonen $d_P : X \times X \to \mathbb{R}$ definert ved
$$d_P(x,y) = \begin{cases} 0, & \text{hvis } x = y, \\ d(x,P) + d(P,y), & \text{hvis } x \neq y \end{cases}$$
definerer en metrikk på X.

(b) Skissér en omegn om et vilkårlig punkt dersom $X = \mathbb{R}^2$ og P er origo.

(c) Metrikken ovenfor kalles *SNCF-metrikken* (etter "Société Nationale des Chemins de Fer français", de franske statsbaner), eller *British Rail-metrikken* (etter det tidligere statlige selskapet som drev jernbanetrafikk i Storbritannia frem til 1997). Hvorfor?

Oppgaver til §10.8

Oppgave 10.62. Vis at d_∞ i Eksempel 10.8.1 oppfyller betingelsene til en metrikk.

Oppgave 10.63. Vis at funksjonen
$$\begin{array}{ccc} \mathbb{R} & \longrightarrow & \ell_\infty \\ t & \mapsto & \left\{ \dfrac{t}{k} \right\}_{k=1}^{\infty} \end{array}$$
er kontinuerlig.

Oppgave 10.64. Vis at alle rommene ℓ_∞, c, c_0 og c_{00} (jf. Eksemplene 10.8.1 og 10.8.2) har naturlig struktur som $\mathbb{R}$-vektorrom (jf. Definisjon 4.5.5).

Oppgave 10.65. Vis at rommene c og c_0 fra Eksempel 10.8.2 er komplette. (Hint: Bruk resultatet fra Oppgave 10.11 eller mim etter beviset for Setning 10.8.3, eller begge deler.)

Oppgave 10.66. Vi refererer til Eksempel 10.8.2 for definisjonen av rommene c og c_0. Betrakt funksjonen
$$\begin{array}{cccc} F: & c & \longrightarrow & \ell_\infty, \\ & \{x(n)\} & \mapsto & \{x_*, x(1) - x_*, x(2) - x_*, x(3) - x_*, \ldots\}, \end{array}$$
hvor $x_* = \lim_{k \to \infty} x(k)$.

(a) Vis at verdimengden til F ligger i c_0.

(b) Vis at F er kontinuerlig.

(c) Vis at $F : c \to c_0$ har en invers, som igjen er kontinuerlig.

(d) Er F en isometri (jf. Merknad 10.2.15 for definisjonen)?

Oppgaver til §10.9

OPPGAVE 10.67. Vis at d_∞ i Eksempel 10.9.1 oppfyller betingelsene til en metrikk.

OPPGAVE 10.68. Definér funksjoner $f_n : [0,1] \to \mathbb{R}$ ved

$$f_n(x) = \begin{cases} 2n^2 x, & \text{hvis } 0 \le x \le \frac{1}{2n}, \\ 2n^2\left(\frac{1}{n} - x\right), & \text{hvis } \frac{1}{2n} < x < \frac{1}{n}, \\ 0, & \text{hvis } \frac{1}{n} \le x \le 1. \end{cases}$$

Vis at følgen $\{f_n\}$ konvergerer punktvis mot funksjonen 0. Avgjør om konvergensen er uniform.

OPPGAVE 10.69. Definér funksjoner $f_n : \mathbb{R} \to \mathbb{R}$ ved

$$f_n(x) = \frac{\sin(nx)}{n}.$$

Avgjør om følgen $\{f_n\}$ konvergerer punktvis og/eller uniformt.

OPPGAVE 10.70. Vis at $\mathcal{C}[a,b]$ har naturlig struktur som $\mathbb{R}$-vektorrom (jf. Definisjon 4.5.5).

OPPGAVE 10.71. La $H : [a,b] \to [c,d]$ være en funksjon mellom to lukkede, begrensede intervaller i $\mathbb{R}$. Betrakt funksjonen

$$\begin{array}{rccc} H^* : & \mathcal{C}[c,d] & \longrightarrow & \mathcal{C}[a,b] \\ & h & \mapsto & f \circ H. \end{array}$$

(a) Begrunn kort at H^* er veldefinert, det vil si at verdimengden virkelig ligger i $\mathcal{C}[a,b]$.
(b) Vis at H^* er kontinuerlig.
(c) Vis at H er surjektiv hvis og bare hvis H^* er injektiv.
(d) Vis at H er injektiv hvis og bare hvis H^* er surjektiv.

OPPGAVE 10.72. Betrakt det kartesiske produktet $Y = \mathcal{C}[0,1] \times [0,1]$ og gjør det til et metrisk rom med en av metrikkene definert i Oppgave 10.59(b) eller 10.60(b) (det spiller ingen rolle for oppgaven hvilken vi bruker). Med andre ord, for $(f,a),(g,b) \in \mathcal{C}[0,1] \times [0,1]$, definerer vi metrikken

$$d_Y\Big((f,a),(g,b)\Big) = d_\infty(f(x),g(x)) + |a-b|$$

eller

$$d_Y\Big((f,a),(g,b)\Big) = \max\Big\{d_\infty(f(x),g(x)), |a-b|\Big\}.$$

Vis at *evalueringsfunksjonen*

$$\begin{array}{rcl} Y = \mathcal{C}[0,1] \times [0,1] & \longrightarrow & \mathbb{R} \\ (f,x) & \mapsto & f(x) \end{array}$$

er kontinuerlig.

Oppgave 10.73. Bevis Setning 10.9.5, ved å ta utgangspunkt i beviset for Setning 10.9.2 og bytte ut $[a, b]$ med X.

Oppgaver til §10.10

Oppgave 10.74. Bevis Setning 10.10.6 ved å mime etter beviset for Setning 10.3.21. Du trenger å vise at Observasjon 10.3.18 og egenskap (292) overføres til topologiske rom.

Oppgave 10.75. La X være et topologisk rom, med topologi $\mathcal{T}$, og $Y \subset X$ en delmengde. Vis at samlingen av delmengder av Y gitt ved

$$\mathcal{T}_Y = \{U \subset Y \mid U = Y \cap V \text{ for en } V \in \mathcal{T}\}$$

definerer en topologi på Y, som hevdet i teksten (*underromstopologien*).

Oppgave 10.76. Verifisér at samlingen $\mathcal{T}$ i Eksempel 10.10.14 definerer en topologi og at denne ikke er Hausdorff når X er uendelig.

Oppgave 10.77. Vi skal vise noen av utsagnene i Eksempel 10.10.12 om torusen T.

(a) Vis at (321) definerer en ekvivalensrelasjon på $\mathbb{R}^2$.
(b) Vis at (322) definerer en topologi på T.
(c) Vis at avbildningen $f : \mathbb{R}^2 \to T$ er kontinuerlig.

Oppgave 10.78. Er torusen (Eksempel 10.10.12) Hausdorff?

Oppgave 10.79. Verifisér at *kvotienttopologien* definert i (323) fra Merknad 10.10.13 er en topologi.

Oppgave 10.80. Vi skal i denne oppgaven se om resultatet fra Oppgave 10.21 kan generaliseres til topologiske rom.

(a) La X være et Hausdorffrom og $x \in X$. Vis at $\{x\}$ er lukket i X.

(b) Vis ved et eksempel, at resultatet fra (a) ikke trenger å være sant dersom vi dropper Hausdorffbetingelsen.

Oppgave 10.81. (a) Verifisér at samlingen $\mathcal{T}$ i Eksempel 10.10.15 på en mengde X definerer en topologi og at denne ikke er Hausdorff når X er overtellbar.

(b) Vis at en følge $\{x_n\}$ i X konvergerer mot $a \in X$ hvis og bare hvis det finnes en $N \in \mathbb{Z}^+$ slik at $x_n = a$ for alle $n \geq N$.

Oppgave 10.82. Vis følgende ekvivalente karakterisering av konvergens, som generaliserer resultatet fra Oppgave 10.13: *En følge $\{a_n\}$ i et topologisk rom konvergerer mot a hvis og bare hvis enhver omegn om a inneholder alle bortsett fra endelig mange a_n.*

Oppgave 10.83. La X være et topologisk rom og $A \subset X$ en delmengde.

Definisjon 9.10.1 generaliseres naturlig til topologiske rom: Et punkt $x \in X$ er et *opphopningspunkt* eller *grensepunkt* til A dersom alle omegner om x i X inneholder et punkt i A utenom a (med andre ord: for enhver omegn V om x er $V \cap (A \setminus \{a\}) \neq \emptyset$).

(a) Vis at hvis $\{x_n\}$ er en følge i A som konvergerer mot $x \in X$, da er x et opphopningspunkt til A. (Dette generaliserer den ene implikasjonen i Setning 9.10.2.)
(b) Vis ved å studere Eksempel 10.10.22 at det motsatte av (a) ikke holder: et opphopningspunkt til A trenger ikke å være grense til en følge i A. (Dette viser at den andre implikasjonen i Setning 9.10.2 ikke generaliseres til topologiske rom, i motsetning til metriske rom, jf. Oppgave 10.14).)
(c) Vis at A er lukket hvis og bare hvis A inneholder alle sine opphopningspunkter.

Oppgave 10.84. Betrakt mengden $X = \{a, b, c\}$ med topologien fra Eksempel 10.10.28. Betrakt $\mathbb{R}$ som et topologisk rom via den euklidske metrikken. Avgjør om funksjonen $f : X \to \mathbb{R}$ gitt ved

$$f(a) = f(c) = 0, \ f(b) = 1$$

er kontinuerlig.

Oppgave 10.85. Vis at resultatene fra Oppgave 10.26 og 10.27 generaliseres til topologiske rom, som hevdet i Merknad 10.10.26:

(a) Vis at hvis $f : X \to Y$ er en kontinuerlig funksjon mellom topologiske rom og $Z \subset X$ er et underrom, da er også restriksjonen $f|_Z : Z \to Y$ kontinuerlig.
(b) La X være et underrom av det topologiske rommet Y og $i : X \to Y$ inklusjonsavbildningen. Vis at i er kontinuerlig.
(c) La X, Y, Z være topologiske rom, $Y \subset Z$ et underrom og $i : Y \to Z$ inklusjonsavbildningen. Vis at $f : X \to Y$ er kontinuerlig hvis og bare hvis $i \circ f : X \to Z$ er kontinuerlig.

Oppgave 10.86. Generaliser *lukkede-mengder kriteriet for kontinuitet* fra Oppgave 10.35 til topologiske rom: La $f : X \to Y$ være en funksjon mellom topologiske rom. Vis at f er kontinuerlig hvis og bare hvis $f^{-1}(Z) \subset X$ er lukket for alle lukkede mengder $Z \subset Y$.

Oppgave 10.87. La K være en kropp. Betrakt K som et topologisk rom ved *Zariski-topologien* (Eksempel 10.10.16). La $F \in K[x]$. Vis at funksjonen

$$\begin{array}{ccc} K & \longrightarrow & K \\ a & \mapsto & F(a) \end{array}$$

er kontinuerlig. (Hint: bruk *lukkede-mengder-kriteriet* fra Oppgave 10.86.)

Oppgave 10.88. Vis at de to definisjonene på sammenhengende topologiske rom i Definisjon 10.10.33 er ekvivalente.

Oppgave 10.89. La X være et topologisk rom. Vis at en union av åpne sammenhengende delmengder av X med ett felles punkt er sammenhengende.

Oppgave 10.90. Bevis Setning 10.10.41. (Hint: Vis at $X \setminus K$ er åpen. La $x \in X \setminus K$. For hver $y \in K$, la U_y og V_y være disjunkte omegner om x og y henholdsvis. Argumentér at det finnes endelig mange $y_1, \ldots, y_n \in K$ slik at $K \subset V_{y_1} \cup \cdots \cup V_{y_n}$. Konkluder at $U = U_{y_1} \cap \cdots \cup U_{y_n}$ er en omegn om x som er disjunkt fra K. Hvor har du brukt Hausdorff-betingelsen?)

Oppgave 10.91. Generalisér resultatet fra Oppgave 10.49 til topologiske rom: La K være et kompakt tolopogisk rom og $A \subset K$ en lukket delmengde. Vis at A er kompakt.

Oppgave 10.92. I denne oppgaven skal vi generalisere resultatet fra Oppgave 10.50 til topologiske rom.

La $f : X \to Y$ være en kontinuerlig bijektiv funksjon fra et kompakt topologisk rom X til et Hausdorffrom Y.

(a) Vis at den inverse funksjonen $f^{-1} : Y \to X$ er kontinuerlig, med andre ord at f er *bikontinuerlig* eller en *homeomorfi*, i språket fra Merknad 9.5.5. (Hint: Bruk *lukkede-mengder-kriteriet* fra Oppgave 10.86, resultatet fra Oppgave 10.91 og Setningene 10.10.39 og 10.10.41.)

(b) Vis ved et moteksempel at resultatet fra (a) ikke gjelder lenger om vi dropper betingelsen om at Y er Hausdorff.

Referanser

[AE] R. A. Adams, C. Essex, *Calculus–A complete course*, ninth edition, Pearson Canada, 2017.

[BLPR] B. Baritompa, R. Lowen, B. M. Polster, M. Ross, *Mathematical table-turning revisited*, Mathematical Intelligencer **29** (2007), 49–58.

[Da] P. J. Davies, *Are there coincidences in mathematics?*, Amer. Math. Monthly **88** (1981), 311-320

[KM] D. Kalish, R. Montague, *Logic: Techniques of Formal Reasoning*, 1964

[La] J. C. Lagarias (ed.), *The Ultimate Challenge: The* $3x+1$ *Problem*, The American Mathematical Society, Providence, 2010.

[Li] T. Lindstrøm, *Kalkulus*, 4. utgave, Universitetsforlaget, 2016.

Begrepsindeks

Navnsindeks

Symbolsindeks

Andreas Leopold Knutsen er Professor i Matematikk ved Universitetet i Bergen. Hans forskningsfelt er *Algebraisk Geometri*, hvor han har publisert over 50 fagfellevurderte artikler med mer enn 30 medforfattere fra hele verden. Han har lang erfaring både som underviser av kurs og foredragsholder ved norske og internasjonale universiteter, og har holdt forelesninger på norsk, engelsk, italiensk, fransk og tysk. Denne boken er basert på erfaringer gjort som underviser av begynneremner for matematikkstudenter ved Universitetet i Bergen.

Andreas Leopold Knutsen nyter *Moules Frites* i Marseille under en konferanse i 2013