Algebra I

Rikard Bøgvad, Qimh Xantcha & Håkan Granath

Matematiska institutionen
Stockholms universitet

Tionde tryckningen
2018

A L G E B R A

Rikard Bøgvad, Qimh Xantcha, Håkan Granath

Tionde tryckningen, 2018

Förord

Detta kompendium innehåller material till första terminens kurs i algebra vid Matematiska institutionen vid Stockholms Universitet, närmare bestämt allt utom linjär algebra. Läsaren och vi är skyldiga ett stort tack till alla studenter och lärare som hjälpt oss korrigera fel och brister i tidigare versioner.

Som utgångsmaterial har vi haft sammanställningar och föreläsningsanteckningar av Ralf Fröberg, Martin Tamm, Yishao Zhou och andra, och så förstås vår personliga ambition att skriva vår generations stora och definitiva algebrakompendium. Paul Vaderlind har gått igenom allt noggrant och givit många kommentarer. Att texten nu förhoppningsvis är läsbar, korrekt och användbar är i mångt och mycket hans förtjänst.

Vi tycker, att motiveringarna ofta kan vara bristfälliga, om varför en viss typ av matematik finns och anses viktig, och vi har därför försökt placera in resultaten i en litet vidare bild. Bäst är, att fråga föreläsaren åtminstone en gång per vecka "varför är det här viktigt, vad används det till, varför existerar det?". Den mesta matematiken har en gång utvecklats för att lösa intensivt spännande och viktiga problem, och fortsätter av bara farten att vara användbar. Algebrans senaste triumf är Google-algoritmen, baserad på ideér, som är förfärligt smarta, men egentligen mest grundläggande linjär algebra.

För tusen år sedan gav Al-Khwarizmi — den förste att använda ordet *al-jabr* — följande programförklaring i sin inflytelserika bok om algebra:

> I shall now teach you how to multiply the unknown numbers, that is to say, the roots, one by the other, if they stand alone, or if numbers are added to them, or if numbers are subtracted from them, or if they are subtracted from numbers; also how to add them one to the other, or how to subtract one from the other.

Bättre kan det egentligen inte uttryckas, vad algebra handlar om. Men om läsaren inte känner sig tillräckligt taggad av Al-Khwarizmi, vill vi bara påpeka: (a) materialet är helt grundläggande för tillämpningar inom alla ämnen som använder sig av matematik (fysik, kemi, datavetenskap, ekonomi, ...), (b) rätt "skitkul" och (c) privatekonomiskt lönsamt — i alla fall problemet om poker i kapitlet om kombinatorik — och sålunda värt allt arbete.

Rikard Bøgvad & Qimh Xantcha

Revideringen inför 10:e tryckningen består av mindre innehållsliga ändringar, en uppdaterad typografi, och korrigeringar av kända tryckfel. Alla eventuella nya fel är faller därmed helt på undertecknad.

Håkan Granath

Innehåll

Kapitel 0

Matematikens kunskapsteori

§1. Rationell och empirisk vetenskap

Innan vi ger oss i kast med matematikstudier på universitetsnivå, kan det vara lämpligt att fundera litet över, vad matematisk kunskap och vetenskap egentligen innebär.

Som student, oavsett ämne, lär man sig (förutom någon form av "hantverksskicklighet") bland annat s.k. *fakta*, vars sanningshalt väl bör vara garanterad av, att de lärs ut av kompetenta lärare vid ett rumsrent universitet. Man accepterar alltså dessa fakta för att auktoriteterna säger så. Förutom detta, så skolas man in i, att själv kunna etablera sanna fakta i de olika ämnena genom forskning. Metoder för att fastställa, vad som är sant, varierar drastiskt från ämne till ämne.

Jämför t.ex. följande påståenden.

- Ljushastigheten i vakuum är ungefär 300 000 km/s.
- Gavrilo Princip mördade Franz Ferdinand (en händelse som utlöste första världskriget).
- Det finns oändligt många primtal.
- En deriverbar funktion, vars derivata är positiv, är växande.

De två första påståendena är empiriska och uttalar sig om den verklighet vi lever i. I filosofisk mening är de *falsifierbara*.

Fysikern kan anställa ett experiment att mäta ljushastigheten och kanske rentav upprepar Michelson–Morleys berömda experiment för att kontrollera, att ljushastigheten är konstant och ej förändras, om ljusstrålen skickas ut från ett objekt i rörelse. Experiment med verkligheten är det som bekräftar att ett påstående är sant.

Historikern kan inte experimentera, men kan studera överlevande dokument och historiska (ibland arkeologiska) lämningar. Av källorna kan man se i vilken mån det går att utläsa, att just Gavrilo Princip var boven, samt i vilken utsträckning detta kan sägas ha utlöst världskriget. Här är förstås dokumenten öppna för tolkning, och historikerna kan förstås munhuggas om, vilka andra skeenden som spelat roll, men det är ändå, i någon mån, en falsifierbar utsaga som går att empiriskt verifiera eller falsifiera. Detta är kriteriet på sanning i de empiriska vetenskaperna.

Vad är det då som har etablerat de två senare, matematiska påståendena som sanna? Det är förstås *beviset*, det som så poetiskt har omtalats som *matematikens själ*. Beviset spelar för matematikern samma roll som experimentet gör för fysikern och kritiskt källstudium för historikern. Detta förklarar varför bevisen intar en så central position i den här boken och i all presentation av matematik. Bevisets funktion är, i första hand, att etablera matematiska sanningar.

Det är också nödvändigt att lära sig förstå och behärska teknikerna däri med syfte att självständigt lösa nya problem i matematiken eller dess tillämpningar. Om man förstår *varför* tecknet på derivatan bestämmer om funktionen är växande eller avtagande i ett visst intervall, så är man förstås bättre rustad att använda partiell information om derivatans storlek, t.ex. att den ligger mellan $\frac{1}{2}$ och 1, för att uppskatta funktionens beteende över ett intervall.

Matematikern, stackarn, har ingen verklighet att gå till för att fysiskt se efter om ett resultat är sant, utan måste klara sig med logisk argumentation. Å andra sidan är matematikern inte bunden till någon torftig *verklighet*, utan äger utomordentlig frihet att välja de modeller och "världar" hon konstruerar. Tag t.ex. begreppet *reella tallinjen*. Liksom andra matematiska idéer

är den en ren abstraktion, som saknar förankring i verkligheten, utan existerar blott i matematikernas böcker och tankar. Förstås är den utvecklad för att användas praktiskt och visar sig verkligen vara ett effektivt verktyg. Men är den verkligen självklart bäst för allt? Varför skall vi ständigt använda dessa utnötta gamla reella tal, oändliga decimalbråk, och inte undersöka varianter där tallinjen är sammansatt av små odelbara minsta delar, som en sorts atomer eller pixlar? Vad händer då? Dessa idéer lägger grunden till ett modernt område kallat *digital geometri*. Matematiker hittar ständigt på nya begrepp och definitioner. Vissa är förstås, subjektivt sett, intressantare och bättre än andra, beroende på vilka fenomen i verkligheten man vill förstå eller beskriva.

Låt oss återgå till de första exemplen, fysikern och historikern, och observera, att det finns en hel del skillnader vad säkerheten av de vunna resultaten anbelangar. Är experimentet för att mäta ljushastigheten verkligen helt säkra? Visst går det t.ex. att föreställa sig en värld, där några relativistiskt kunniga men klåtentakliga trafikpoliser sett till, att ljuset i tätorten verkligen håller hastighetsbegränsningen och inte blåser iväg med mer än 300 000 km/s, medan det däremot på den galaktiska landsbygden kan gå betydligt vildare till. Eller visst går det att argumentera för att Franz Ferdinands död bara var en förevändning för ett krig alla egentligen drömde om, eller att det var en komplott av det österrikiska hovet? Speciellt troligt låter det förstås inte, men *helt* säkra kan vi inte vara. Det är som i en rättegång: de empiriska vetenskapsmännen förväntas bara ställa det utom allt *rimligt* tvivel, att det var X, som mördade Y. Då och då är det förstås ändå Z, som är skurken.

Däremot är det obestridligen och oåterkalleligen så, att de två matematiska satserna ovan är sanna. Det går inte, med mindre man förvrider ordens mening och perverterar språket, att hävda existensen av ett största primtal.

Skulle en känd fysiker hävda, att ljushastigheten varierar beroende på platsen i universum, skulle vi känna oss tvungna, att lyssna på argumenten, för att kunna avgöra om denna nya teori vore rimlig. Kanske skulle vi känna oss glada för att universum är ändå intressantare, än vad det kan tyckas vara i trakterna kring Stockholm.

Men att någon påstår sig ha funnit ett exakt uttryck för π som rationellt tal, det intresserar oss inte. Vi har läst och förstått beviset, som ställer irrationaliteten av π bortom allt tvivel. Inte bara utom allt *rimligt* tvivel, utan verkligen bortom *allt* tvivel. Vore det så en världsberömd matematiker, så kan vi med bestämdhet slå fast, att vederbörande gått och blivit en smula *fnoskig* på gamla dagar. Motsvarande reaktion skulle inte möta en åldersstigen fysiker med oortodoxa idéer. Här är alltså en viktig skillnad mellan de båda ämnena: Matematiker kan bli senila och passé, men även den mest ålderstigne fysikers idéer måste tyvärr tagas på allvar.

Det här är kanske lätt att missförstå som lokalpatriotiskt skryt, ungefär att matematik minsann står sig gott i jämförelse med andra vetenskaper, precis som Halmstad BK efter några öl. Det är det nu inte, utan är en naturlig följd av att matematiken är en *rationell* vetenskap (jämte filosofien), inte *empirisk* (som i princip alla andra riktiga vetenskaper), och därför följer andra regler. Matematiken lever i sitt eget abstrakta universum. Hon kan därför bara uttala sig om sig själv, inte om världen utanför, den "riktiga världen", det universum och den planet vi befolkar.

Matematiker måste sålunda ha andra sätt att förhålla sig till begreppet *sanning*. Matematikens bevisregler applicerade på t.ex. juridiken hade inneburit att ingen någonsin blivit dömd för något brott (utom möjligen för division med 0), och skulle alltså ha inneburit att juridiken hade blivit meningslös och oanvändbar.

Fysiker, som är tvingade till det mest intima umgänget med oss matematiker, tycker ofta att vi är pedantiska och småaktiga *petits-maîtres*. Beträffande Taylor-utvecklingen av exponentialfunktionen (studerad i analysen),

$$e^x = 1 + x + \frac{1}{2}x^2 + \cdots,$$

sade oss en känd fysiker en gång torrt, att en fysiker är intresserad av approximationen $e^x \approx 1+x+\frac{1}{2}x^2$ (giltig då $x \approx 0$), medan en matematiker endast intresserar sig för prickarna $\cdots$ (felet vid approximationen). Snärtigt så det förslår. För fysikerns del är det själva approximationen, som är det väsentliga i sammanhanget, eftersom den gör att han kan approximera hejvilt, så som fysiker älskar att göra. Matematikern kräver exakthet i varje steg. Detta är en vanlig och relativt sund frustration hos fysikerna. Hade man på matematiskt manér krävt kvantmekanikens

logiska och matematiska förenlighet med då kända fysikaliska teorier (början av 1900-talet), så hade det, sorry, inte blivit någon kvantmekanik alls.

§2. Matematiska modeller

Matematikens logiska struktur är emellertid inte enbart ett självändamål, utan också ett villkor för dess praktiska användbarhet. Matematik är ju, i mångt och mycket, utvecklat för att vara naturvetenskapernas språk, med vilket man kan bygga (förenklade) tankemodeller av verklighetens komplicerade fenomen. Då ligger det förstås i sakens natur, att man inte skall behöva springa till Mamma Natur för ständig bekräftelse. Inuti den matematiska modellen resonerar man exakt och logiskt, vilket leder till konkreta förutsägelser om världens beskaffenhet. Dess tillämpbarhet kan sedan studeras via experiment och observationer av verkligheten. Gav den matematiska modellen goda förutsägelser, anses den vara en duglig approximation av verkligheten.

Tag till exempel Newtons analys av planetrörelse, en av startpunkterna för både den matematiska analysen och för den fysikaliska vetenskapen. Hans utgångspunkt var en enda fysikalisk princip, att planeter (liksom alla kroppar) påverkar varandra med något han kallade "krafter" enligt *Gravitationslagen*

$$F = \frac{Gm_1m_2}{r^2}.$$

Med denna som enda axiom, lyckades han matematiskt bevisa, vad Brahe och Kepler redan funnit genom astronomiska observationer, nämligen att planetbanorna är ellipser med solen i ena brännpunkten (Keplers första lag för planetrörelse). *Givet* att Gravitationslagen stämmer, är det alltså en *matematisk nödvändighet*, att planeterna kretsar i ellipser. Detta senare faktum visar sig experimentellt stämma, och den newtonska mekaniken tycks alltså vara en god modell för astrofysiken.

Även Einstein tog avstamp i en fysikalisk princip, den om ljushastighetens invarians för alla obervatörer, vilken tycktes belagd genom Michelson–Morleys ovanbemälda experiment. Härur kunde han härleda den speciella relativitetsteorien rent matematiskt. Teoriens förutsägelser, varibland tidsdilationen, längdkontraktionen och ekvivalensen $E = mc^2$ mellan energi och massa, har sedermera experimentellt verifierats.

I bägge fallen var det förstås av största vikt, att de matematiska resonemangen (efter det att de fysikaliska axiomen valts) var stringenta och logiskt bindande, så att eventuella korrekta förutsägelser gjorda genom modellen verkligen troliggjorde denna.

§3. Matematisk teori

De gamla egyptierna kände Pythagoras sats, men för dem var den ett empiriskt och experimentellt faktum. I lantmäteriet kom de i kontakt med en massa rätvinkliga trianglar och observerade, att egenskapen $a^2 + b^2 = c^2$ tycktes gälla. De hade också en "formel" $A = \frac{1}{4}(a + c)(b + d)$ för arean av en godtycklig fyrhörning med sidorna a, b, c, d. Empiriskt tycktes även denna äga beständighet, men den ger enbart en approximation, bättre ju mera rektangelaktig fyrhörningen är. Egyptierna gjorde ingen åtskillnad mellan dessa. För dem var matematiken lika empirisk som fysiken eller astronomien.

Matematiken som vetenskap danades av de gamla grekerna. Här föddes idén om att matematiska påståenden kan och måste *bevisas*. Pythagoras sägs exempelvis vara den förste, som bevisade den sats, som bär hans namn (huruvida detta är sant kan diskuteras).

En *sats* är ett sant påstående av större intresse. Den åtföljs av ett noggrant *bevis*, som egentligen är ett logiskt argument för att påståendet i satsen är sant. Samma sats kan ofta bevisas på många olika sätt — t.ex. finns det en bok med 150 olika bevis för Pythagoras sats. Det finns många olika sätt att förstå varför den är sann. Egentligen skiljer ingenting nämnvärt mellan beviset och argumentationen i en lösning av en elementär räkneuppgift. I bägge fallen skall det produceras en logiskt korrekt och oemotsäglig demonstration av att något är sant. En mindre intressant och ofta teknisk hjälpsats kallas *lemma*.

Frågan är då, vad satserna bevisas från. Från redan kända och i någon mening enklare satser, naturligtvis, som i sin tur måste bevisas utifrån ännu enklare satser, och så vidare. Matematiken

är en pyramid, ett Babelstorn, som för varje år byggs allt högre och mäktigare. Men av vad material är då pyramidens fundament gjutet av? Ett fundament måste nämligen finnas, om vi inte vill råka in i cirkelresonemang.

Några grundläggande, obevisbara sanningar måste vi ha att utgå från. Dessa är den matematiska teoriens *axiom* eller *postulat*. Axiomen kan sägas definiera begreppet sanning inom den matematiska teorien. Om dessa är eller anses som sanna, så måste allt vi kan bevisa från dem också med nödvändighet vara sant.

Euklides var den förste att inse behovet av elementära "sanningar" att utgöra en grund för matematiken. I sitt monumentala verk *Elementa* bygger han upp all sin tids matematiska vetenskap (d.v.s. geometri och talteori), utifrån en handfull postulat, varifrån man successivt bevisar allt mer avancerade satser. Första boken kulminerar i Pythagoras sats. Två av Euklides postulat är:

- Mellan två punkter går det en och endast en rät linje.
- Alla räta vinklar är lika stora.

Tydligen uttalar sig Euklides postulat om saker som punkter, räta linjer och räta vinklar. Dessa måste då definieras i teorien. *Definitionen* talar om precis vad som hädanefter kommer att menas med en viss term. Det kan betyda, att ett välkänt ord, t.ex. *funktion*, ges en annorlunda innebörd än den man är van vid från vardagslivet. De är alltid så kallade namndefinitioner, d.v.s. de definierar hur ett ord skall användas. De har också en absolut karaktär — ett objekt är antingen av den definierade typen eller inte. Jämför med samhällsvetarnas definition av begreppet *demokrati* som *folkstyre*. Denna definition kan lätt urarta i politiska slagsmål. (Är Kina en demokrati eller inte?)

Euklides försöker tappert definiera allt. Det går inte. Samma problem vidlåder definitionerna som satserna. Någonting måste vi ha att utgå ifrån, om vår teori inte skall bli cirkulär och obrukbar. I botten av teorien, jämte axiomen, ligger därför en samling odefinierbara *grundbegrepp*. I *Elementa* kan begreppen *punkt* och *linje* sägas vara grundbegrepp. En *triangel* kan däremot definieras utifrån dessa som en samling av tre punkter och linjerna mellan dessa.

I denna framställning av algebran kommer vi att ta de *reella talen* som grundbegrepp. Vi definierar dem inte, utan de förutsättes existera. För dessa antar vi en samling axiom, som fastlägger egenskaperna för addition, multiplikation och ordningsrelationen har. Från axiomen kommer vi sedan att bevisa de övriga räknelagar vi känner behov av; dessa blir då våra satser.

§4. Matematiska bevis

Matematiska satser har ofta formen av *implikationer*: om vissa förutsättningar gäller, så är en viss slutsats sann. Exempel på sådana satser, från olika områden av matematiken, kan vara:

- Om talet p är ett primtal större än 3, så kan p skrivas på formen $p = 6k \pm 1$.
- Om $p(x)$ är ett polynom (med komplexa koefficienter), så har $p(x)$ något (komplext) nollställe.
- Om ABC är en triangel, i vilken sidorna $AB = AC$, så är vinklarna $\angle B = \angle C$.
- Om funktionen $f(x)$ är deriverbar i en punkten $x = a$, så är $f(x)$ även kontinuerlig i $x = a$.

Bevis för sådana satser faller oftast i någon av de tre kategorierna direkta bevis, indirekta bevis och motsägelsebevis. En fjärde specialvariant är *induktionsbeviset*, som behandlas separat i ett senare kapitel.

Det *direkta beviset* är det vanligaste. Här startar man från de givna förutsättningarna i satsen eller problemet, och resonerar sig fram till slutsatsen.

Exempel 1.
Vi ger ett direkt bevis för påståendet: *Om n är ett udda heltal, så är även n^2 udda.* Eftersom

n är udda, finns det ett heltal k, så att $n = 2k + 1$. Det ger att

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

som har resten 1 vid division med 2, och är alltså ett udda tal.

Det *indirekta beviset* går i motsatt riktning. Här startar man med att anta (läs: låtsas) slutsatsen vara falsk, och resonerar sig fram till att förutsättningen (eller förutsättningarna) då inte heller kan gälla.

Exempel 2.
Vi ger ett indirekt bevis för påståendet: *Om n är ett udda heltal, så är även n^2 udda.* Ett indirekt bevis går som följer. Antag slutsatsen vara falsk; antag alltså n^2 vara jämn. Då är n^2 delbart med 2, alltså är även n delbart med 2, således jämnt. Men detta strider mot den givna förutsättningen, att n skulle vara udda. Talet n^2 kan alltså inte vara jämnt under rådande förutsättning, utan måste vara udda.

Motsägelsebeviset, slutligen, fungerar som följer. Antingen är slutsatsen sann eller så är den falsk, alltså dess motsats är sann. Kan vi visa, att motsatsen leder till en uppenbar omöjlighet, så har vi eliminerat möjligheten, att motsatsen är sann, och kan dra slutsatsen att vår önskade slutsats måste vara sann.

En sådan resonemang kallas på latin *Reductio ad absurdum*, d.v.s. reduktion till det absurda. Det är en teknik som ofta användes i vardagslivet, t.ex. i familjegräl. Tonårssonen menar sin förälder vara en elak och snål jävel, och föräldern (som t.ex. är lektor i matematik), avkräver honom bevis. Då ges kanske svaret: "Om du hade varit snäll eller givmild, så hade du köpt en ny och bättre dator." För att göra detta till ett perfekt och stringent motsägelsebevis, så fattas bara en omformulering: "Låt oss antaga att du är en snäll eller givmild förälder. Då inses lätt att du hade köpt en nyare dator, vilket uppenbart är falskt." (Se här matematikens användbarhet.)

Exempel 3.
Här är ett motsägelsebevis till en väldigt berömd och mer än 2000 år gammal sats: *Talet $\sqrt{2}$ är irrationellt.*

Antag motsatsen, d.v.s. att $\sqrt{2}$ är rationellt och alltså att det finns heltal m och n, så att $\sqrt{2} = \frac{m}{n}$. Vi väljer ett slutförkortat sådant bråk $\frac{m}{n}$. Kvadrering av båda leden ger $2 = \frac{m^2}{n^2}$; omflyttning ger $m^2 = 2n^2$. Alltså är m^2 jämnt, och talet m måste därför vara jämnt (förra exemplet!). Det finns därför ett heltal p, sådant att $m = 2p$. Insättning ger $4p^2 = 2n^2$, varav $2p^2 = n^2$. Men detta innebär att n också är jämnt. Alltså var både m och n jämna, och hade den gemensamma faktorn 2. Men detta strider mot att bråket $\frac{m}{n}$ valdes slutförkortat!

Konklusionen är, att det omöjligtvis kan finnas något slutförkortat bråk $\frac{m}{n} = \sqrt{2}$. Talet $\sqrt{2}$ kan därför inte vara rationellt, för detta leder ju till en absurditet, utan motsatsen måste vara sann. Talet $\sqrt{2}$ *är* därför irrationellt.

Del I

Reell algebra

Kapitel 1

Symbolisk algebra

§1. Varför algebra när det finns miniräknare?

Vi har förstås stött på många algebraiska samband inom naturvetenskapen, t.ex. hur tillryggalagd sträcka s beror av (konstant) hastighet v och tid t:

$$s = tv.$$

Vi behöver inte ha konkreta tal för att kunna dra slutsatser ur detta samband (även kallat Svenssons TV efter en mytisk matematiklärare Svensson...). Som ett väldigt banalt exempel kan vi ju t.ex. se, att då vi dubblar tiden, blir den färdade sträckan s_2 dubbelt så lång:

$$s_2 = (t + t)v = tv + tv = 2s.$$

Beroende på vad man vet och vill veta, är det förstås också behändigt att kunna omformulera $s = tv$ till $v = \frac{s}{t}$ eller $t = \frac{s}{v}$. Detta är exempel där vi räknar algebraiskt med bokstäver som reella tal, utan att behöva veta vad varje bokstav står för just i skrivande ögonblick. De samband vi får gäller för alla tal. Matematikens effektivitet spirar ur just denna allmängiltighet. (Kanske just detta exempel var tämligen avslaget, men det är bara att titta i vilken fysik- eller kemibok som helst för att se häftigare saker.)

Precis som i exemplet använder algebran symboler istället för tal — oftast latinska bokstäver, men ibland också av lång tradition bokstäver från det grekiska alfabetet.

§2. Räknelagar

För att kunna genomföra algebraiska räkningar krävs en mer medveten och systematisk styrning av räkningarna än vid vanlig aritmetik ("sifferräkning"). Vissa tekniker, som i konkreta räkningar blivit så självklara, att vi förlorat dem ur sikte, är nu kraftfulla verktyg för att manipulera och förenkla uttryck. En del av dem bär namn av räkneregler eller *räknelagar*. Det banalaste exemplet är kanske, att ordningen kan omkastas vid multiplikation och addition, så att

$$ab = ba \qquad \text{och} \qquad a + b = b + a. \tag{1}$$

Den som älskar snitsiga namn kan här lära sig, att dessa räkneregler kallas *kommutativa lagarna* för multiplikation respektive addition. De äger giltighet för alla tal, reella som komplexa.

T.ex. säger (1) att $2 \cdot 3 = 3 \cdot 2$, men också att $217 \cdot 3x = 3x \cdot 217$ för ett tal x, vilket som helst. (Vi vet förstås också att $217 \cdot 3x = (217 \cdot 3)x = 651x$.)

Observera, att vi för det mesta inte bryr oss om att skriva ut multiplikationstecknet $\cdot$, när det är fråga om multiplikation av symboler: uttrycket ab skall alltså tolkas som a gånger b. Däremot skriver vi $2 \cdot 3$ för att skilja det från 23. För att ytterligare förvirra, så skall ett bråk på blandad form $3\frac{1}{7}$ inte tolkas som en multiplikation, utan som en *addition* $3 + \frac{1}{7}$. De matematiska symbolerna är tyvärr inte helt entydiga och lätta att begripa sig på. Ofta är det sammanhanget som avgör tolkningen.

§3. Parentes om parenteser och prioriteringsregler

För att fastslå en entydig innebörd av uttryck som $2+3\cdot5$, har man enats om i vilken ordning räkneoperationerna må utföras, en s.k. *prioriteringsordning*. Multiplikation och division genomförs alltid före addition och subtraktion. Alltså är

$$2 + 3 \cdot 5 = 2 + (3 \cdot 5) = 17.$$

Vill vi upphäva denna naturens ordning och istället genomföra additionen först, tillgriper vi parenteser, och skriver t.ex.

$$(2 + 3) \cdot 5 = 25.$$

På samma sätt är[1]

$$3 + 3/6 = 3\frac{1}{2} = 3.5, \qquad \text{medan} \qquad (3 + 3)/6 = 1.$$

Operationerna indelas ofta i två grupper: i den ena gruppen ingår addition och subtraktion och i den andra multiplikation och division. Operationerna i den andra gruppen utförs alltid före operationerna i den första. Inom samma grupp utförs operationerna däremot från vänster till höger. Alltså är

$$3/3 \cdot 3 = (3/3) \cdot 3 = 3, \qquad \text{men} \qquad 3/(3 \cdot 3) = \frac{1}{3}.$$

Parenteser fyller alltså mer än bara ett dekorativt syfte. De skadar aldrig, utan bör gärna tillgripas så fort missförstånd skulle kunna uppstå.

Den, som inte kan få nog av prioriteringsregler, kan gärna googla på följande minnesramsa, som kodar prioriteringsordning:

Please Excuse My Dear Aunt Sally
(Parentheses, Exponentiation, Multiplication/Division, Addition/Subtraction)

Den finns bland annat på Wikipedia, som för övrigt varmt kan rekommenderas som komplement till kurslitteraturen för sina välskrivna matematikartiklar och instruktiva exempel (den engelska är fylligare än den svenska).

§4. Axiomen för reella tal

Som utgångspunkt för vår framställning av algebran tar vi de reella talen $\mathbb{R}$, jämte operationerna addition och multiplikation. Dessa styrs av räknelagar, vilka vi tar till axiom. Vi begynner med addition.

Axiom 1.1: Räknelagar för addition

Addition av reella tal lyder under följande räknelagar.

A1. *Associativa lagen för addition:* $(a + b) + c = a + (b + c)$.

A2. *Kommutativa lagen för addition:* $a + b = b + a$.

A3. *Nolla för addition:* $a + 0 = a$.

A4. *Additiva inverser:* För varje tal a finns en unik *additiv invers* $-a$, sådan att $a + (-a) = 0$.

De första tre lagarna torde alla vara välbekanta eller rentav självklara. Subtraktion lyser med sin frånvaro, bortsett från ett litet minustecken i fjärde lagen. Denna räcker dock för att definiera operationen subtraktion, vilket är den vanliga proceduren.

[1] I denna text väljer vi att, i stället för decimalkomma, använda den internationellt mera accepterade *decimalpunkten*. Vi skriver alltså 3.5 i stället för $3,5$.

Saker blir ibland klarare ju mer filosofisk man är, i varje fall upp till en viss gräns. Vad är alltså egentligen differensen $a - b$ av två reella tal a och b? Uppenbarligen är $(a - b) + b = a$, så att differensen $a - b$ är den entydiga lösningen till ekvationen $x + b = a$.

Detta är ett elegant sätt att definiera differensen, som träffar just den matematiskt väsentliga egenskapen hos subtraktion, förmågan att lösa ekvationer. (Ekonomerna anser eventuellt andra egenskaper vara väsentligare, som subtraktionens förmåga att balansera en budget.) Denna definition har förstås nackdelen, att den inte förser oss med någon information om, hur differensen skall beräknas eller om den ens finns. Men det kan man lösa; om inte annat, så vet närmsta lilla mobiltelefon hur.

Skämt åsido, så garanteras existensen och entydigheten av en lösning till $x + b = a$ just av den fjärde räknelagen ovan. Ty adderas $-b$ till båda leden, fås

$$(x + b) + (-b) = a + (-b).$$

Vänsterledet kan förenklas enligt

$$(x + b) + (-b) \overset{A1}{=} x + (b + (-b)) \overset{A4}{=} x + 0 \overset{A3}{=} x.$$

Ekvationen har därför den entydiga lösningen $x = a + (-b)$. Vi väljer att taga detta till definition av subtraktion.

Definition 1.2

Subtraktion av reella tal definieras av formeln

$$a - b = a + (-b).$$

Eftersom $a + (-a) = 0$ och samtidigt $(-a) + a = 0$, ser vi direkt, att den additiva inversen till $-a$ är a. Vi kan alltså skriva $-(-a) = a$.

Exempel 1.
Räknelagarna för subtraktion fås enkelt från reglerna för addition. Till exempel kan vi bevisa den välbekanta formeln

$$a - (b + c) = a - b - c$$

på följande vis. Eftersom

$$\begin{aligned}(b + c) + ((-b) + (-c)) &\overset{A2}{=} (c + b) + ((-b) + (-c)) \\ &\overset{A1}{=} c + (b + (-b)) + (-c) \overset{A4}{=} c + 0 + (-c) \overset{A3}{=} c + (-c) \overset{A3}{=} 0,\end{aligned}$$

så är $(-b) + (-c) = -(b + c)$ den additiva inversen till $b + c$. Då har vi, enligt ovanstående definition av subtraktion,

$$a - (b + c) = a + (-(b + c)) = a + ((-b) + (-c)) \overset{A1}{=} (a + (-b)) + (-c) = (a - b) - c.$$

Räknelagarna för multiplikation kommer härnäst.

Axiom 1.3: Räknelagar för multiplikation

Multiplikation av reella tal lyder under följande räknelagar.

M1. *Associativa lagen för multiplikation:* $(ab)c = a(bc)$.

M2. *Kommutativa lagen för multiplikation:* $ab = ba$.

M3. *Etta för multiplikation:* $a \cdot 1 = a$.

M4. *Multiplikativa inverser:* För varje tal $a \neq 0$ finns en unik *multiplikativ invers* a^{-1}, sådan att $a \cdot a^{-1} = 1$.

Precis som de additiva inverserna medgav en definition av subtraktion ovan, låter de multiplikativa inverserna oss definiera division. Det leder till bråkräkning, och behandlas utförligt i ett senare avsnitt.

Slutligen finns det en räknelag, som sammanknyter addition och multiplikation.

Axiom 1.4: Distributiva lagen

Addition och multiplikation av reella tal lyder under följande räknelag.

D. *Distributiva lagen:* $(a+b)c = ac+bc$.

Kombineras Distributiva lagen med Kommutativa lagen, får vi lagen $c(a+b) = ca+cb$. Distributiva lagen kan vidare enkelt utvidgas till att gälla fler termer inom parentesen, t.ex. fem:

$$(a+b+c+d+e)f = af+bf+cf+df+ef.$$

Exempel 2.
Den distributiva lagen talar om hur vi kan eliminera parenteser, men vi kan också omvänt använda den för att *faktorisera* uttryck. Till exempel har vi

$$a^2b+ab^2 = a(ab+b^2) = ab(a+b).$$

Räknelagen $a \cdot 0 = 0$ fanns inte med bland de nio räknelagarna ovan. Det beror på, att den kan härledas ur de givna.

Sats 1.5

För varje reellt tal a gäller att

$$a \cdot 0 = 0.$$

Bevis

Vi har

$$a \cdot 0 \overset{A3}{=} a \cdot (0+0) \overset{D}{=} a \cdot 0 + a \cdot 0.$$

Subtraktion av $a \cdot 0$ från båda leden, eller ekvivalent addition med den additiva inversen $-a \cdot 0$, ger $0 = a \cdot 0$.

Vi är nu i stånd att besvara den eviga frågan, varför "minus gånger minus är plus".

Sats 1.6

För alla reella tal a och b gäller att

$$a(-b) = -ab = (-a)b \qquad \text{och} \qquad (-a)(-b) = ab.$$

Bevis

Vi visar att $a(-b) = -ab$. Vi har

$$0 \overset{\text{Sats 1.5}}{=} a \cdot 0 \overset{A4}{=} a(b+(-b)) \overset{D}{=} ab + a(-b).$$

Talet $a(-b)$ är därför den additiva inversen till ab (tillsammans ger de ju noll), så att $a(-b) = -ab$. Identiteten $(-a)b = -ab$ visas på samma sätt. Slutligen har vi, enligt det redan visade, att

$$(-a)(-b) = -a(-b) = -(-ab) = ab.$$

Exempel 3.
Från satsen följer det, att

$$(-1) \cdot a = -(1 \cdot a) = -a.$$

§5. Rationella uttryck

Vad gäller *heltalsbråk* eller *rationella tal* som

$$\frac{1}{2} = 0.5 \qquad \text{och} \qquad \frac{4}{3} = 1.333\ldots$$

är det inget problem att räkna ut (approximativa) värden för de reella tal de står för, d.v.s. de första decimalerna i deras decimalutvecklingar.[2] Vi vill nu kunna förstå och använda brutna uttryck, vari de ingående kvantiteterna är okända, som t.ex.

$$\frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}},$$

som beskriver hur massan förändras med hastigheten i relativitetsteorien. Hur man räknar abstrakt med dem är formulerat i ett antal räkneregler, så löjligt få att lära sig utantill, att det bara visar hur mycket bättre det är för en lat, nöjeslysten person att studera matematik än, säg, ärvdabalken (för att nu ta något extremt tråkigt).

Minns räknelagen M4 ovan: För varje reellt tal $a \neq 0$ garanteras existensen av en unik multiplikativ invers a^{-1}. Den multiplikativa inversen a^{-1} till talet a skrivs även $\frac{1}{a}$. Eftersom $a^{-1}a = aa^{-1} = 1$, så är a den multiplikativa inversen till a^{-1}, alltså $(a^{-1})^{-1} = a$.

Definition 1.7

Division av reella tal definieras av formeln

$$\frac{a}{b} = ab^{-1}, \qquad b \neq 0.$$

Division med noll är alltså odefinierat. Det har sin grund i, att 0 saknar multiplikativ invers; det kan ju omöjligtvis finnas något reellt tal x, med egenskapen att $0 \cdot x = 1$.

Nu kommer vi till principen för förlängning och förkortning av bråk. Att

$$\frac{1}{2} + \frac{1}{3} = \frac{5}{6}$$

övertygade vi oss intuitivt om redan i småskolan genom att skära en pizza i sex delar och räkna sjättedelar (eventuellt äta dem). Men "när man växer mer och mer, blir ju kraven litet fler" (Eva Remaeus i *Fem myror är fler än fyra elefanter*). Vi vill nu genomföra ett matematiskt strängt bevis med axiomen som enda grund.

Sats 1.8

För alla reella tal a, b, c och d, där $b, d \neq 0$, gäller att

$$\frac{a}{b} = \frac{ad}{bd}.$$

[2] Vi använder ordet reellt tal för ett decimalbråk med ändligt eller oändligt antal decimaler. Den djupa frågan om, vad ett reellt tal egentligen *är*, lämnas till högre kurser i analysens grunder.

Bevis

Eftersom

$$(bd)\cdot(b^{-1}d^{-1}) \overset{M2}{=} (db)\cdot(b^{-1}d^{-1}) \overset{M1}{=} d\cdot(bb^{-1})\cdot d^{-1} \overset{M4}{=} d\cdot 1\cdot d^{-1} \overset{M3}{=} d\cdot d^{-1} \overset{M4}{=} 1,$$

så är $b^{-1}d^{-1}$ den multiplikativa inversen till bd, alltså

$$(bd)^{-1} = b^{-1}d^{-1}.$$

Nu följer enkelt att

$$\frac{ad}{bd} = (ad)(bd)^{-1} = adb^{-1}d^{-1} = ab^{-1}dd^{-1} = ab^{-1} = \frac{a}{b}.$$

Exempel 4.
Ett rent numeriskt exempel är

$$\frac{17}{34} = \frac{1\cdot 17}{2\cdot 17} = \frac{1}{2}.$$

Två algebraiska exempel är

$$\frac{x^3y+xy^2}{x^2y+xy^3} = \frac{xy(x^2+y)}{xy(x+y^2)} = \frac{x^2+y}{x+y^2}$$

och

$$\frac{1}{\sqrt{x+1}} = \frac{1\cdot\sqrt{x+1}}{\sqrt{x+1}\cdot\sqrt{x+1}} = \frac{\sqrt{x+1}}{x+1}.$$

Vi redogör nu för hur man tillämpar de fyra räknesätten på bråk (ibland kallade rationella uttryck, kanske för att skilja dem från irrationella utbrott, typ usch) för att få nya rationella uttryck. Addition och subtraktion sker genom förlängning av uttrycken, så att de får samma nämnare:

Sats 1.9

För alla reella tal a, b, c och d, där $b, d \neq 0$, gäller att

$$\frac{a}{b}+\frac{c}{d} = \frac{ad}{bd}+\frac{bc}{bd} = \frac{ad+bc}{bd}.$$

Bevis

Första steget, förlängning, visade vi i föregående sats. Nu är

$$\frac{ad}{bd}+\frac{bc}{bd} = (ad)(bd)^{-1}+(bc)(bd)^{-1} = (ad+bc)(bd)^{-1} = \frac{ad+bc}{bd}$$

enligt definitionen av division och Distributiva lagen.

Exempel 5.
Ett numeriskt exempel på addition av bråk är

$$\frac{1}{10}+\frac{3}{7} = \frac{1\cdot 7+10\cdot 3}{10\cdot 7} = \frac{37}{70}.$$

Två algebraiska exempel är

$$\frac{1}{x}-\frac{1}{y} = \frac{y}{xy}-\frac{x}{xy} = \frac{y-x}{xy}$$

och

$$2y + \frac{1}{x-1} = \frac{2y(x-1)}{x-1} + \frac{1}{x-1} = \frac{2yx - 2y + 1}{x-1}.$$

Multiplikation är enklare, ty det kräver inte lika nämnare:

Sats 1.10

För alla reella tal a, b, c och d, där $b, d \neq 0$, gäller att

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

Bevis

Vi har

$$\frac{a}{b} \cdot \frac{c}{d} = ab^{-1} \cdot cd^{-1} = acb^{-1}d^{-1} = ac(bd)^{-1} = \frac{ac}{bd}$$

enligt definitionen av division samt Associativa och Kommutativa lagen.

Exempel 6.

Ett numeriskt exempel är

$$\frac{1}{2} \cdot \frac{3}{4} \cdot \frac{5}{6} = \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6} = \frac{5}{16}$$

och ett algebraiskt är

$$\frac{x^2y - xy^2}{x+y} \cdot \frac{x+y}{x^2 - 2xy + y^2} = \frac{xy(x-y) \cdot (x+y)}{(x+y) \cdot (x-y)^2} = \frac{xy}{x-y}.$$

Division, slutligen, skall som vanligt vara värst, men är lätt att komma ihåg via mellanledet:

Sats 1.11

För alla reella tal a, b, c och d, där $b, d \neq 0$, gäller att

$$\frac{a}{b} \Big/ \frac{c}{d} = \frac{a}{b} \cdot \frac{d}{c} = \frac{ad}{bc}.$$

Bevis

Eftersom

$$\frac{c}{d} \cdot \frac{d}{c} = \frac{cd}{dc} = 1,$$

så är $\frac{d}{c}$ den multiplikativa inversen till $\frac{c}{d}$, alltså

$$\left(\frac{c}{d}\right)^{-1} = \frac{d}{c}.$$

Definitionen av division ger nu

$$\frac{a}{b} \Big/ \frac{c}{d} = \frac{a}{b} \cdot \left(\frac{c}{d}\right)^{-1} = \frac{a}{b} \cdot \frac{d}{c} = \frac{ad}{bc}.$$

Division med bråket $\frac{c}{d}$ är alltså ekvivalent med multiplikation med $\frac{d}{c}$. Några speciella varianter är extra viktiga att lägga på minnet:

$$1 \Big/ \frac{c}{d} = \frac{d}{c} \qquad \text{och} \qquad 1 \Big/ \frac{1}{d} = d.$$

Månget skolbarn har väl legat vaket om natten, grubblande över, varför $3/\frac{1}{2} = 3 \cdot 2$. I beviset ovan ges alltså den algebraiska förklaringen. Intuitivt förstås detta enklast med s.k. *innehållsdivision:* Hur många halvlitersflaskor kan fyllas med 3 liter saft?

Exempel 7.
Vi önskar lösa ut x ur ekvationen $ax = 7$ (som funktion av a) och sedan beräkna x för $a = \frac{2}{3}$. Genom att dividera med a (tydligen är $a \neq 0$, ty om $a = 0$ får vi ekvationen $0 = 7$, som ju saknar lösningar) får vi $x = \frac{7}{a}$. Om $a = \frac{2}{3}$ är alltså

$$x = \frac{7}{2/3} = \frac{7 \cdot 3}{2} = \frac{21}{2}.$$

Exempel 8.
Vi kan förenkla

$$\frac{\frac{1}{x} + \frac{1}{y}}{\frac{1}{x} - \frac{1}{y}} = \frac{\frac{y}{xy} + \frac{x}{xy}}{\frac{y}{xy} - \frac{x}{xy}} = \frac{y+x}{xy} \Big/ \frac{y-x}{xy} = \frac{y+x}{xy} \cdot \frac{xy}{y-x} = \frac{y+x}{y-x}.$$

§6. Kvadrerings- och konjugatreglerna

Låt oss nu härleda några ytterligare räknelagar. Härledningarna visas kanske pinsamt utförligt för att illustrera tekniker som är användbara i mer komplicerade sammanhang.

Sats 1.12: Kvadreringsregeln

För alla reella tal a och b gäller att

$$(a+b)^2 = a^2 + 2ab + b^2.$$

Bevis

Distributiva lagen D gäller för alla tal, oavsett om de kallas a och b eller något mer fantasifullt. Alltså kan vi tillämpa denna räkneregel, $(a+b)c = ac + bc$, på det speciella talet $c = a + b$ för att få

$$(a+b)^2 = (a+b)(a+b) = (a+b)c = ac + bc = a(a+b) + b(a+b).$$

Tillämpas nu Distributiva lagen på de två parenteserna i högerledet får vi

$$a(a+b) + b(a+b) = aa + ab + ba + bb = a^2 + 2ab + b^2,$$

som önskat.

Du har kanske lärt dig att göra beviset ovan genom rita pilar mellan alla möjliga produkter. Poängen med att göra det på ovanstående sätt är att det är systematiskt och fungerar bättre i mer komplicerade räkningar, när antalet parenteser är många. Man behöver förstås inte införa något c, utan nöja sig med att tänka sig det. Då ser den första räkningen ut som

$$(a+b)^2 = (a+b)(a+b) = (a+b)a + (a+b)b.$$

Kvadreringsregeln $(a+b)^2 = a^2+2ab+b^2$ leder i två riktningar. Startar vi med vänsterledet och ersätter det med högerledet, har vi multiplicerat ihop parenteserna och förenklat uttrycket. Går vi i andra riktningen, så har vi faktoriserat $a^2+2ab+b^2$ till $(a+b)^2$. Beroende på sammanhanget kan vi, litet förvirrande, se bägge operationerna som förenklingar. De två olika uttrycken innehåller ju samma information om samma tal, men formaterad på olika sätt.

Exempel 9.
Vi önskar faktorisera $4x^2 + 4xy + y^2$. Genom att titta på högerledet i Kvadreringsregeln och försöka passa in de givna termerna i detta, ser vi att, med valet $a = 2x$ och $b = y$, Kvadreringsregeln ger

$$4x^2 + 4xy + y^2 = (2x)^2 + 2 \cdot (2x)y + y^2 = (2x + y)^2.$$

Som i beviset för Kvadreringsregeln kan vi med hjälp av Distributiva lagen räkna ut

$$(a + b)(c + d) = (a + b)c + (a + b)d = ac + bc + ad + bd.$$

Sats 1.13: Konjugatregeln

För alla reella tal a och b gäller att

$$(a + b)(a - b) = a^2 - b^2.$$

Bevis

Vi kan helt enkelt sätta in $c = a$ och $d = -b$ i formeln ovan. Gör vi i stället räkningen från grunden, får vi

$$(a + b)(a - b) = (a + b)a - (a + b)b = a^2 + ba - ab - b^2 = a^2 - b^2$$

enligt räknelagarna (vilka?).

Exempel 10.
Enligt Konjugatregeln är

$$(3x + 2y)(3x - 2y) = (3x)^2 - (2y)^2 = 9x^2 - 4y^2.$$

Omvänt kan vi med Konjugatregeln faktorisera

$$8a^2b^2c^2 - 2a^2b^2 = 2a^2b^2(4c^2 - 1) = 2a^2b^2(2c - 1)(2c + 1).$$

Exempel 11.
Ett avancerat exempel är

$$x^4 + 64 = x^4 + 16x^2 + 64 - 16x^2 = (x^2 + 8)^2 - (4x)^2 = (x^2 + 8 + 4x)(x^2 + 8 - 4x),$$

där både Kvadrerings- och Konjugatregeln nyttjats.

Vi kan använda förlängning tillsammans med Konjugatregeln för att förenkla en del algebraiska uttryck som innehåller kvadratrötter. Om

$$p\sqrt{A} + q\sqrt{B}$$

är ett rotuttryck, kallas

$$p\sqrt{A} - q\sqrt{B}$$

dess *konjugat*.

Exempel 12.
Bråket nedan har nämnare $\sqrt{2} - \sqrt{3}$. Förlänger vi detta med konjugatkvantiteten $\sqrt{2} + \sqrt{3}$ kan vi sedan utnyttja Konjugatregeln.

$$\frac{1}{\sqrt{2} - \sqrt{3}} = \frac{1 \cdot (\sqrt{2} + \sqrt{3})}{(\sqrt{2} - \sqrt{3}) \cdot (\sqrt{2} + \sqrt{3})} = \frac{\sqrt{2} + \sqrt{3}}{(\sqrt{2})^2 - (\sqrt{3})^2} = \frac{\sqrt{2} + \sqrt{3}}{-1} = -\sqrt{2} - \sqrt{3}.$$

(Detta är ett exempel på ett snyggt trick, som ibland, men tyvärr rätt sällan, är användbart.)

§7. Binomialsatsen

Nu skall vi bege oss bortom futtiga två parenteser.

Sats 1.14: Kuberingsregeln

För alla reella tal a och b gäller att

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3.$$

Bevis

Vi har $(a+b)^3 = (a+b)^2(a+b)$ enligt definitionen av vad exponenter innebär (mer om detta senare). Tillämpas Distributiva lagen och sedan Kvadreringsregeln, får vi

$$(a+b)^3 = (a+b)^2(a+b) = (a+b)^2a + (a+b)^2b = (a^2+2ab+b^2)a + (a^2+2ab+b^2)b.$$

Åter kan vi tillämpa Distributiva lagen på högerledets två parenteser och få

$$(a+b)^3 = a^3 + 2aba + b^2a + a^2b + 2ab^2 + b^3 = a^3 + 3a^2b + 3ab^2 + b^3,$$

vilket var vad vi ville komma till.

Ett allmänt uttryck för $(a+b)^n$ ges av det så kallade *Binomialsatsen*. Först måste vi införa litet notation.

Definition 1.15

För naturliga tal n, definieras n-**fakultet** genom formeln

$$n! = 1 \cdot 2 \cdots (n-1) \cdot n, \qquad n \geq 1,$$

och dessutom $0! = 1$.

T.ex. är

$$0! = 1 \qquad 3! = 3 \cdot 2 \cdot 1 = 6$$
$$1! = 1 \qquad 4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$$
$$2! = 2 \cdot 1 = 2 \qquad 5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$$

Den säregna konventionen $0! = 1$ kommer att visa sig mycket praktisk.

Vi inför nu beteckningen

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Dessa är de så kallade **binomialkoefficienterna**. Till exempel är

$$\binom{4}{2} = \frac{4!}{2!2!} = \frac{24}{2 \cdot 2} = 6.$$

Vi har alltid

$$\binom{n}{0} = \binom{n}{n} = \frac{n!}{0!n!} = 1$$

och

$$\binom{n}{1} = \binom{n}{n-1} = \frac{n!}{1!(n-1)!} = n.$$

Vi kommer att återvända till dessa tal och ge dem en kombinatorisk tolkning. Just nu är vi bara intresserade av de algebraiska aspekterna.

Sats 1.16: Binomialsatsen

Låt a och b vara reella tal, och låt n vara ett naturligt tal. Då gäller att

$$(a+b)^n = \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \cdots + \binom{n}{k}a^{n-k}b^k + \cdots + \binom{n}{n-1}ab^{n-1} + \binom{n}{n}b^n.$$

Beviset kommer i slutet av kursen, i kapitlet om kombinatorik. Ett **binom** är en summa av två termer, som $x+y$, härav namnet binomialkoefficienter för talen $\binom{n}{k}$.

Exempel 13.
För $n = 4$ ger Binomialsatsen

$$\begin{aligned}(a+b)^4 &= \binom{4}{0}a^4 + \binom{4}{1}a^3b + \binom{4}{2}a^2b^2 + \binom{4}{3}ab^3 + \binom{4}{4}b^4 \\ &= \frac{4!}{0!4!}a^4 + \frac{4!}{1!3!}a^3b + \frac{4!}{2!2!}a^2b^2 + \frac{4!}{3!1!}ab^3 + \frac{4!}{4!0!}b^4 \\ &= a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4.\end{aligned}$$

§8. Olikheter

För olikheter behövs friska axiom. Vi utgår från ordningsrelationen $a < b$ för reella tal, som förstås utläses "a är mindre än b". Skrivsättet $a \leq b$ betyder att $a < b$ eller $a = b$.

Axiom 1.17: Räknelagar för olikheter

Olikheter för reella tal lyder under följande räknelagar.

O1. *Reflexiva lagen.* $a \leq a$.

O2. *Antisymmetriska lagen.* $a \leq b$ och $b \leq a$ medför $a = b$.

O3. *Transitiva lagen.* $a \leq b$ och $b \leq c$ medför $a \leq c$.

O4. *Trikotomilagen.* För två reella tal a och b gäller precis en av möjligheterna $a < b$, $a = b$ eller $a > b$.

O5. *Additiv isotoni.* $a \leq b$ medför $a + c \leq b + c$.

O6. *Multiplikativ isotoni.* $a, b \geq 0$ medför $ab \geq 0$.

För att få en intuitiv bild av tal och olikheter är reella tallinjen oslagbar. Olikheten $a < b$ tolkas som att punkten (som svarar mot talet) a ligger till vänster om punkten b på tallinjen. Detta låter oss skapa en konkret bild av sådant abstrakt nonsens som den transitiva lagen O3 ovan. I bilden säger den bara, att om a ligger till vänster om b, och b ligger till vänster om c, så ligger a till vänster om c.

För att få $a + c$ från a, flyttar vi punkten a precis c steg från a, åt höger, om c är positivt, eller åt vänster, om c är negativt. Vi kan alltså tolka addition som förflyttning av punkter längs tallinjen.

Vi kan också göra oss en bild av varför den additiva isotonien O5 kan förväntas gälla. Lagen säger följande i geometriskt tungomål: Om a ligger till vänster om b, och vi flyttar båda dessa punkter c steg (åt vänster eller åt höger) till $a + c$ respektive $b + c$, så liggger $a + c$ till vänster om $b + c$. Effektiv problemlösning handlar mycket om att skapa sig enkla intuitiva bilder som dessa, ett slags principskisser över hur situationen ser ut.

Vi kommer inte att behöva hänvisa explicit till lagarna O1–O3 ovan. De uttrycker mer eller mindre vad man förväntar sig av en ordningsrelation, men vi formulerade dem ovan för fullständighetens skull. Enligt Trikotomilagen O4 gäller, för ett reellt tal a, precis en av möjligheterna $a > 0$, $a = 0$ eller $a < 0$. Talen $a > 0$ kallas **positiva**, talen $a < 0$ kallas **negativa**. Subtraktion med a i bägge leden av $a > 0$ ger $0 > -a$. Således är $-a$ negativ, då a är positiv, och omvänt.

Additiv isotoni O5 betyder, att samma tal kan adderas eller subtraheras till en olikhet. Exempelvis är $1 < x - 2$ ekvivalent med $3 < x$. Däremot är det som bekant inte sant, att en olikhet kan multipliceras med valfritt tal. Multipliceras olikheten $2 < 3$ med 2, fås den sanna olikheten $4 < 6$, men multiplikation med -1 ger den falska olikheten $-2 < -3$. Lugn, det är inte kaos. Det som händer, när vi multiplicerar med ett negativt tal, är att olikheten byter riktning. Vi kan förstå varför genom ett exempel.

Exempel 14.
Antag att $a \leq b$, och att vi vill multiplicera bägge sidor med med -1. Det kan vi åstadkomma på följande vis. Subtrahera först a från bägge led till $0 \leq b - a$. Subtraheras sedan b från bägge led fås $-b \leq -a$, som önskat.

Vi bevisar formellt en generell regel.

Sats 1.18

Om $a \leq b$, så är

$$\begin{cases} ac \leq bc & \text{om } c > 0 \\ ac \geq bc & \text{om } c < 0. \end{cases}$$

Motsvarande gäller ifall det råder sträng olikhet $a < b$.

Bevis

Vi har $bc - ac = (b - a)c$ och $b - a \geq 0$ enligt antagandet. Om då $c > 0$, ger den multiplikativa isotonien

$$bc - ac = (b - a)c \geq 0.$$

Addition av ac till båda leden ger $bc \geq ac$ enligt additiv isotoni.

Vi har också $ac - bc = (b - a)(-c)$. Om $c < 0$, är $-c > 0$, och multiplikativ isotoni ger åter

$$ac - bc = (b - a)(-c) \geq 0.$$

Addition av bc till båda leden ger $ac \geq bc$ enligt additiv isotoni.

Det är välkänt, att produkten av två positiva, liksom två negativa, tal är positiv, men produkten av ett positivt och ett negativt tal är negativ. Vi formulerar och bevisar även detta formellt.

Sats 1.19

Produkten ab är positiv om a och b har samma tecken, men negativ om a och b har olika tecken.

Bevis

Om $a, b > 0$, så är $ab \geq 0$ enligt multiplikativ isotoni. Eftersom $ab = 0$ inte kan komma på fråga (i så fall vore en av a och b lika med 0), så är alltså $ab > 0$. Om $a > 0 > b$, så är $a, -b > 0$. Enligt vad vi redan bevisat, är då $-ab = a(-b) > 0$, så att $ab < 0$. Övriga fall lämnas åt läsaren att tänka igenom.

Exempel 15.
Det följer att en kvadrat alltid är icke-negativ, ty $a^2 = a \cdot a$ är produkten av antingen två positiva tal, två negativa tal eller två nollor. I de två första fallen är produkten positiv, i det tredje fallet noll.

Olikheter kan i allmänhet ej kvadreras, ty kvadrering av $-3 < 2$ ger den falska olikheten $9 < 4$. En olikhet *får* kvadreras, då båda led är icke-negativa:

Sats 1.20

Låt $a, b \geq 0$. Då är $a \leq b$ om och endast om $a^2 \leq b^2$.

Bevis

Om $a \leq b$ är $b - a \geq 0$. Då är $b^2 - a^2 = (b + a)(b - a)$ produkten av två icke-negativa faktorer och därför icke-negativ enligt föregående sats. Om omvänt $a > b \geq 0$, så är $b - a < 0$ och $b + a > 0$, så att $b^2 - a^2$ är produkten av en positiv och en negativ faktor, och därför negativ.

§9. Medelvärden

Det mest använda medelvärdet, vårt vanliga, är det aritmetiska medelvärdet.

Definition 1.21

Det **aritmetiska medelvärdet** av två tal a och b är

$$A(a, b) = \frac{a + b}{2}.$$

Notera att medelvärdet ligger instängt mellan talen:

$$a \leq A(a, b) = \frac{a + b}{2} \leq b, \qquad \text{om } a \leq b.$$

(Enkel övning!) Det är en väsentlig egenskap, för att något skall förtjäna att kallas medelvärde. Det följer förstås, att $A(a, a) = a$, när de ingående talen är lika.

Ibland kan det vara lämpligt, att vikta de ingående talen olika, och betrakta t.ex. det viktade aritmetiska medelvärdet $\frac{9}{10}a + \frac{1}{10}b$. Orsaken kan vara, att man anser mätvärdet a från en viss källa vara 9 gånger mer troligt att vara korrekt än värdet b från en mer slarvig och kanske suspekt källa. Då skulle man kunna välja att bara ange a, men då slänger man samtidigt bort information, en av Kyrkans sju dödssynder. (Bland de övriga märks att dela med 0, och känna åtrå efter sin nästas datoralgebrasystem.) I stället tilldelar man alltså den suspekta informationen mindre vikt.

Det aritmetiska medelvärdet av flera tal $a_1, \ldots, a_n$ definieras på samma sätt som

$$A(a_1, \ldots, a_n) = \frac{a_1 + a_2 + \cdots + a_n}{n}.$$

Att summera talen och dividera med antalet tal är ett sätt att få fram ett medelvärde. Men det finns andra.

Definition 1.22

Det **geometriska medelvärdet** av två icke-negativa tal a och b är

$$G(a, b) = \sqrt{ab}.$$

Talen måste här förutsättas vara icke-negativa, för att kvadratroten säkert skall kunna utdragas. Även det geometriska medelvärdet har egenskaperna att

$$a \leq G(a,b) = \sqrt{ab} \leq b, \qquad \text{om } a \leq b,$$

samt att, $G(a,a) = a$, när de ingående talen är lika.

Det geometriska medelvärdet av flera icke-negativa tal $a_1, \ldots, a_n$ definieras på samma sätt som

$$G(a_1, \ldots, a_n) = \sqrt[n]{a_1 a_2 \cdots a_n}.$$

Exempel 16.
Talen 2 och 18 har det aritmetiska medelvärdet $\frac{2+18}{2} = 10$ och det geometriska medelvärdet $\sqrt{2 \cdot 18} = 6$.

I exemplet är $A \geq G$, och detta är inte någon slump. Det är ett allmänt faktum, känt som *Aritmetisk–geometriska olikheten*. Vi formulerar den för två involverade tal, trots att resultatet äger beständighet även för flera.

Sats 1.23: Aritmetisk–geometriska olikheten

För icke-negativa tal a, b gäller olikheten

$$A(a,b) = \frac{a+b}{2} \geq \sqrt{ab} = G(a,b),$$

med likhet precis då $a = b$.

Bevis

En kvadrat är alltid icke-negativ enligt ovan; sålunda är

$$0 \leq (\sqrt{a} - \sqrt{b})^2 = a + b - 2\sqrt{ab},$$

vilket är ekvivalent med $\frac{a+b}{2} \geq \sqrt{ab}$.

Övningar

1.1. Utveckla

(a) $(a-b)^3$;
(b) $(a+b+c)^2$.

1.2. Faktorisera

(a) $x^3 - 9x$;
(b) $a^2 - 9a + ab - 9b$;
(c) $(x+a)^2 - b^2$;
(d) $x^6 - x^4 + x^2 - 1$.

1.3. Förenkla

$$\frac{2}{3x+9} + \frac{x}{x^2-9} - \frac{1}{2x-6}.$$

1.4. Förenkla

(a) $\dfrac{x^4 - y^4}{x+y}$;

(b) $\dfrac{\frac{16x^4}{81} - y^4}{\frac{2x}{3} + y}$;

(c) $\dfrac{\frac{1}{x+1} + \frac{1}{x-1}}{\frac{1}{x-1} - \frac{1}{x+1}}$.

1.5. Förenkla så långt som möjligt

(a) $\dfrac{\frac{1}{ab}}{\frac{1}{a} + \frac{1}{b}}$;

(b) $\dfrac{b}{a^2 + ab} + \dfrac{1}{a+b}$;

(c) $\dfrac{\frac{1}{a-b} - \frac{1}{a}}{\frac{1}{a} - \frac{1}{a+b}}$.

1.6. Förläng bråket

$$\frac{2+\sqrt{5}}{2-\sqrt{5}},$$

så att nämnaren blir fri från kvadratrötter.

1.7. Utveckla $(a+b)^5$.

1.8. Vilken är den konstanta termen i utvecklingen av

$$\left(2x^2 + \frac{1}{4x}\right)^{99} ?$$

1.9. Bestäm koefficienten för x^2 i utvecklingen av

$$\left(\frac{x}{3} - \frac{1}{x^3}\right)^{10}.$$

1.10. Visa, att olikheten

$$\frac{a}{b} + \frac{b}{a} \geq 2$$

gäller för alla positiva tal a och b.

1.11. Det *harmoniska medelvärdet* av två positiva tal a och b definieras som

$$H(a,b) = \frac{2}{\frac{1}{a} + \frac{1}{b}}.$$

(a) Finn det aritmetiska, geometriska och harmoniska medelvärdet av talen 1 och 2.

(b) Visa, att det harmoniska medelvärdet uppfyller kravet på ett medelvärde, nämligen $a \leq H(a,b) \leq b$, om $a \leq b$.

(c) Bevisa den *geometrisk–harmoniska olikheten*

$$G(a,b) \geq H(a,b).$$

1.12. (a) Visa följande generalisering av Konjugatregeln:

$$a^3 - b^3 = (a-b)(a^2 + ab + b^2).$$

(b) Hur skulle uttrycket $a^3 + b^3$ kunna faktoriseras?

(c) Visa den allmänna identiteten

$$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \cdots + ab^{n-2} + b^{n-1}).$$

1.13. Visa identiteten

$$a^2(a+b)^2 + b^2(a+b)^2 + a^2b^2 = (a^2 + b^2 + ab)^2.$$

1.14. Visa, att

$$\sqrt{1 + n(n+1)(n+2)(n+3)}$$

är ett heltal för varje heltal n.

Kapitel 2

Talföljder och summor

§1. Aritmetiska talföljder och summor

Om Gauss barndom berättades det, att han hade en folkskolelärare av den gamla folkilskna sorten (den där sorten, ni vet, som egentligen inte vill vara lärare). En dag gav han barnen i uppdrag att summera de första hundra positiva heltalen. Han tänkte sig nog hinna med både en och två koppar kaffe i lugn och ro på lärarrummet. Barnen stönade och det blev ett förfärligt kackalorum av kritors raspande på små griffeltavlor (barnen hade i regel inga anteckningsböcker). Magistern vände segervisst ryggen till.

Historien har två slut. Ena versionen är, att magistern knappt hunnit vända ryggen till, förrän det hördes ett ivrigt klapprande av träskor, då lille Gauss skyndade att visa resultatet av sin uträkning. Andra versionen är, att magistern återvände, tämligen kaffestinn, och fann alla barnen, sånär som på ett, ha räknat fel någonstans längs vägen. Endast Gauss hade bestått provet.

I princip är det förstås ingen konst, givet att vi har en timmes välbetald arbetstid och tillräckligt med papper. Men hur gjorde då Gauss? Han satte

$$S = 1 + 2 + 3 + \cdots + 98 + 99 + 100,$$

och exekverade sedan följande trick. Addera S med sig själv i omvänd ordning:

$$\begin{array}{rcccccccccccccc} S & = & 1 & + & 2 & + & 3 & + & \cdots & + & 98 & + & 99 & + & 100 \\ S & = & 100 & + & 99 & + & 98 & + & \cdots & + & 3 & + & 2 & + & 1 \\ \hline 2S & = & 101 & + & 101 & + & 101 & + & \cdots & + & 101 & + & 101 & + & 101 \end{array}$$

Summan i högerledet är $100 \cdot 101 = 10\,100$, varav $S = 5050$. Detta var vad som stod skrivet på Gauss tavla. Säger historien. Och som alla bra historier, så är den uppdiktad. Den tycks sakna belägg från Gauss själv.

Följden $1, 2, 3, \ldots, 100$ är ett exempel på en *aritmetisk talföljd*. Differensen mellan två konsekutiva tal är konstant, här 1.

Definition 2.1

Talföljden

$$a_0,\ a_1,\ \ldots,\ a_{n-1},\ a_n,\ a_{n+1},\ \ldots$$

är **aritmetisk**, om differensen mellan två på varandra följande tal i följden är konstant d, alltså $a_n - a_{n-1} = d$ för varje n.

Startande med a_0, fås nästa tal i följden genom att addera d, d.v.s. $a_1 = a_0 + d$. Därefter följer

$$\begin{aligned} a_2 &= a_1 + d = (a_0 + d) + d = a_0 + 2d \\ a_3 &= a_2 + d = (a_0 + 2d) + d = a_0 + 3d, \end{aligned}$$

och allmänt

$$a_n = a_0 + nd.$$

Exempel 1.
Hur många tal innehåller den aritmetiska talföljden

$$13,\ 10,\ \ldots,\ -44\ ?$$

Kalla följden $a_0, a_1, \ldots, a_n$, så att antalet tal i följden är $n+1$. Uppenbarligen är differensen $d = 10 - 13 = -3$, det första talet $a_0 = 13$ samt det sista talet $a_n = -44$. Från $-44 = a_0 + nd = 13 - 3n$ fås $n = 19$. Antalet tal i följden är $19 + 1 = 20$.

På samma sätt som i det inledande exemplet kan vi lätt beräkna summan av en aritmetisk talföljd.

Sats 2.2

Summan av talen i den aritmetiska talföljden $a_0, a_1, a_2, \ldots, a_n$ är

$$a_0 + a_1 + a_2 + \cdots + a_n = \frac{1}{2}(n+1)(a_0 + a_n).$$

Satsen kan formuleras som att den aritmetiska summan är medelvärdet av första och sista termen, gånger antalet termer i följden. Klokt kan vara, att memorera den som

$$\frac{1}{2}(\text{antalet termer})(\text{första termen} + \text{sista termen}).$$

Bevis

Vi försöker upprepa succén från Gauss pojkstreck. Först observerar vi, att

$$a_k + a_{n-k} = (a_0 + kd) + (a_0 + (n-k)d) = a_0 + (a_0 + nd) = a_0 + a_n$$

för varje k, eftersom $a_k = a_0 + kd$ enligt ovan. Därefter adderar vi summan till sig själv i omvänd ordning:

$$\begin{array}{rcccccccccc} S & = & a_0 & + & a_1 & + & \cdots & + & a_{n-1} & + & a_n \\ S & = & a_n & + & a_{n-1} & + & \cdots & + & a_1 & + & a_0 \\ \hline 2S & = & (a_0 + a_n) & + & (a_0 + a_n) & + & \cdots & + & (a_0 + a_n) & + & (a_0 + a_n) \end{array}$$

Summan är $2S = (n+1)(a_0 + a_n)$, varav den önskade formeln följer.

Exempel 2.
Som ett enkelt specialfall har vi

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

Exempel 3.
Vi skall beräkna summan av alla udda tal mellan 3 och 10 011. Dessa bildar en aritmetisk följd med differensen $d = 2$. Den allmänna formeln är $a_k = 3 + 2k$, där första talet $a_0 = 3$ och sista talet $10011 = a_n = 3 + 2n$, vilket ger $n = 5004$. Följden innehåller 5005 tal och dess summa är

$$3 + 5 + \cdots + 10\ 011 = \frac{1}{2} \cdot 5005 \cdot (3 + 10\ 011) = 25\ 060\ 035.$$

§2. Geometriska talföljder och summor

I de aritmetiska talföljderna fick vi nästa tal i följden genom att addera ett visst tal d till det föregående. Den *geometriska talföljden* springer ur samma idé, men byter additionen mot

multiplikation.

Exempel 4.
I den geometriska talföljden

$$1,\ \frac{1}{10},\ \frac{1}{100},\ \frac{1}{1000},\ \frac{1}{10\,000},\ldots$$

fås varje tal genom att multiplicera föregångaren med $\frac{1}{10}$. Vi kan också uttrycka det så, att kvoten mellan två på varandra följande element i följden är konstant $\frac{1}{10}$.

Definition 2.3

Talföljden

$$a_0,\ a_1,\ \ldots,\ a_{n-1},\ a_n,\ a_{n+1},\ \ldots$$

är **geometrisk**, om kvoten mellan två på varandra följande tal i följden är konstant k, alltså $\frac{a_n}{a_{n-1}} = k$ för varje n.

Startande med a_0, fås nästa tal i följden genom att multiplicera med k, d.v.s. $a_1 = a_0k$. Därefter följer

$$a_2 = a_1k = (a_0k)k = a_0k^2$$
$$a_3 = a_2k = (a_0k^2)k = a_0k^3,$$

och allmänt

$$a_n = a_0k^n.$$

Exempel 5.
Hur många tal innehåller den geometriska talföljden

$$256,\ 128,\ 64,\ \ldots,\ \frac{1}{4}\ ?$$

Kalla följden $a_0, a_1, \ldots, a_n$, så att antalet tal i följden är $n+1$. Uppenbarligen är kvoten $k = \frac{128}{256} = \frac{1}{2}$, det första talet $a_0 = 256$ samt det sista talet $a_n = \frac{1}{4}$. Från $\frac{1}{4} = a_0k^n = 256 \cdot (\frac{1}{2})^n$ fås $2^n = 1024$ och $n = 10$. Antalet tal i följden är $10 + 1 = 11$.

(Du kan väl alla 2-potenser upp till $2^{10} = 1024$? Mer seriöst, så är logaritmer, som lärs ut i analysen, det mer systematiska och icke-nerdiga sättet att hitta det n, för vilket $2^n = 1024$. Tages nämligen tvålogaritmen av bägge sidor, erhålles $n = \log_2 1024 = 10$, enligt räknedosan.)

Om schackspelets uppfinnare berättas det (och det är säkert påhitt, eftersom vi säger så), att han fick ynnesten att kunna utbedja sig valfri gåva av Persiens konung. Hans önskan föreföll i förstone vara av det modesta slaget, för ty han önskade sig allenast en mängd ris: 1 riskorn på schackbrädets första ruta, 2 korn på den andra, 4 på den tredje, 8 på den fjärde, och så vidare. Särskilt modest var det ju nu inte. Bara på den sextiofjärde rutan ensam skulle det alltså torna upp sig en hög med 2^{63} riskorn (och det är väldigt *mycket* ris). Frågan vi ställer oss är, hur mycket ris det totalt skulle krävas.

Vi önskar alltså beräkna den geometriska summan (med kvoten 2)

$$S = 1 + 2 + 4 + 8 + \cdots + 2^{63}$$

och brukar härvidlag ett trick liknande Gauss snilleblixt. Vi subtraherar summan från sig själv, men multiplicerad med 2 och förskjuten ett steg:

$$\begin{array}{rcccccccccccccc}
S &=& 1 &+& 2 &+& 4 &+& \cdots &+& 2^{62} &+& 2^{63} & & \\
2S &=& & & 2 &+& 4 &+& \cdots &+& 2^{62} &+& 2^{63} &+& 2^{64} \\
\hline
-S &=& 1 &+& 0 &+& 0 &+& \cdots &+& 0 &+& 0 &-& 2^{64}
\end{array}$$

Prislappen för schackspelet belöpte sig således på $S = 2^{64} - 1$ riskorn.

Som en övning i överslagsräkning kan vi förströ oss med att approximera detta tal. Eftersom $2^{10} = 1024 \approx 10^3$, kan vi avrunda

$$2^{64} - 1 \approx 2^{64} = 2^4 \cdot (2^{10})^6 \approx 16 \cdot (10^3)^6 = 16 \cdot 10^{18}.$$

Ett riskorn väger cirka 25 mg, vilket ger en massa på

$$25 \cdot 16 \cdot 10^{18} \text{ mg} = 400 \cdot 10^{18} \text{ mg} = 400 \cdot 10^{12} \text{ kg} = 400 \cdot 10^9 \text{ t},$$

tusen gånger världens samlade risproduktion.

Det är lätt att generalisera schackproblemet till en formel för en godtycklig geometrisk summa.

Sats 2.4

Summan av talen i den geometriska talföljden $a_0, a_1, a_2, \ldots, a_n$ är

$$a_0 + a_1 + a_2 + \cdots + a_n = \frac{a_0 - a_{n+1}}{1 - k},$$

om $k \neq 1$ är kvoten.

Som stöd för minnet, kan vi tänka oss formeln som

$$a_0 + a_1 + a_2 + \cdots + a_n = \frac{\text{första termen} - \text{supersista termen}}{1 - \text{kvoten}},$$

där den *supersista* termen är a_{n+1}, den term, som skulle följa efter den sista a_n (vi tackar våra påhittiga studenter).

Om kvoten är $k = 1$ fungerar inte formeln, men då är å andra sidan följden konstant, och vi saknar behov av formel.

Bevis

Vi upprepar idén från schackproblemet. Först observerar vi, att $a_m - a_{m-1}k = 0$ för varje m, ty $\frac{a_m}{a_{m-1}} = k$. Vi subtraherar summan från sig själv, multiplicerad med kvoten k:

$$\begin{array}{rcccccccccccc} S &=& a_0 &+& a_1 &+& a_2 &+& \cdots &+& a_{n-1} &+& a_n \\ kS &=& & & a_0k &+& a_1k &+& \cdots &+& a_{n-2}k &+& a_{n-1}k &+& a_nk \\ \hline S - kS &=& a_0 &+& 0 &+& 0 &+& \cdots &+& 0 &+& 0 &-& a_nk \end{array}$$

Summan är $(1-k)S = a_0 - a_nk = a_0 - a_{n+1}$, varav den önskade formeln följer.

Exempel 6.
Ett viktigt specialfall är

$$1 + k + k^2 + \cdots + k^n = \frac{1 - k^{n+1}}{1 - k}.$$

Exempel 7.
Vi skall beräkna den geometriska summan

$$512 + 256 + \cdots + \frac{1}{4}.$$

Kvoten är uppenbarligen $\frac{1}{2}$, och talet som skulle komma efter $\frac{1}{4}$ (den supersista termen), är $\frac{1}{8}$. Därmed är summan lika med

$$\frac{512 - \frac{1}{8}}{1 - \frac{1}{2}} = 1024 - \frac{1}{4} = 1023\frac{3}{4}.$$

Exempel 8.
Betrakta den geometriska summan

$$1+\frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2^n}=\frac{1-\frac{1}{2^{n+1}}}{1-\frac{1}{2}}=2-\frac{1}{2^n}.$$

När n går mot oändligheten, går $\frac{1}{2^n}$ mot 0. I gränsen får vi summan av den oändliga *geometriska serien*

$$1+\frac{1}{2}+\frac{1}{4}+\cdots=2.$$

§3. Summa- och produktnotation

Istället för att skriva en summa, exempelvis summan av de första tio heltalskvadraterna, som

$$S=1+4+9+16+25+\cdots+100,$$

varest den intelligente läsaren förväntas gissa mönstret, skriver man vanligen

$$S=\sum_{k=1}^{10}k^2.$$

Denna notation uttrycker att vi tar summan av alla tal k^2, då heltalet k promenerar från 1 till 10. Vill man istället taga produkten av dessa kvadrattal, skriver man

$$1\cdot 4\cdot 9\cdot\cdots\cdot 100=\prod_{k=1}^{10}k^2.$$

Namnet på variabeln är totalt oväsentligt. Vilken bokstav som helst duger, så att kvadratsumman ovan lika gärna kunde skrivas

$$\sum_{i=1}^{10}i^2=\sum_{p=1}^{10}p^2=\sum_{\ddot\smile=1}^{10}\ddot\smile^2.$$

Variabeln må förstås inte dyka upp på något annat ställe i summan. Vi kan skriva

$$n+4n+9n+\cdots+100n=\sum_{k=1}^{10}k^2n,$$

men

$$\sum_{n=1}^{10}(n^2\cdot n)=1^3+2^3+3^3+\cdots+10^3$$

betyder något annat.

Vad menar vi med

$$\sum_{k=10}^{10}k^2?$$

Det är ont om heltal mellan 10 och 10. Det finns bara ett, nämligen 10, så tolkningen är

$$\sum_{k=10}^{10}k^2=10^2.$$

Givet en talföljd $a_m,\ldots,a_n$, definierar vi, mera allmänt,

$$\sum_{k=m}^{n}a_k=a_m+a_{m+1}+\cdots+a_n,$$

och

$$\prod_{k=m}^{n} a_k = a_m \cdot a_{m+1} \cdot \cdots \cdot a_n.$$

Vi ser enkelt, att

$$\sum_{k=m}^{n}(a_k + b_k) = \sum_{k=m}^{n} a_k + \sum_{k=m}^{n} b_k \qquad \text{och} \qquad \prod_{k=m}^{n} a_k b_k = \prod_{k=m}^{n} a_k \cdot \prod_{k=m}^{n} b_k.$$

Exempel 9.
Summan

$$\sum_{k=1}^{10}(3k + 1 + 3^k)$$

är en summa av den aritmetiska följden $3k+1$ och den geometriska följden 3^k. Den aritmetiska går från 4 (när $k = 1$) till 31 (när $k = 10$) i tio steg och har summan

$$\sum_{k=1}^{10}(3k + 1) = \frac{1}{2} \cdot 10(4 + 31) = 175.$$

Den geometriska går från 3 (när $k = 1$) till 3^{10} (när $k = 10$) och har summan

$$\sum_{k=1}^{10} 3^k = \frac{3 - 3^{11}}{1 - 3} = \frac{3^{11} - 3}{2}.$$

Den sökta summan är därför

$$\sum_{k=1}^{10}(3k + 1 + 3^k) = 175 + \frac{3^{11} - 3}{2}.$$

Övningar

2.1. (a) Skriv summan och produkten av de första n positiva heltalen med hjälp av summa- resp. produktnotation.

(b) Skriv Binomialsatsen med hjälp av summanotation.

2.2. (a) Skriv summan av de första n jämna heltalen med summanotation, samt beräkna denna.

(b) Skriv summan av de första n udda heltalen med summanotation, samt beräkna denna.

2.3. Beräkna summan $\sum_{k=1}^{n}(3k - 2)$.

2.4. Ett antikvariat erbjuder dig att få en bok för varje två böcker du lämnar in. Du har 1024 lästa böcker. Hur många böcker kan du maximalt läsa för dessa?

2.5. Beräkna följande geometriska summor.

(a) $1 + 2 + \cdots + 512$;

(b) $3 + 1 + \frac{1}{3} + \frac{1}{9} + \frac{1}{27}$;

(c) $1 - 2 + 4 - \cdots + 1024$;

(d) $1 + x + x^2 + \cdots + x^9$;

(e) $1 - x + x^2 + \cdots - x^9$.

2.6. Beräkna summorna och produkterna

(a) $\sum_{k=1}^{3} \frac{1}{k}$;

(b) $\sum_{i=0}^{10} 3 \cdot 2^i$;

(c) $\sum_{i=m}^{n} 3 \cdot 2^i$, där $m \leq n$;

(d) $\sum_{i=0}^{2n} 3 \cdot x^{-i}$;

(e) $\sum_{i=0}^{2n} 3 \cdot x^{-n}$;

(f) $\prod_{k=1}^{n} 4$;

(g) $\prod_{k=1}^{n} \frac{k}{k+1}$.

2.7. Kan det finnas en icke-konstant följd av tre tal, som är både aritmetisk och geometrisk?

2.8. Talen

$$\frac{1}{b-a}, \quad \frac{1}{2b}, \quad \frac{1}{b-c}$$

bildar en aritmetisk talföljd. Visa, att a, b, c bildar en geometrisk talföljd.

Kapitel 3

Potenser

Kom ihåg, att uttrycket 3^2 är definierat som $3 \cdot 3$, och att på samma sätt är $3^5 = 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3$. Mera allmänt är

$$3^n = 3 \cdot \cdots \cdot 3$$

produkten av n stycken treor för ett positivt heltal n.

Vi skall i detta kapitel definiera och studera potenser a^b, där $a > 0$, för successivt allmännare exponenter.

§1. Positiva heltalsexponenter

Definition 3.1

Låt $a > 0$ och låt n vara ett positivt heltal. Då definieras **potensen** a^n med **bas** a och **exponent** n som

$$a^n = a \cdot \cdots \cdot a \qquad (n \text{ faktorer}).$$

Vi observerar, att

$$a^2 \cdot a^5 = (a \cdot a) \cdot (a \cdot a \cdot a \cdot a \cdot a) = a^{2+5},$$

och på samma sätt att

$$a^m a^n = a^{m+n}$$

för godtyckliga positiva heltalsexponenter m och n. Vad gäller division av potenser, så har vi

$$\frac{a^5}{a^2} = \frac{a \cdot a \cdot a \cdot a \cdot a}{a \cdot a} = a \cdot a \cdot a = a^{5-2},$$

och, mera allmänt,

$$\frac{a^m}{a^n} = a^{m-n}$$

för $m > n$ (exponenten $m - n$ måste vara ett positivt heltal).

Nu sammanfattar vi de viktiga lagarna för potenser. Och sorry, de måste läras utantill. Enda trösten är att de är försvinnande få jämfört med antalet ben i människokroppen (stackars medicinare), samt att de förstås är så tillfredsställande att använda som raggningsfraser på krogen. (I kaffet använder vi bitsocker, på krogen raggsockor.) Ett gott sätt att lägga upp förståelsen för lagarna är, att tänka igenom dem ordentligt för positiva heltal eller räkna några konkreta numeriska exempel.

Sats 3.2: Potenslagarna

Låt $a, b > 0$. För positiva heltalsexponenter m och n gäller följande räknelagar.

- $a^m a^n = a^{m+n}$.
- $a^m / a^n = a^{m-n} \quad (m > n)$.
- $(a^m)^n = a^{mn}$.
- $a^m b^m = (ab)^m$.
- $a^m / b^m = (a/b)^m$.

Bevis

Vi argumenterade ovan för de första två av dessa. Lagen för exponentiering, $(a^m)^n = a^{mn}$, är även den en omedelbar konsekvens av definitionen. Antalet faktorer a i

$$(a^m)^n = a^m \cdot a^m \cdot \cdots \cdot a^m = (a \cdot \cdots \cdot a)(a \cdot \cdots \cdot a) \cdots (a \cdot \cdots \cdot a)$$

är ju precis mn (vi har n parenteser med m stycken a i varje). Återstående två lagar lämnas åt läsaren.

§2. Naturliga exponenter

Vi skall nu successivt utvidga definitionen till exponenter, som inte nödvändigtvis är positiva heltal. Det sker då med uttalat mål och syfte, att *potenslagarna skall fortfara att gälla.*

Låt oss börja med att lista ut, vad a^0 rimligen borde betyda. Kräver vi fortsatt giltighet av potenslagen $\frac{a^m}{a^n} = a^{m-n}$, så måste vi med nödvändighet ha

$$1 = \frac{a}{a} = \frac{a^1}{a^1} = a^{1-1} = a^0.$$

Vi tvingas alltså utvidga begreppet på följande vis.

Definition 3.3

För varje $a > 0$ definierar vi $a^0 = 1$.

Sats 3.4: Potenslagarna

Låt $a, b > 0$. För naturliga exponenter m och n gäller följande räknelagar.

- $a^m a^n = a^{m+n}$.
- $a^m / a^n = a^{m-n} \quad (m \geq n)$.
- $(a^m)^n = a^{mn}$.
- $a^m b^m = (ab)^m$.
- $a^m / b^m = (a/b)^m$.

Bevis

Vi argumenterar för den första lagen. Vi känner redan dess giltighet för positiva m och n. Om nu t.ex. $n = 0$, så säger lagen $a^m \cdot a^0 = a^{m+0}$, vilket stämmer, då vi ju definierade $a^0 = 1$. På samma sätt fungerar lagen om $m = 0$.

Låt oss angripa den tredje lagen. Om $m = 0$, så säger $(a^m)^n = a^{mn}$ att $(a^0)^n = a^{0\cdot n}$. Men enligt definitionen av a^0 är vänsterledet $1^n = 1$, liksom högerledet, så detta är sant. Om $n = 0$, så säger identiteten att $(a^m)^0 = a^{m\cdot 0}$, vilket också är sant, ty båda led har definierats till 1.

De andra lagarna visas med liknande argument.

§3. Heltaliga exponenter

Vi undersöker nu negativa exponenter. Antag, att det finns ett sätt, att tilldela a^n ett reellt värde för godtyckliga heltal n (både positiva och negativa), som för det första överensstämmer med definitionen när n är positiv eller noll, och för det andra uppfyller potenslagen $a^m \cdot a^n = a^{m+n}$. Låt oss leka med detta antagande. Vi ser, att

$$a^m a^{-m} = a^{m+(-m)} = a^0 = 1;$$

alltså måste vi i så fall ha $a^{-m} = \frac{1}{a^m}$. Då är t.ex.

$$a^{-1} = \frac{1}{a}, \qquad a^{-2} = \frac{1}{a^2}, \qquad \dots,$$

och alla potenser av a med negativ exponent är bestämda utav denna formel.

Detta visar, hurusom vi tvingas till följande definition för negativa exponenter, om vi vill behålla våra potenslagar.

Definition 3.5

Låt $a > 0$ och n vara ett positivt heltal. Då definierar vi

$$a^{-n} = \frac{1}{a^n} = \frac{1}{a \cdot \dots \cdot a} \qquad (n \text{ faktorer i nämnaren}).$$

Sats 3.6: Potenslagarna

Låt $a, b > 0$. För heltaliga exponenter m och n gäller följande räknelagar.

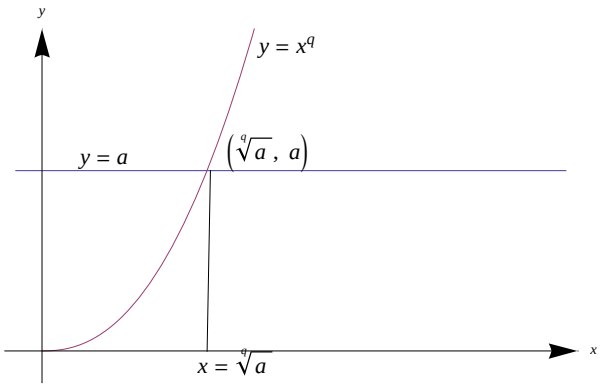
- $a^m a^n = a^{m+n}$.
- $a^m / a^n = a^{m-n}$.
- $(a^m)^n = a^{mn}$.
- $a^m b^m = (ab)^m$.
- $a^m / b^m = (a/b)^m$.

Bevis

Låt oss nöja oss med att visa den tredje lagen. Vi känner den redan för naturliga exponenter. Nu har vi ett antal andra fall att gå igenom, nämligen när m eller n är negativa. Antag t.ex. att $m = -p$, där p är positiv. Då är enligt definitionen $a^m = \frac{1}{a^p}$, så

$$(a^m)^n = \left(\frac{1}{a^p}\right)^n = \underbrace{\frac{1}{a^p} \cdot \dots \cdot \frac{1}{a^p}}_{n \text{ faktorer}} = \frac{1}{\underbrace{a^p \cdot \dots \cdot a^p}_{n \text{ faktorer}}} = \frac{1}{(a^p)^n} = \frac{1}{a^{pn}}$$

Men $mn = -pn$, så enligt definitionen är också $a^{mn} = \frac{1}{a^{pn}}$; således gäller $(a^m)^n = a^{mn}$. Nu har vi bara ytterligare två fall till att gå igenom, om vi inte fuskar och naivt accepterar att denna identitet, som bara har sisådär 300 års daglig användning på nacken, är sann



FIGUR 1: Grafen till $y = x^q$, då $q > 0$.

trots sin ungdom. Skämt åsido är det alltid en bra idé, att räkna igenom något fall till, för att öva sin skicklighet i att förstå och använda sig av definitioner, men det överlämnas åt läsaren.

§4. Rationella exponenter

Brutna tal som exponenter är väldigt användbara, och vi skall strax se varför. Låt oss mjukstarta med ett enkelt exempel i stil med det som motiverade övergången från positiva till negativa exponenter.

Vad borde menas med $a^{\frac{1}{2}}$, $a^{\frac{1}{3}}$ eller $a^{\frac{2}{3}}$, givet att dessa tal nu går att definiera meningsfullt? Nyckeln är, att potenslagarna bör fortsätta att gälla. Under denna förutsättning måste vi ha

$$a^{\frac{1}{2}} \cdot a^{\frac{1}{2}} = a^{\frac{1}{2}+\frac{1}{2}} = a^1 = a,$$

och alltså är $x = a^{\frac{1}{2}}$ en lösning till ekvationen $x^2 = a$. Men detta tal har redan ett namn! Den unika positiva lösningen till $x^2 = a$ går ju under namnet $x = \sqrt{a}$, varför $a^{\frac{1}{2}} = \sqrt{a}$ är den enda rimliga tolkningen ($-\sqrt{a}$ skulle förstås också fungera, men det verkar föga meningsfullt att tilldela $a^{\frac{1}{2}}$ ett negativt värde, då a ju är positiv).

På samma sätt är

$$(a^{\frac{1}{3}})^3 = a^{\frac{1}{3}\cdot 3} = a^1 = a,$$

vilket tvingar fram $a^{\frac{1}{3}} = \sqrt[3]{a}$. Slutligen måste

$$a^{\frac{2}{3}} = (a^{\frac{1}{3}})^2 = (\sqrt[3]{a})^2.$$

Vägledda av detta, definierar vi nu potenser med rationella exponenter.

Definition 3.7

Låt $a > 0$. För ett positivt heltal q, definierar vi $a^{\frac{1}{q}} = \sqrt[q]{a}$ som den unika positiva lösningen till ekvationen $x^q = a$. Ytterligare, då p är ett godtyckligt heltal, definierar vi

$$a^{\frac{p}{q}} = (a^{\frac{1}{q}})^p = (\sqrt[q]{a})^p.$$

För full förståelse är det lämpligt, att rita grafen till funktionen $y = x^q$ (Figur 1). Vad menas nämligen med "den unika positiva lösningen till $x^q = a$"? Vi startar alltså med ett tal a och vill hitta ett x så att $x^q = a$. Studera funktionsgrafen $y = x^q$. En lösning till ekvationen $x^q = a$ svarar i figuren mot (x-koordinaten för) en punkt, där grafen och den horisontella linjen $y = a$ skär varandra. Påståendet, att det finns en unik lösning, säger alltså, att det finns en och endast en skärningspunkt, vilket verkar troligt med figuren i beaktande.

En sladdrig figur är förstås inte nog. Strängt taget måste vi här ha ett lemma, som garanterar oss en unik lösning till $x^q = a$. Beviset kräver dock litet analys.

Lemma 3.8

Ekvationen $x^q = a$ har en unik positiv lösning, då $a > 0$ och q är ett positivt heltal.

Bevis

Funktionen $y = x^q$ är strängt växande för $0 < x < \infty$, ty $y' = qx^{q-1} > 0$. Eftersom $y(0) = 0$, men $y \to \infty$ då $x \to \infty$, följer det av Satsen om mellanliggande värden, att funktionen måste antaga värdet a i en unik punkt.

Med risk för att verka tjatiga (vi *är* tjatiga), så kommer här ännu en uppdaterad version av våra kära potenslagar. Vi avstår från beviset.

Sats 3.9: Potenslagarna

Låt $a, b > 0$. För rationella exponenter r och s gäller följande räknelagar.

- $a^r a^s = a^{r+s}$.
- $a^r / a^s = a^{r-s}$.
- $(a^r)^s = a^{rs}$.
- $a^r b^r = (ab)^r$.
- $a^r / b^r = (a/b)^r$.

Exempel 1.
Vi kan förenkla enligt potenslagarna

$$b^{\frac{1}{2}} / b^{\frac{1}{3}} = b^{\frac{1}{2} - \frac{1}{3}} = b^{\frac{1}{6}} = \sqrt[6]{b}.$$

Exempel 2.
För att förenkla uttryck som innehåller rotsymboler är det fördelaktigt, att ersätta dem med motsvarande potenser, för att sedan kunna använda potenslagarna:

$$\sqrt[7]{a^3 \sqrt[6]{a^3} \sqrt{a^7}} = \left(a^3 (a^3)^{\frac{1}{6}} (a^7)^{\frac{1}{2}} \right)^{\frac{1}{7}} = (a^3 a^{\frac{1}{2}} a^{\frac{7}{2}})^{\frac{1}{7}} = (a^7)^{\frac{1}{7}} = a.$$

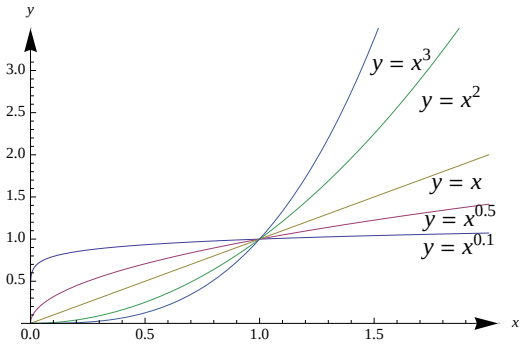
§5. Reella exponenter

Några ord slutligen, om hur man kan definiera a^x, där nu x tillåts vara ett godtyckligt reellt tal. Tag t.ex. 2^π. Hur klämmer vi åt den? Vi vet, att vi kan (i princip, kanske med en del jobbig räkning) approximera π (liksom alla reella tal) med ändliga decimalbråk med precis den noggrannhet, som vi någonsin skulle kunna önska, exempelvis 3.1415 (fyra decimaler) eller 3.1415 926 535 (tio decimaler).

Nu har vi redan genomfört definitionen av både

$$2^{3.1415} \qquad \text{och} \qquad 2^{3.1415\ 926\ 535}$$

i förra avsnittet. Exponenterna är ju rationella tal. Vi förväntar oss nog, att 2^π — om nu detta tal existerar — bör gränsa ganska nära till dessa tal, och att det högra torde vara en bättre approximation än det vänstra. Ju bättre närmevärden till π, desto bättre närmevärden för 2^π borde vi få. Läsaren (som möjligen läser endimensionell analys parallellt) anar ett gränsvärde



FIGUR 2: Potensfunktionen $y = x^r$ för positiv exponent r.

lura någonstans i kulissen. På sant matematiskt manér kan vi nu vända på resonemanget och *definiera* 2^π som gränsvärdet för nedanstående talföljd.

$$\begin{aligned}
2^{3.14} &= 8.815\,240\,92\ldots \\
2^{3.141\,5} &= 8.824\,411\,08\ldots \\
2^{3.141\,59} &= 8.824\,961\,59\ldots \\
2^{3.141\,592\,653\,5} &= 8.824\,977\,82\ldots \\
2^{3.141\,592\,653\,589\,793\,238\,462\,643\,383\,27} &= 8.824\,977\,82\ldots
\end{aligned}$$

De åtta första decimalerna i 2^π spirar fram och stabiliserar sig så sakteliga under våra ögon.

Definition 3.10

Låt $a > 0$. För ett godtyckligt reellt tal

$$x = c.d_1d_2d_3d_4\ldots$$

med heltalsdelen c och decimalerna $0 \leq d_i \leq 9$, definierar vi

$$a^x = \lim_{n\to\infty} a^{c.d_1d_2d_3\ldots d_n}.$$

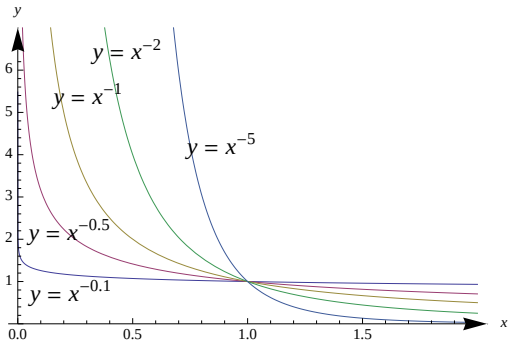
Återigen är det en hel del arbete för att bevisa, att definitionen fungerar (varför i all världen existerar gränsvärdet alltid?). Detta måste åtföljas av en massa träigt arbete för att visa potenslagarnas fortsatta giltighet. Potenslagarna är förstås inte träiga, utan *eleganta*, men att kontrollera deras giltighet är det! Vi formulerar dem inte en gång till; nu får det vara nog.

§6. Potensfunktionernas grafer

Några nätta små bilder kan månne liva upp den tunga teorien? I Figur 2 och 3 finns några piffiga grafer av potensfunktioner för att åskådliggöra deras beroende av exponenten. För positiva exponenter är potensfunktionen $y = x^r$ växande för $x \geq 0$. Ju större exponenten, desto snabbare växer den för $x \geq 1$. Beteendet mellan $0 \leq x \leq 1$ kan förtjäna särskild granskning. För att ge ordentliga bevis för dessa egenskaper krävs idéer från analysen.

För negativa exponenter är $y = x^r$ avtagande för $x \geq 0$. Ju mindre exponenten, desto snabbare avtar den för $x \geq 1$.

Alla kurvor $y = x^r$ passerar förstås punkten $(1, 1)$, både för positiv och negativ exponent.



FIGUR 3: Potensfunktionen $y = x^r$ för negativ exponent r.

ÖVNINGAR

3.1. Förenkla

(a) $\frac{4^3 \cdot 5^4}{10^6}$;

(b) $(\frac{1}{9})^{-\frac{1}{2}}$;

(c) $(\sqrt{81})^{\frac{3}{4}}$;

(d) $(\frac{1}{3})^{-2}$;

(e) 2^{2^3};

(f) $(2^2)^3$.

3.2. Förenkla

(a) $\frac{a^{-5.1}a^{1.9}}{a^{-3.2}}$;

(b) $\frac{a\sqrt[3]{a}}{\sqrt{a\sqrt{a}}}$;

(c) $\frac{\sqrt[4]{x^8y^7}}{\sqrt{x^4y^3}}$;

(d) $\sqrt{\sqrt[3]{x}}\sqrt[6]{x^{-1}}(x^2)^{\frac{2}{3}}$.

3.3. Förenkla

(a) $5^x + 5^{x+2}$;

(b) $e^x + e^{x+1}$;

(c) $(5^x + 5^{-x})^2 - 5^{2x} - 5^{-2x}$;

(d) $\frac{1}{5^x} + \frac{1}{5^{x+1}}$.

3.4. Förenkla

(a) $\sum_{k=1}^{n} x^k$;

(b) $\prod_{k=1}^{n} x^k$.

Kapitel 4

Ekvationer och olikheter

§1. Ekvationer

En ekvation har formen "något uttryck = något annat uttryck", där de två uttrycken beror av en eller flera obekanta. Att lösa ekvationen är förstås att hitta de explicita värden på de obekanta som löser ekvationen.

För att sammanfatta de viktigaste principerna vid ekvationslösning, så söker vi i första hand transformera den givna ekvationen till en enklare *ekvivalent* ekvation, d.v.s. en ekvation med precis samma lösningar. Sådana kan ofta fås genom att utföra en eller flera av följande operationer.

- Addera eller subtrahera samma ting från bägge sidor av ekvationen, typ $2x + 3 = 5$ är ekvivalent med $2x = 2$.
- Multiplicera eller dividera bägge sidor av ekvationen med en nollskild kvantitet, typ $xe^x = 2e^x$ är ekvivalent med $x = 2$ (eftersom alltid $e^x \neq 0$).
- Algebraiskt omformulera ena eller båda leden var för sig, typ ekvationen $1 = x^2 + 2x + 1$ är ekvivalent med $1 = (x + 1)^2$.

Här är det lämpligt att göra en utvikning om begreppen implikation och ekvivalens. *Implikationen* $\Rightarrow$ markerar att vi drar en slutsats, exempelvis

$$x \geq 2 \quad \Rightarrow \quad x^2 \geq 4.$$

Det kan översättas i ord till: "Om ett tal är större än eller lika med 2, så är dess kvadrat större än eller lika med 4." För att verifiera att det är sant, tittar vi på alla fall när $x \geq 2$ är sant, och räknar ut att $x^2 \geq 2^2 = 4$. Vi behöver förstås inte studera alla de fall när påståendet $x \geq 2$ är falskt; dem är vi inte intresserade av.

Däremot så är påståendet

$$x^2 = 4 \quad \Rightarrow \quad x = 2$$

falskt. Om vi vet att kvadraten på ett tal är 4, så behöver talet inte nödvändigtvis vara 2. Talet kunde ju förstås lika gärna vara -2. Korrekt implikation går åt motsatt håll:

$$x^2 = 4 \quad \Leftarrow \quad x = 2.$$

Vi kan också ge mer fullständig information genom en *ekvivalens*:

$$x^2 = 4 \quad \Leftrightarrow \quad (x = 2 \quad \text{eller} \quad x = -2).$$

De två påståendena eller ekvationerna är här ekvivalenta, vilket betyder, att de är sanna samtidigt och falska samtidigt. De har alltså precis samma lösningar. Inga lösningar har tappats längs vägen (vilket vore katastrof), och inga har heller tillkommit längs vägen, så kallade *falskrötter*. Dessa kan man dock inte alltid undvika.

§2. Förstagradsekvationer

De enklaste ekvationerna är *förstagradsekvationerna*, exempelvis

$$3x = 17 \qquad \text{och} \qquad 4x + 17 = 2x + 5.$$

Dessa kan, som läsaren vet, faktiskt alltid lösas. Detsamma gäller andragradsekvationerna, som strax följer.

Senare i kursen ämnar vi studera tredje- och fjärdegradsekvationer i samband med polynom. I den linjära algebran behandlas metoder att lösa system av förstagradsekvationer med flera obekanta. De otäcka ekvationer, man möter i tillämpningarna (fysik, kemi och sådana biämnen), kan däremot sällan lösas exakt, men det finns tekniker för att lösa ekvationer approximativt.

Exempel 1.
Som enkel fingerövning, låt oss tillsammans lösa ekvationen

$$4x + 17 = 2x + 5.$$

Vi använder principen att, om man multiplicerar med eller adderar lika storheter till bägge sidor av en likhet, så fortsätter likheten att gälla. Elimineras först termerna som innehåller x från ena sidan, så fås

$$(4x + 17) - 2x = (2x + 5) - 2x,$$

Förenklar vi denna ekvation får vi $2x + 17 = 5$. Nu subtraherar vi 17 från bägge sidorna och får

$$(2x + 17) - 17 = 5 - 17.$$

Denna förenklas till $2x = -12$, som löses genom division med 2 på bägge sidor. Då får vi till slut $x = -6$.

Detta var en lösning med mycket text, där varje steg förklarats noga i ord. Fullt så noggranna förklaringar krävs förstås inte på universitetet (vi förutsätter ändå, att förstagradsekvationens lösning är välbekant). En vanlig lösning läroverksstudenter presenterar kan däremot se ut ungefär som följer.

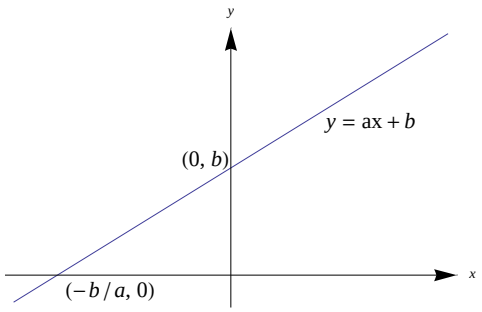
$$\begin{aligned} 4x + 17 &= 2x + 5 \\ 2x + 17 &= 5 \\ 2x &= -12 \\ x &= -6. \end{aligned}$$

Räkningarna är förstås klara, men detta ger ju en ganska ofullständig och tafflig bild av det matematiska resonemanget. Hur vet vi att det inte tillkommer falskrötter? Och än värre, hur vet vi att det inte förloras några lösningar längs vägen? Det syns inte av resonemanget. Räkningarna måste kompletteras, antingen med några förklaringar i ord som ovan eller med ekvivalenspilar:

$$\begin{aligned} & 4x + 17 = 2x + 5 \\ \Leftrightarrow\ & 2x + 17 = 5 \\ \Leftrightarrow\ & 2x = -12 \\ \Leftrightarrow\ & x = -6. \end{aligned}$$

Detta är en fullständig lösning värdig en universitetsstudent. Ekvivalenspilarna mellan till exempel $2x = -12$ och $x = -6$ anger, att dessa ekvationer är ekvivalenta, alltså att de har samma lösningar. Ett annat sätt att formulera det är, att om den ena gäller, så gäller även den andra, och omvänt. Vi kan gå både framlänges och baklänges i räkningarna, och är därför säkra på, att rötterna till den ena ekvationen också är rötter till den andra. Inga rötter har tappats längs vägen, och inga falskrötter kan ha tillkommit.

Sett från ett abstrakt perspektiv består lösningen av ekvationen i byggandet av en kedja av ekvivalenta ekvationer, d.v.s. ekvationer med precis samma lösningar, och där den sista ekvationen har den simpla formen $x =$ tal.



FIGUR 1: Grafen till ett förstagradspolynom.

Lösningen av förstagradsekvationen

$$ax + b = 0$$

kan grafiskt tolkas som att vi söker skärningen $x = -\frac{b}{a}$ mellan x-axeln och den räta linjen

$$y = ax + b.$$

Koefficienten a är riktningskoefficienten för linjen, medan b kan tolkas som y-koordinaten för skärningen mellan y-axeln och linjen.

Exempel 2.
Följande ekvation är väsentligen en förstagradsekvation:

$$\frac{4x-1}{3x+1} = 1.$$

Uttrycket på vänster sida är meningslöst, när nämnaren är noll. Detta sker för $x = -\frac{1}{3}$. Detta värde på x kan alltså inte vara en lösning. Låt oss sedan undersöka övriga tänkbara värden på x. Vi antar då $x \neq -\frac{1}{3}$, så att $3x + 1 \neq 0$. Då kan vi multiplicera bägge sidor av ekvationen med $3x + 1$ och få den ekvivalenta ekvationen $4x - 1 = 3x + 1$. Omedelbara manipulationer ger att denna ekvation har den enda lösningen $x = 2$.

Ett rent symboliskt sätt att redovisa detta är via logiska pilar:

$$\begin{aligned} & \frac{4x-1}{3x+1} = 1 \\ \Rightarrow \quad & 4x - 1 = 3x + 1 \\ \Leftrightarrow \quad & x = 2. \end{aligned}$$

Observera den enkelriktade implikationspilen. Den signalerar att räkningen inte nödvändigtvis kan omvändas. Resonemanget fungerar i ena riktningen, men inte i den andra. Detta beror på, att vi, för att gå bakåt, måste dividera med $3x+1$, och *detta tal behöver inte nödvändigtvis vara nollskilt*. Eftersom lösningen tydligen var $x = 2$, är dock divisionen tillåten och resonemanget omvändbart. Det visar, att $x = 2$ verkligen är en äkta rot. Det hade förstås också kunnat inses genom prövning i den ursprungliga ekvationen:

$$\frac{4 \cdot 2 - 1}{3 \cdot 2 + 1} = \frac{7}{7} = 1.$$

Vi är pedantiska, det medges. Men det har vi dessvärre fog för. Det var viktigt, att göra analysen av nämnaren ovan, vilket nästa exempel visar.

Exempel 3.
Låt oss lösa

$$\frac{6x+3}{2x+1} = 1$$

på ett lagom söndagszombieartat vis. Vi sätter igång att multiplicera med $2x+1$ och får ekvationen $6x+3 = 2x+1$ med lösningen $x = -\frac{1}{2}$. Succé! Hurra!

Nej, inte hurra. Vi skall inte inbilla oss, att detta är en rot till ekvationen. Vänsterledet är ju inte meningsfullt för $x = -\frac{1}{2}$, eftersom vi då försöker dela 0 med 0, vilket är rätt meningslöst. (Visserligen påstod en svensk morgontidning nyligen att en engelsman löst ”det tusenåriga problemet om vad $\frac{0}{0}$ egentligen är för något”, men det var nog journalistisk desperation och längtan efter att det äntligen skulle hända något en ovanligt mordfattig dag.) Den korrekta slutsatsen är, att ekvationen saknar lösningar.

En mer kortfattad lösning skulle kunna se ut som följer.

$$\begin{aligned} & \frac{6x+3}{2x+1} = 1 \\ \Rightarrow \quad & 6x+3 = 2x+1 \\ \Leftrightarrow \quad & x = -\frac{1}{2}. \end{aligned}$$

Prövning i den ursprungliga ekvationen visar dock, att $x = -\frac{1}{2}$ är en falskrot. Ekvationen saknar alltså lösning.

§3. Andragradsekvationer

Nu skall vi lösa andragradsekvationer.

Exempel 4.
En ekvation som $x^2 - 2 = 0$ har ju lösningarna $x = \pm\sqrt{2}$, och på samma sätt ser vi bums, att $x^2 + 2 = 0$ inte har några lösningar alls. Om vi sedan, med detta enkla fall i bakhuvudet, tittar på en ekvation som $x^2 + 2x - 1 = 0$, så ser vi kanske att vi kan skriva om den till

$$0 = x^2 + 2x - 1 = (x^2 + 2x + 1) - 2 = (x+1)^2 - 2.$$

Men då är

$$\begin{aligned} & (x+1)^2 - 2 = 0 \\ \Leftrightarrow \quad & (x+1)^2 = 2 \\ \Leftrightarrow \quad & x+1 = \pm\sqrt{2} \\ \Leftrightarrow \quad & x = -1 \pm \sqrt{2} \end{aligned}$$

och vi har hittat de två rötterna till ekvationen!

Kvadratkomplettering är ett sätt att försöka göra exemplets resonemang så allmänt som möjligt. På så vis kan vi härleda den ökända formeln för lösning av andragradsekvationer. Vi svenskar har givit den smeknamnet *pq-formeln*; våra danska bröder kallar den *ab-formeln* (tydligen är de vana vid andra beteckningar). Den danska kärleken till drastiska formuleringar har för övrigt döpt Trigonometriska ettan

$$\cos^2 x + \sin^2 x = 1$$

till *Idiotformeln*.

Sats 4.1: *pq*-formeln

Den reella andragradsekvationen

$$x^2 + px + q = 0$$

har de reella rötterna

$$x = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q},$$

förutsatt att uttrycket under rottecknet är icke-negativt.

Bevis

Tittar vi tillbaka på exemplet vill vi skriva om de två första termerna som en kvadrat. Från kvadreringsregeln får vi

$$\left(x + \frac{p}{2}\right)^2 = x^2 + 2 \cdot \frac{p}{2}x + \left(\frac{p}{2}\right)^2 = x^2 + px + \left(\frac{p}{2}\right)^2,$$

vilket vi kan skriva om som

$$x^2 + px = \left(x + \frac{p}{2}\right)^2 - \left(\frac{p}{2}\right)^2.$$

Alltså är

$$0 = x^2 + px + q = \left(x + \frac{p}{2}\right)^2 - \left(\frac{p}{2}\right)^2 + q \quad \Leftrightarrow \quad \left(x + \frac{p}{2}\right)^2 = \left(\frac{p}{2}\right)^2 - q.$$

Nu har vi äntligen kommit till en av den lätta sortens ekvationer. Då uttrycket $p^2 - 4q \geq 0$ finner vi de reella lösningarna genom att dra kvadratrötter:

$$\begin{aligned} & 0 = x^2 + px + q \\ \Leftrightarrow \quad & \left(x + \frac{p}{2}\right)^2 = \left(\frac{p}{2}\right)^2 - q \\ \Leftrightarrow \quad & x + \frac{p}{2} = \pm\sqrt{\left(\frac{p}{2}\right)^2 - q} \\ \Leftrightarrow \quad & x = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q} \end{aligned}$$

Notera ekvivalenspilarna hela vägen!

Lösningen kan även skrivas

$$x = \frac{-p \pm \sqrt{p^2 - 4q}}{2}.$$

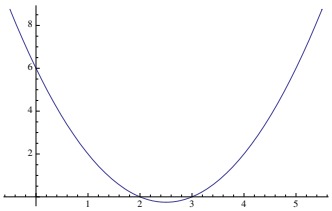
Uttrycket $p^2 - 4q$ är *diskriminanten* till polynomet $x^2 + px + q$, så benämnd, emedan den diskriminerar mellan de tre fallen: två reella rötter, en reell dubbelrot och två irreella rötter. Vi ser att $p^2 - 4q = 0$ ger en dubbelrot, medan $p^2 - 4q > 0$ ger två reella rötter.

Om diskriminanten $p^2 - 4q < 0$, saknas uppenbarligen reella lösningar, men samma räkningar och samma formel duger faktiskt för att uttrycka ekvationens irreella rötter. Vi har visserligen inte introducerat komplexa tal än, men exemplet nedan visar hur det fungerar.

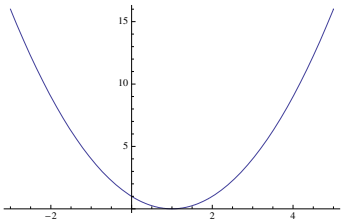
I de följande tre exemplen diskuterar vi tre olika andragradsekvationer $p(x) = 0$. Motsvarande grafer $y = p(x)$ illustrerar de tre olika fallen, som kan uppkomma: två, en eller noll reella rötter.

Exempel 5.
Ekvationen

$$x^2 - 5x + 6 = 0$$



FIGUR 2: Grafen till $y = x^2 - 5x + 6$.



FIGUR 3: Grafen till $y = x^2 - 2x + 1$.

löses med pq-formeln till

$$x = \frac{5}{2} \pm \sqrt{\left(\frac{5}{2}\right)^2 - 6} = \frac{5}{2} \pm \sqrt{\frac{1}{4}} = \frac{5}{2} \pm \frac{1}{2},$$

så att rötterna är $x = 2$ eller $x = 3$, vilka syns i Figur 2 som skärningspunkterna mellan grafen till $y = x^2 - 5x + 6$ och x-axeln.

Exempel 6.
Ekvationen

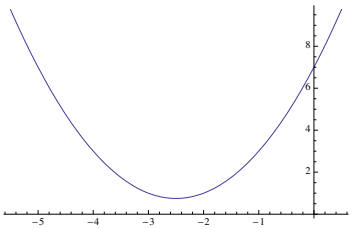
$$x^2 - 2x + 1 = 0$$

har dubbelroten $x = 1$, och vi ser att grafen i Figur 3 mycket riktigt tangerar x-axeln i denna punkt.

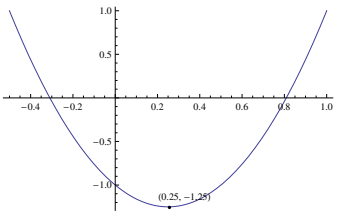
Exempel 7.
Slutligen visas i Figur 4 grafen för $y = x^2 + 5x + 7$, för vilken pq-formeln ger rötterna

$$x^2 + 5x + 7 = 0 \quad \Leftrightarrow \quad x = -\frac{5}{2} \pm \sqrt{\left(\frac{5}{2}\right)^2 - 7} = -\frac{5}{2} \pm \sqrt{-\frac{3}{4}}.$$

Normalt kan vi inte dra roten ur negativa tal, men i just detta fall fungerar faktiskt den naiva tolkningen $x = -\frac{5}{2} \pm \frac{\sqrt{3}}{2}i$. Rötterna är irreella, och grafen skär ej x-axeln.



FIGUR 4: Grafen till $y = x^2 + 5x + 7$.



FIGUR 5: Grafen till $y = 4x^2 - 2x - 1$.

Beviset för *pq*-formeln ger en bra illustration till hur bevis ofta innehåller idéer, som är användbara i andra sammanhang. Här är det den algebraiska teknik, som bär namnet *kvadratkomplettering*.

Exempel 8.
Enklaste sättet att bestämma det minsta värde, som funktionen

$$f(x) = 4x^2 - 2x - 1$$

kan antaga, är att kvadratkomplettera. Skriv

$$f(x) = 4x^2 - 2x - 1 = 4\left(x^2 - \frac{1}{2}x - \frac{1}{4}\right) = 4\left(\left(x - \frac{1}{4}\right)^2 - \frac{1}{16} - \frac{1}{4}\right)$$
$$= 4\left(\left(x - \frac{1}{4}\right)^2 - \frac{5}{16}\right) = 4\left(x - \frac{1}{4}\right)^2 - \frac{5}{4}.$$

Kvadrater är alltid större än eller lika med 0, så därför är

$$\left(x - \frac{1}{4}\right)^2 \geq 0,$$

med likhet precis när $x = \frac{1}{4}$. Alltså är

$$f(x) = 4\left(x - \frac{1}{4}\right)^2 - \frac{5}{4} \geq -\frac{5}{4},$$

med likhet precis när $x = \frac{1}{4}$. Se grafen i Figur 5.

§4. Rotekvationer

Vi skall studera tredje- och fjärdegradsekvationer längre fram. Just nu lämnar vi emellertid det temat till förmån för rotekvationer.

Exempel 9.
Vi önskar lösa ekvationen

$$\sqrt{x+1} + x = 11. \tag{1}$$

Här är det naturligt att först subtrahera x från bägge sidor, så att kvadratroten blir ensam kvar på ena sidan och vi kan kvadrera bägge sidor. Alltså får vi först

$$\sqrt{x+1} = 11 - x. \tag{2}$$

Detta är en ekvation som har precis samma lösningar som den ursprungliga, alltså en ekvivalent ekvation. Men om vi nu kvadrerar bägge sidor, ledande till

$$x + 1 = (11 - x)^2, \tag{3}$$

så har vi inte längre en ekvation som är ekvivalent med den ursprungliga. Varje lösning till (2) är en lösning till (3), men inte tvärtom. Att vi vet ett tals kvadrat räcker ju inte, för att vi skall kunna bestämma talet ifråga. Det är bara bestämt upp till tecken: $y = 2$ ger att $y^2 = 4$, men inte tvärtom, eftersom $y^2 = 4$ löses av både $y = \pm 2$. Från och med nu måste vi därför spana efter falskrötter, men det är inget att böla för. Det är bara att lösa

$$x + 1 = (11 - x)^2 \quad \Leftrightarrow \quad x^2 - 23x + 120 = 0$$

(med pq-formeln eller kvadratkomplettering), vilket ger kandidaterna $x = 8$ eller $x = 15$, och sedan pröva lösningarna i den ursprungliga ekvationen (1). Talet $x = 8$ är en äkta rot, ty

$$\sqrt{x+1} + x = \sqrt{8+1} + 8 = 11.$$

Men för $x = 15$ får vi

$$\sqrt{x+1} + x = \sqrt{15+1} + 15 = 19 \neq 11.$$

Detta är alltså en falsk rot till den ursprungliga ekvationen (1).

En kortfattad symbolisk lösning ges av följande.

$$\begin{aligned} & \sqrt{x+1} + x = 11 \\ \Leftrightarrow \quad & \sqrt{x+1} = 11 - x \\ \Rightarrow \quad & x + 1 = (11 - x)^2 = 121 - 22x + x^2 \\ \Leftrightarrow \quad & x^2 - 23x + 120 = 0 \\ \Leftrightarrow \quad & x = 8 \quad \text{eller} \quad x = 15. \end{aligned}$$

Prövning av rötterna som ovan utvisar falskheten hos $x = 15$, medan $x = 8$ är ekvationens enda äkta rot.

Det är instruktivt att utröna exakt var, hur och varför den eländiga falskroten gör sin sordida entré. Vi utmärkte platsen för det hemska brottet med en enkelriktad implikationspil i räkningen ovan. Notera, att *båda* ekvationerna

$$\sqrt{x+1} = 11 - x \qquad \text{och} \qquad \sqrt{x+1} = -11 + x$$

kvadreras till $(x+1)^2 = (11-x)^2$. Vår äkta rot $x = 8$ löser den vänstra av dessa, och därmed den kvadrerade ekvationen. Den falska roten $x = 15$ löser inte den vänstra, men väl den högra, och kommer därför också att lösa den kvadrerade ekvationen. Härifrån kommer den alltså.

Vi kan också ge en repris av räkningarna för det numeriska värdet $x = 15$ och finner

$$\begin{aligned} & \sqrt{15+1} + 15 = 19 \neq 11 \\ \Leftrightarrow \quad & \sqrt{15+1} = 4 \neq -4 = 11 - 15 \\ \Rightarrow \quad & 15 + 1 = (11 - 15)^2 = 16 = 121 - 22 \cdot 15 + 15^2. \end{aligned}$$

Vid implikationspilen transformeras den falska identiteten $4 \neq -4$ till en sann genom kvadreringen.

§5. Substitution

Så några exempel på hur variabelsubstitution kan vara behjälplig.

Exempel 10.
Substitutionen $t = x^2$ transformerar fjärdegradsekvationen

$$x^4 - 3x^2 + 2 = 0$$

till andragradsekvationen $t^2 - 3t + 2 = 0$, med rötterna $t = 1$ och $t = 2$. Nu har $x^2 = t = 1$ rötterna $x = \pm 1$ och $x^2 = t = 2$ rötterna $x = \pm\sqrt{2}$. Ekvationens fyra rötter är därmed ± 1 och

$\pm\sqrt{2}$.

Exempel 11.
Vi önskar lösa ekvationen

$$\sin^2 x - 4\sin x + 3 = 0.$$

Här är det naturligt att dela upp lösningen i två steg. Först tar vi reda på vilka värden $\sin x$ kan ha, och sedan vilka värden på x som är lösningar. Sätt alltså först $t = \sin x$. Då blir ekvationen $t^2 - 4t + 3 = 0$, som (enligt den vanliga formeln) har lösningarna $t = 1$ och $t = 3$. I nästa steg skall vi lösa ekvationen $\sin x = 1$, som har rötterna $x = \frac{\pi}{2} + 2n\pi$, där n är ett godtyckligt heltal, och ekvationen $\sin x = 3$, som saknar lösningar.

§6. Olikheter löses som ekvationer... nästan

En olikhet som

$$x + 2 \leq 5,$$

där problemet är att bestämma de reella tal x, för vilka olikheten är sann, kan vi lösa genom att subtrahera 2 från bägge sidor. Vi får $(x + 2) - 2 \leq 5 - 2$, alltså $x \leq 3$. Detta är Axiom O5, additiv isotoni, och alltså exakt samma teknik som för linjära ekvationer.

På liknande vis kan vi lösa $2x \geq 3$ genom att multiplicera med $\frac{1}{2}$ på bägge sidor:

$$2x \geq 3 \quad \Leftrightarrow \quad 2x \geq 3 \quad \Leftrightarrow \quad x \geq \frac{3}{2}.$$

Mer generellt är principen att

$$a \leq b \quad \Leftrightarrow \quad ac \leq bc, \qquad \text{om } c > 0.$$

Detta är ena delen av Sats 1.18 i kapitel 1. Men här kommer nu en viktig skillnad gentemot ekvationslösning. Sats 1.18 säger också, att multiplikation med ett negativt tal vänder olikheten:

$$a \leq b \quad \Leftrightarrow \quad ac \geq bc, \qquad \text{om } c < 0.$$

Kontentan är att, om vi bara håller reda på tecknet på allt vi multiplicerar (eller dividerar med), och byter riktning på våra olikheter, när det är negativt, så kan vi lösa olikheter precis som ekvationer.

Exempel 12.
Vi löser olikheten $3x + 7 < -5x + 15$ som

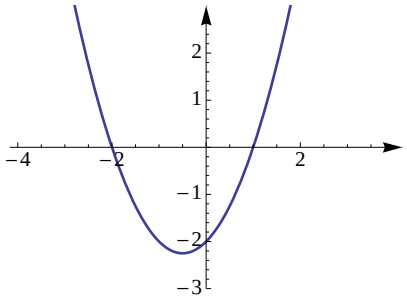
$$\begin{aligned} & 3x + 7 < -5x + 15 \\ \Leftrightarrow \quad & 8x + 7 < 15 \\ \Leftrightarrow \quad & 8x < 8 \\ \Leftrightarrow \quad & x < 1. \end{aligned}$$

§7. Teckenstudium

Exempel 13.
Antag att vi vill veta när funktionen $f(x) = (x-1)(x+2)$ är positiv, d.v.s. vi vill lösa olikheten $(x - 1)(x + 2) \geq 0$. Vi vet ju (Sats 1.19 i kapitel 1), att produkten av två positiva, liksom två negativa, tal är positiv, och att produkten av ett positivt och ett negativt tal är negativ. Om vi alltså vet tecknet på respektive faktor i $(x - 1)(x + 2)$ kan vi lösa problemet.

x				-2			1			
$x+2$	$-$	$-$	$-$	0	$+$	$+$	$+$	$+$	$+$	$+$
$x-1$	$-$	$-$	$-$	$-$	$-$	$-$	0	$+$	$+$	$+$
$f(x)=(x-1)(x+2)$	$+$	$+$	$+$	0	$-$	$-$	0	$+$	$+$	$+$

TABELL 1: Teckenstudium för $f(x) = (x-1)(x+2)$.



FIGUR 6: Grafen till $y = (x+2)(x-1)$.

Men detta är lätt att utröna. Faktorn $x-1$ är 0 när $x = 1$, positiv när $x > 1$, samt negativ när $x < 1$. Den andra faktorn $x+2$ är 0 när $x = -2$, positiv när $x > -2$, samt negativ när $x < -2$. Vi kan alltså dela upp tallinjen i tre intervall $x < -2$, $-2 < x < 1$ och $1 < x$. I varje intervall vet vi tecknet på faktorerna och därmed på funktionen $f(x)$ i varje intervall. I brytpunkterna är funktionen noll.

I Tabell 1 är den översta raden en tallinje, med $x = -2$ och $x = 1$ markerade. Varje rad under tallinjen hör till en funktion, och under varje punkt på tallinjen (nåja, inte *alla* punkter...) står tecknet för denna funktion. Ur de två första radernas tecken vid ett visst x kan vi få tecknet för produkten $f(x)$ enligt principen, att samma tecken ger positiv produkt, olika tecken ger negativ produkt.

Från tabellen kan vi avläsa, att $f(x) \geq 0$ precis då $x \leq -2$ eller $x \geq 1$. Vi ser samma resultat i grafen till $y = f(x)$ i Figur 6, som ligger ovanför x-axeln precis när funktionen är positiv.

Samma teknik fungerar, då det är fråga om division eller fler än två faktorer.

Exempel 14.
Vi försöker lösa olikheten

$$f(x) = \frac{(x-1)(x+1)e^x}{(x-3)(x+5)} > 0$$

med samma metod. Vi har nu fem faktorer. De fyra linjära faktorerna $x+5$, $x+1$, $x-1$ och $x-3$ växlar tecken i $x = -5$, $x = -1$, $x = 1$ respektive $x = 3$. Den femte e^x är alltid positiv och spelar därför ingen roll för resultatet. (Vi kan dela med den och få en ekvivalent olikhet.)

Tabell 2 ger teckenstudium för denna funktion. Principen för att få den sista raden är i detta fall, att produkten av ett jämnt antal negativa faktorer är positiv, medan ett udda antal negativa faktorer ger en negativ produkt. Vid $x = -5$ och $x = 3$ är funktionen inte definierad, något som vi markerat med ett frågetecken.

Ur tabellen kan vi slutligen avläsa, att $f(x) > 0$ precis då $x < -5$, $-1 < x < 1$ eller $x > 3$.

Exempel 15.
När är

$$\frac{1}{(x-2)(x-4)} \geq \frac{1}{(x-1)(x-2)} \; ?$$

Vi kan lösa detta genom att skriva om olikheten, så att det blir en fråga om teckenstudium. Olikheten är ekvivalent med att $r(x) \geq 0$, där

x			-5				-1		1		3			
$x+5$	$-$	$-$	0	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
$x+1$	$-$	$-$	$-$	$-$	$-$	$-$	0	$+$	$+$	$+$	$+$	$+$	$+$	$+$
$x-1$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	0	$+$	$+$	$+$	$+$	$+$
$x-3$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	0	$+$	$+$	$+$
$f(x)$	$+$	$+$	$?$	$-$	$-$	$-$	0	$+$	0	$-$	$?$	$+$	$+$	$+$

TABELL 2: Teckenstudium för $f(x) = \frac{(x-1)(x+1)e^x}{(x-3)(x+5)}$.

x				1		2				4			
$x-1$	$-$	$-$	$-$	0	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
$x-2$	$-$	$-$	$-$	$-$	$-$	0	$+$	$+$	$+$	$+$	$+$	$+$	$+$
$x-4$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	0	$+$	$+$	$+$
$r(x)$	$-$	$-$	$-$	$?$	$+$	$?$	$-$	$-$	$-$	$?$	$+$	$+$	$+$

TABELL 3: Teckenstudium för $r(x) = \frac{3}{(x-1)(x-2)(x-4)}$.

$$r(x) = \frac{1}{(x-2)(x-4)} - \frac{1}{(x-1)(x-2)}$$
$$= \frac{x-1}{(x-1)(x-2)(x-4)} - \frac{x-4}{(x-1)(x-2)(x-4)} = \frac{3}{(x-1)(x-2)(x-4)}.$$

Teckenstudium finns i Tabell 3. Ur denna avläser vi svaret $1 < x < 2$ eller $x > 4$. (För $x = 1, 2, 4$ är nämnaren noll och kvoten odefinierad, så dessa punkter skall definitivt inte ingå i svaret.) Vi ser samma resultat i grafen $y = r(x)$, som ligger ovanför x-axeln precis när funktionen är positiv (Figur 7).

Varning. I detta exempel skulle man också lätt, men *felaktigt*, kunnat "tänka" så här: Multiplicera bägge sidor i

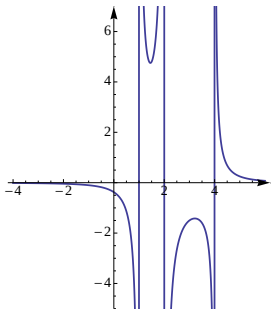
$$\frac{1}{(x-2)(x-4)} \geq \frac{1}{(x-1)(x-2)}$$

med $(x-2)(x-1)(x-4)$, så blir vi av med bråkskräpet. Det ger

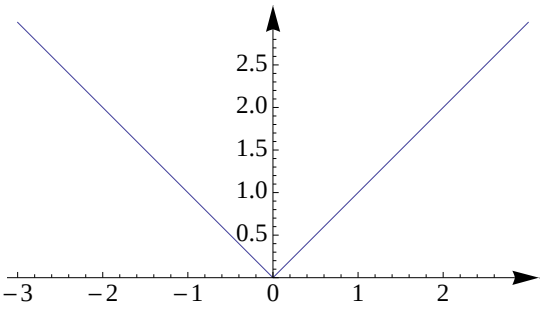
$$x - 1 \geq x - 4 \quad \Leftrightarrow \quad 3 \geq 0,$$

vilket alltid är sant. Man skulle därför förledas att tro olikheten alltid vara sann. Problemet med detta resonemang är, att tecknet på det vi multiplicerar med *varierar med* x. Det är ibland negativt, och i dessa fall skulle alltså olikheten behöva vändas.

Ibland kan man förstås gå rakt på:



FIGUR 7: Grafen till $y = \frac{3}{(x-1)(x-2)(x-4)}$.



FIGUR 8: Grafen till $y = |x|$.

Exempel 16.
När är

$$\frac{x-2}{x^2+4} \geq 0 \ ?$$

En kvadrat kan aldrig vara negativ, så $x^2 + 4 \geq 4 > 0$. Alltså kan vi multiplicera med $x^2 + 4$ och få den ekvivalenta olikheten $x - 2 \geq 0$, med lösningen $x \geq 2$.

§8. Absolutbelopp

Läsaren har säkert redan mött begreppet absolutbelopp. Absolutbeloppet $|a|$ mäter avståndet från talet a till 0 längs reella talaxeln. T.ex. är $|3| = 3$ och, mer generellt, om $a \geq 0$ så är $|a| = a$. För negativa tal som -3 har vi $|-3| = 3$ eller $|-5| = 5$. Det är alltså lätt att beskriva, vad det innebär att taga absolutbeloppet av ett reellt tal, som vi möter ansikte mot ansikte: vi plockar bort ett eventuellt minustecken. Hur skall vi formulera detta strikt matematiskt?

Vi tittar på fallet då a är negativt igen. Då är $a = -b$, där b är positivt. Avståndet från a till origo är just b, så vi har att $|a| = b$. I termer av a är $b = -a$, så här har vi faktiskt klistrat på ett minustecken: $|a| = b = -a$. Alltså är t.ex. $|-3| = -(-3) = 3$.

Definition 4.2

Om a är ett reellt tal, så definieras dess **absolutbelopp** av

$$|a| = \begin{cases} a & \text{om } a \geq 0 \\ -a & \text{om } a \leq 0. \end{cases}$$

Talet $a = 0$ täcks av båda fallen i definitionen, men de ger upphov till samma värde $|0| = 0 = -0$.

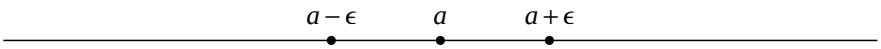
I Figur 8 har vi ritat upp grafen $y = |x|$ till absolutbeloppet som funktion av x. Observera att funktionsvärdena ju är längder av sträckor, och därmed aldrig negativa. Därför ligger grafen i övre planhalvan. Observera vidare att $|x| = |-x|$, vilket svarar mot grafens symmetri kring y-axeln.

Skepnaden av absolutbeloppets definition, med fallindelning, kan verka främmande för den, som är svag för funktioner med stadiga och rejäla formler likt $x^2(e^x + 23)$, men är egentligen inte konstigare, än att det kostar olika för icke-studenter och studenter att åka buss. Först avgör ju biljettförsäljaren om x är student eller ej, varefter biljettpriset $B(x)$ bestäms med detta rön som grundval. Även här härskar en falluppdelning.

För övrigt går det visst att producera en explicit formel. Den ser ut så här:

$$|a| = \sqrt{a^2}.$$

Vi kan se detta direkt från Definition 4.2. Kvadratroten ur ett positivt tal b är alltid den *positiva* lösningen till $x^2 = b$. Ekvationen $x^2 = a^2$ har de två rötterna $\pm a$, och kvadratroten är alltså den



Figur 9: Intervallet $|x-a| \leq \epsilon$.

av dessa som är positiv. Men detsamma är ju sant för $|a|$. Detta är nyttigt att ha som reflex: $|a|$ *är det av de två talen* $\pm a$*, som är positivt.*

Exempel 17.
Om vi skall förenkla

$$\frac{\sqrt{(x-1)^2}}{x-1},$$

är det frestande att draga kvadratroten ur kvadraten, $\sqrt{(x-1)^2} = x-1$, och få kvoten till 1. Men detta är fel, ty $\sqrt{(x-1)^2} = x-1$ är bara sant, om $x-1 \geq 0$. Om istället $x-1 < 0$ så är $\sqrt{(x-1)^2} = -(x-1)$, och kvoten blir -1. Det korrekta sättet att förenkla är alltså att skriva

$$\frac{\sqrt{(x-1)^2}}{x-1} = \frac{|x-1|}{x-1} = \begin{cases} \frac{x-1}{x-1} = 1 & \text{om } x > 1 \\ -\frac{x-1}{x-1} = -1 & \text{om } x < 1. \end{cases}$$

(Uttrycket är odefinierat om $x = 1$.)

Absolutbeloppet $|a-b|$ mäter avståndet mellan talen a och b på reella tallinjen. Detta är en viktig och användbar intuition och förklarar, till en del, varför absolutbelopp så ofta dyker upp. Avståndet mellan t.ex. -2 och 3 är $|3-(-2)| = 5$. Det kvittar i vilken ordning vi skriver dem: $|a-b| = |b-a|$.

I Figur 9 ser vi, att $|x-a| = \epsilon$ inträffar precis när avståndet från x till a är ϵ, d.v.s. precis för $x = a+\epsilon$ eller $x = a-\epsilon$. Då har vi förflyttat oss från a med avståndet ϵ antingen åt höger eller vänster. Intervallet $a-\epsilon \leq x \leq a+\epsilon$ mellan dessa två tal består av de punkter, vars avstånd till a är mindre än eller lika med ϵ, d.v.s. de punkter som uppfyller $|x-a| \leq \epsilon$. Absolutbeloppet erbjuder därmed ett behändigt sätt att beskriva intervall.

Exempel 18.
Talet 3.14 är ett närmevärde med tre signifikanta siffror till π. Att de tre siffrorna är signifikanta innebär, att $|\pi - 3.14| \leq 0.5 \cdot 10^{-2} = 0.005$, eller att $3.135 \leq \pi \leq 3.145$, d.v.s. att π ligger instängt i ett visst intervall, symmetriskt kring 3.14.

§9. Ekvationer med absolutbelopp

Falluppdelning är väsentlig vid lösning av problem innehållande absolutbelopp.

Exempel 19.
Vi vill lösa ekvationen

$$|x-3| + 2x = 0.$$

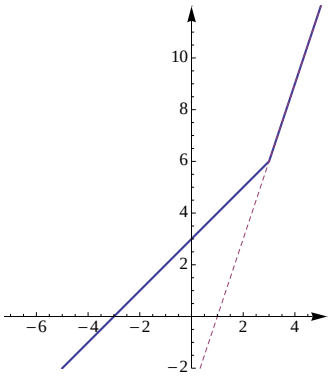
För att kunna räkna på detta, vill vi gärna bli av med absolutbeloppstecknen. Det kan vi göra, om vi vet tecknet på $x-3$, eftersom det följer ur definitionen att

$$|x-3| = \begin{cases} x-3 & \text{om } x-3 \geq 0 \\ -(x-3) & \text{om } x-3 \leq 0. \end{cases}$$

Den punkt, där vi växlar rad i definitionen, är $x = 3$. Första raden i definitionen används när $x \geq 3$, andra raden när $x \leq 3$. För att eliminera de hemska absolutbeloppen, bör vi alltså dela upp lösningen i två fall.

Fall 1: $x \geq 3$. I detta intervall är

$$0 = |x-3| + 2x = (x-3) + 2x = 3x - 3,$$



FIGUR 10: Grafen till $y = |x-3| + 2x$ och $y = 3x - 3$ (streckad).

med enda lösningen $x = 1$. Det är bara den lilla haken, att vi just nu letar lösningar i intervallet $x \geq 3$, och häri ligger inte talet 1. Alltså är detta inte en giltig lösning, utan en falskrot! (Vad händer, om vi försöker med $x = 1$ i den ursprungliga ekvationen?) Vi drar slutsatsen, att ekvationen inte har några lösningar $x \geq 3$.

Fall 2: $x \leq 3$. I detta intervall är

$$0 = |x-3| + 2x = -(x-3) + 2x = x + 3,$$

med enda lösningen $x = -3$. Denna lösning uppfyller verkligen villkoret $x \leq 3$, och är därmed en äkta lösning till den ursprungliga ekvationen.

Vi kan se detta fenomen på grafen till $y = |x-3| + 2x$ i Figur 10. Vi letar nollställen till funktionen, d.v.s. punkter där grafen skär x-axeln. Till vänster om $x = 3$ sammanfaller grafen med linjen $y = x + 3$, och de har det gemensamma nollstället $x = -3$. Till höger om $x = 3$ sammanfaller $y = |x-3| + 2x$ i stället med linjen $y = 3x - 3$, som syns streckad i bilden. Den streckade linjen skär x-axeln i $x = 1$, som inte ligger på kurvan $y = |x-3| + 2x$.

Exempel 20.
Som ett litet mer involverat exempel, löser vi ekvationen

$$|x-1| + |x+1| = 3.$$

För att kunna tillämpa våra kunskaper om linjära ekvationer på detta vill vi, som i förra exemplet, gärna bli av med absolutbeloppstecknen. Det kan vi göra, om vi vet tecknet på *både* $x-1$ och $x+1$. Brytpunkten när vi växlar rad i definitionen av $|x-1|$ är $x = 1$, och för $|x+1|$ är den $x = -1$. Vi kan alltså dela upp lösningen i tre fall, som svarar mot de tre intervall som dessa två punkter delar in reella talaxeln i. I vart och ett av intervallen kan vi ersätta absolutbeloppen med ett enda linjärt uttryck.

Fall 1: $x \geq 1$. Då är både $x+1$ och $x-1$ positiva, så att $|x+1| = x+1$ och $|x-1| = x-1$. Alltså reduceras ekvationen till

$$3 = |x-1| + |x+1| = (x-1) + (x+1) = 2x,$$

med lösningen $x = \frac{3}{2}$. Eftersom $3/2 \geq 1$, så är detta alltså en giltig lösning.

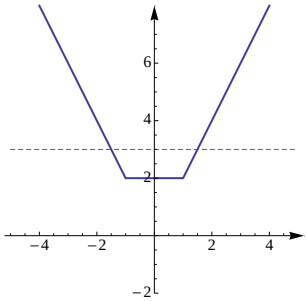
Fall 2: $-1 \leq x \leq 1$. Då är $x - 1$ negativ och $|x-1| = -(x-1)$, medan $x + 1$ fortfarande är positiv och alltså $|x+1| = x+1$. Alltså reduceras ekvationen till

$$3 = |x-1| + |x+1| = -(x-1) + (x+1) = 2.$$

Likheten $2 = 3$ är förstås nonsens; hur vi än väljer x, kan vi aldrig få $2 = 3$. Det blir inga lösningar i detta fall.

Fall 3: $x \leq -1$. Då är både $x - 1$ och $x + 1$ negativa, så att $|x-1| = -(x-1)$ och $|x+1| = -(x+1)$. Därmed reduceras ekvationen till

$$3 = |x-1| + |x+1| = -(x-1) - (x+1) = -2x,$$



FIGUR 11: Grafen till $y = |x-1| + |x+1|$ och $y = 3$.

med lösningen $x = -\frac{3}{2}$. Denna lösning är giltig, ty den är mindre än -1.

Ekvationen har alltså precis två lösningar, $x = \pm\frac{3}{2}$.

Vi kan studera kurvan $y = |x-1| + |x+1|$ i Figur 11 för att se hur detta tar sig uttryck. Denna gång är vi ute efter x-koordinater för skärningen med den horisontella linjen $y = 3$, streckad i figuren. Vi fann två värden $x = \pm\frac{3}{2}$, som också är uppenbara i figuren.

Vi föreställer oss nu, att linjen $y = 3$ förskjutes uppåt eller nedåt, d.v.s. vi studerar linjen $y = b$ för varierande b. När $b = 3$ har vi två skärningspunkter med grafen, d.v.s. två lösningar till $|x-1| + |x+1| = b$. Om b sedan ökar, kommer vi fortfarande att ha två lösningar, liksom också i början när b minskar. Men när så $b = 2$ dyker det plötsligt upp oändligt många lösningar, vilka sedan åter försvinner när $b < 2$. Ekvationen har inga lösningar alls för dessa värden på b. Detta illustrerar olika fall som kan förekomma vid studium av absolutbeloppsekvationer.

ÖVNINGAR

4.1. Kvadratkomplettera polynomet $p(x) = 2x^2 - x - 5$, lös ekvationen $p(x) = 0$, samt bestäm det minsta värde som polynomet kan antaga.

4.2. Lös ekvationerna

(a) $\sqrt{2x+3} = x$;

(b) $\sqrt{2x+3} = -x$;

(c) $\sqrt{x+2}\sqrt{x-1} = x$.

4.3. Finn alla reella lösningar till ekvationen

$$\sqrt{x} + \sqrt{x+1} = \frac{3}{2}.$$

4.4. Finn alla reella x, som uppfyller ekvationerna

(a) $x^4 - x^2 - 2 = 0$;

(b) $e^x + \frac{1}{e^x} = 2$;

(c) $\sin^8 x = 1$.

4.5. Lös olikheterna

(a) $3x + 1 > 5x + 2$;

(b) $(x-3)(x+3) \leq x^2$;

(c) $(x+1)(x-2) \leq x^2 + 2x$.

4.6. Lös olikheterna

(a) $x^2 < 4$;

(b) $x^2 > 4$;

(c) $(x+1)^2 > (x+5)^2$.

4.7. Visa, att olikheten

$$x^2 + xy + y^2 \geq 0$$

gäller för alla reella tal x och y. När råder likhet?

4.8. Lös olikheten

$$\frac{3x+1}{x+2} < 2.$$

4.9. Lös olikheterna

(a) $\frac{x^2+1}{x} < x$;

(b) $\frac{2x^2}{x+2} < x - 2$;

(c) $\frac{x^2+2}{x^2+1} > 1$.

4.10. Lös olikheten

$$\frac{3x+1}{x+2} < \frac{2}{x+3}.$$

4.11. Lös olikheten

$$x(x-1)(x-2)(x-3)(x-4)(x-5)(x-6) > 0.$$

4.12. Bestäm (a är ett reellt tal)

(a) $|5|$;

(b) $|-5|$;

(c) $\sqrt{5^2}$;

(d) $\sqrt{(-5)^2}$;

(e) $\sqrt{a^2}$;

(f) $\sqrt{(-a)^2}$.

4.13. Finn alla reella x, som uppfyller

(a) $|x| = 3$;

(b) $|x| = 0$;

(c) $|x| = -3$;

(d) $|x-1| = 3$;

(e) $|2x-1| = 3$;

(f) $|1-x| = 3$.

4.14. Rita mängden av de reella tal x, för vilka

(a) $|x| \leq 3$;

(b) $|x| \geq 2$;

(c) $|x-1| \leq 5$;

(d) $|x+2| < 3$.

4.15. Lös ekvationerna

(a) $|x| - x = 2$;

(b) $x^2 + 2|x| - 3 = 0$;

(c) $x^2 + 2|x+1| - 1 = 0$.

4.16. Finn de reella x, som uppfyller ekvationen $|x+1|+|x-1|=4$.

4.17. Utred för vilka reella konstanter a ekvationen

$$\sqrt{x+a}-\sqrt{x-a}=1$$

har reella lösningar, och bestäm dessa i förekommande fall.

4.18. Visa, att uttrycket

$$\sqrt{x-4\sqrt{x-1}+3}+\sqrt{x-6\sqrt{x-1}+8}$$

är konstant för $5 \leq x \leq 10$.

4.19. Talen a, b och c uppfyller olikheterna

$$|a| \geq |b+c|, \quad |b| \geq |c+a| \quad \text{och} \quad |c| \geq |a+b|.$$

Bevisa att $a+b+c=0$.

Del II

Talteori

Kapitel 5

Heltal

I naturvetarens verktygslåda utgör olika sorters tal förstås en stor del. Historiskt har man, allt eftersom nya problem dykt upp, hittat på nya typer. Allra först kom förstås de positiva hela talen $\mathbb{Z}^+ = \{1, 2, 3, \dots\}$, som användes för att beräkna *antal*. Hur många kossor gick ut att beta om morgonen? Hur många kossor kom tillbaka på kvällen? (Att det möjligen inte var samma kossor spelade kanske mindre roll.)

Betraktat t.ex. som militär teknologi var det säkert en revolution — för första gången fick man en billig metod att avgöra om man kunde plundra grannbyn — man *räknade* helt enkelt hur många de var och *jämförde* med hur många man själv var och hade så en enkel tumregel för om man skulle vinna eller inte. (Som så många matematiska modeller innebar detta förstås en grov men samtidigt effektiv förenkling av verkligheten.)

Efter många hundra år införde man talet 0 (kanske som en symbol för en walk-over situation när grannbyn mangrant tagit semester) och fick så $\mathbb{N} = \{0, 1, 2, 3, \dots\}$, mängden av hela tal större än eller lika med 0. Bråktal och negativa tal är utvidgningar av detta vinnande koncept, och så småningom har man också hittat på reella tal (oändliga decimaltal) och rentav komplexa tal (och det finns fler varianter).

Det finns flera poänger med att inte bums kila vidare till, säg, reella tal, funktioner och analys, utan först se vad man kan göra enbart med hela tal. En poäng är, att bygga upp en känsla för hur matematisk teoribyggnad ser ut och hur komplicerade saker den kan leda till.

En annan och kanske oväntad sådan poäng är teoriens aktualitet och användbarhet. Primtal är väsentliga i vissa viktiga tillämpningar av matematik från de sista decennierna. De är nämligen basen för de flesta moderna krypterings- och kodningsmetoder, och det som följer i denna text (och mycket mer) används sålunda varje gång man stoppar in sitt kontokort i en bankomat. Dessa tillämpningar är rätt nya, men baserade på resultat som varit kända i mer än 2000 år. Grunderna för talteorien står redan i Euklides *Elementa*, skriven cirka 300 f.Kr. Detta illustrerar att, även om en matematisk teori kan tyckas vara ren vetenskap, enbart intressant ur ett vetenskapligt perspektiv och i sin egen rätt, så hindrar det inte dess användbarhet. Det visar också, att matematiken aldrig blir färdig, inte på långa vägar när.

§1. Primtal och sammansatta tal

Låt oss titta på, hur man kan konstruera positiva hela tal. Från 2 och 3 kan vi med addition bilda t.ex. $5 = 2+3$, $6 = 3+3$, och så vidare. I själva verket kan vi redan med 1 bilda $5 = 1+1+1+1+1$. Och vi inser förstås, att med 1 som den enda typen av legobit (men i obegränsad upplaga) kan vi få alla positiva heltal med upprepad addition. Och för att få de negativa talen sedan, så sätter vi bara ett minustecken framför. Inte speciellt fantasieggande...

Låt oss nu i stället försöka konstruera de positiva hela talen, men med hjälp av multiplikation i stället för addition. Från 2 och 3 kan vi då med multiplikation bilda t.ex.

$$\begin{array}{lll} 2 & 4 = 2 \cdot 2 & 8 = 2 \cdot 2 \cdot 2 \\ 3 & 6 = 2 \cdot 3 & 12 = 2 \cdot 2 \cdot 3 \quad \dots \end{array}$$

Några ögonblicks eftertanke får oss att inse, att vi inte kan få t.ex. 5 som en produkt av 2 och 3. Och situationen blir inte bättre av, att vi lägger till 1 till legobitarna. För att kunna skriva

alla heltal som produkter måste vi alltså ha tillgång till åtminstone 2 och 3 och 5. Men även då får vi inte med 7. Så då måste vi ha 7 också. Men då kan vi fortfarande inte skriva 11 som en produkt av våra legotal.

Situationen är alltså helt annorlunda och mycket mer komplicerad än den var med addition. När världen känns krånglig och svår finns det inget (utom kanske kaffe med något *starkt* i och en deckare) som är så lugnande som en definition:

Definition 5.1

Ett positivt heltal $n > 1$ kallas **sammansatt**, om det är en produkt $n = ab$ av två positiva heltal $a, b > 1$. (Av detta följer nödvändigtvis, att $1 < a < n$ och $1 < b < n$.) Om $n > 1$ inte är sammansatt, kallas det för ett **primtal**.

Talet 1 är, enligt definitionen, varken primtal eller sammansatt, utan faller i en egen kategori. Denna konvention gör (så småningom) teorien lättare att hantera.

Exempel 1.
Talet 6 är ett sammansatt heltal ty $6 = 2 \cdot 3$. Uppenbarligen är 2 ett primtal, ty vi kan uppenbarligen inte skriva $2 = ab$ för två positiva heltal $a, b > 1$. Talen 3, 5, 7 och 11, som dök upp i resonemanget ovan, är också primtal.

Definition 5.2

Ett heltal b kallas **delare** till heltalet n, om det finns ett heltal a, så att $n = ab$. Vi säger också att n **delas av** b. Vi skriver detta som $b \mid n$.

Man ser lätt att $-b \mid n$ om och endast om $b \mid n$.

Exempel 2.
Observera, att definitionen inte bara täcker positiva heltal, utan alla heltal. T.ex. delas 6 av -2, eftersom $6 = (-2) \cdot (-3)$, och alltså gäller det att $-2 \mid 6$. En fullständig lista på alla delare av 6 är $\pm 1, \pm 2, \pm 3, \pm 6$.

Exempel 3.
Vilket tal n som helst delas alltid av sig själv och av 1, eftersom $n = n \cdot 1$. Det positiva talet $n > 1$ är alltså ett primtal om och endast om dess enda positiva delare är n och 1.

Exempel 4.
Varje tal b är en delare i 0 (ty $0 = 0 \cdot b$), medan 0 själv aldrig kan dela ett nollskilt tal n (ty $a \cdot 0 = 0 \neq n$).

§2. Primtalsfaktorisering

Framdeles skall vi mest syssla med positiva tal. Primtalen är de tal, som vi fann ovan svara mot ett slags legobitar för att bygga upp tal med multiplikation. Argumentet för att varje positivt heltal kan skrivas som en produkt av primtal är lätt att ge, och det skall vi göra nu. Först skall vi se det i ett exempel, med ett konkret heltal, och sedan ge ett allmänt bevis, som täcker alla positiva heltal. I matematiska texter finns oftast bara det allmänna argumentet för ett påstående, och då är det ett mycket effektivt sätt att förstå argumentet, att själv gå igenom det i ett speciellt fall.

Exempel 5.
Tag ett tal, t.ex. 3102. Precis som alla tal större än 1 är det antingen ett primtal eller ett sammansatt tal. Man ser att

$$3102 = 2 \cdot 1551,$$

så vårt tal är alltså sammansatt.

Talet 2 konstaterade vi ovan vara ett primtal, men vi kan fortsätta med 1551. Precis som alla tal större än 1, är det antingen primtal eller sammansatt. Om det är sammansatt, så måste det vara delbart med något tal mindre än 1551. Låt oss kontrollera dessa. Talet är inte delbart med 2, men däremot med 3, ty $1551 = 3 \cdot 517$. Hittills har vi fått faktoriseringen

$$3102 = 2 \cdot 3 \cdot 517,$$

Talet 3 är ett primtal, men vi kan fortsätta med 517. Det är inte delbart med 2 och därför inte heller med något jämnt tal överhuvudtaget. Vidare är det inte delbart med 3, ty $\frac{517}{3} = 172.333\ldots$, ej heller med 5 eller 7, ty $\frac{517}{5} = 103.4$ och $\frac{517}{7} = 73.8571\ldots$. Vi behöver inte kontrollera om 517 är delbart med 9 eller 10, för gällde det t.ex. att $517 = 9 \cdot a$ för något a, så vore ju också $517 = 3 \cdot 3a$, så att 517 då vore delbart med 3. Emellertid är 517 delbart med 11, ty $517 = 11 \cdot 47$.

För att sammanfatta, har vi alltså hittills fått

$$3102 = 2 \cdot 3 \cdot 11 \cdot 47.$$

Genom att pröva att dela 47 med alla tal mindre än 47, ser man strax, att 47 är ett primtal. Ovan har vi alltså skapat en framställning av 3102, som produkten av primtal.

När ett positivt heltal $n > 1$, som 3102 i exemplet, skrivits som en produkt $n = p_1 p_2 \cdots p_r$ av primtal, säger vi att talet *primtalsfaktoriserats*. Observera, att r får vara 1, så att ett primtal är sin egen primtalsfaktorisering.

Talet i exemplet är förstås bara ett av oändligt många tal. Vi kan pröva fler tal och se, att argumentet i exemplet väsentligen fungerar på samma sätt. Sedan är vi nog rätt övertygade om, att varje tal torde kunna primtalsfaktoriseras. Men vi *vet* inte att det är sant, förrän vi givit ett strängt bevis, d.v.s. ett korrekt och logiskt argument, som visar att satsen är sann för *alla* heltal. Som tur är, kan vi analysera exemplet ovan och inse, varför samma teknik kommer att fungera för vilket heltal (större än 1) som helst.

Sats 5.3

Varje heltal större än 1 kan primtalsfaktoriseras, alltså skrivas som produkten av ett eller flera primtal. (En alternativ formulering är: För varje heltal $n \geq 2$ finns det primtal $p_1, \ldots, p_r$, sådana att $n = p_1 p_2 \cdots p_r$.)

Bevis

Även det här beviset formuleras rätt pladdrigt. Poängen är, att satsen, ur en viss synvinkel sett, är ett rätt starkt påstående. Det uttalar sig nämligen om existensen av något, som vi rent konkret inte har en isbits chans i helvetet att hitta. Redan för tal med bara tusen futtiga siffror, kommer det att krävas en dator stor som själva universum och mer tid än dess återstående livslängd för att hitta deras primtalsfaktorisering.

Lyckligtvis behöver vi inte hitta konkreta faktoriseringar för alla tal. Vi behöver bara visa på en principiell metod, och den kan vi hitta i exemplet. Det startade med, att vi hade ett tal $n > 1$, och så konstaterade vi att, precis som alla positiva heltal större än 1, är det antingen primtal eller sammansatt tal. Notera nu att, om n är ett primtal, är vi klara: satsen är sann för detta tal. (Ett primtal är sin egen primtalsfaktorisering.) Annars, om alltså n är sammansatt, finns det tal $a > 1$ och $b > 1$, sådana att $n = ab$. Hade vi nu vetat att både a och b vore produkter av primtal, skulle vi vara färdiga. I förstone kan detta tyckas vara ett cirkelbevis. Men vad vet vi om a och b? Nå, åtminstone är de strikt mindre än n.

Så låt oss lägga upp beviset som ett tankeexperiment. Låt oss tänka oss, hurusom vi går igenom heltalen i storleksordning, startande med 2. Om satsen inte är sann för alla tal, så kommer vi förr eller senare till ett första motexempel, alltså ett tal n, som inte är en produkt av primtal. Eftersom n inte är ett primtal, finns det heltal $a, b > 1$, sådana att

$n = ab$. Men vi visste ju, att alla tal strikt mindre än n är produkter av primtal. Speciellt gäller detta för a och b. Men då är förstås också deras produkt $n = ab$ en produkt av primtal. Detta strider ju klart mot tanken, att vårt n skulle vara det minsta motexemplet. Alltså kan det inte finnas sådana n; satsen är således sann.

En metod för att finna primtalsfaktoriseringar får vi på köpet. Givet ett tal n, kontrollerar vi för alla tal b, med $1 < b < n$, om $b \mid n$. Hittar vi inga delare, är n ett primtal, och om vi hittar en faktorisering $n = ab$, så tillämpar vi samma procedur på a och b.

Exempel 6.
Man kan lätt visa att, om n är ett positivt tal, sådant att inga heltal a med $1 < a \leq \sqrt{n}$ delar n, så är n ett primtal. (Se övningarna.) Detta kan användas för att med huvudräkning visa, att 113 är ett primtal. Och för att kontrollera, om 8521 är ett primtal (med en miniräknare), behöver vi högst genomföra 30 divisioner, om vi kan multiplikationstabellen (och det kan vi väl?) och därmed vet, vilka tal under 100 som är primtal.

Exempel 7.
Följande sätt att konstruera primtal kallas *Eratosthenes såll*. Det har den fördelen, att man inte ens behöver kunna räkna ut kvoter av heltal, och är egentligen alldeles densamma tanke, som vi startade det här kapitlet med: primtalen är de legobitar, som behövs för att konstruera alla tal med multiplikation.

Skriv upp alla tal från 2 till (säg) 30 i ordningsföljd:

	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30

Idén är, att eliminera alla sammansatta tal genom att systematiskt ta bort alla multipler av primtal.

Det första talet 2 i listan måste vara ett primtal — det är ju inte delbart med något tal, som är mindre än sig själv. De tal, som är delbara med 2, är precis vartannat tal i listan: $2, 4, 6, 8, \ldots$. Stryk dem allihop utom det första talet 2:

	2	3	~~4~~	5	~~6~~	7	~~8~~	9	~~10~~
11	~~12~~	13	~~14~~	15	~~16~~	17	~~18~~	19	~~20~~
21	~~22~~	23	~~24~~	25	~~26~~	27	~~28~~	29	~~30~~

Nästa tal, som nu står ostruket i listan, alltså talet 3, måste vara ett primtal — det försvann ju inte, då vi strök multiplerna av 2. De tal, som är delbara med 3, är precis vart tredje tal i listan: $3, 6, 9, 12, \ldots$. Stryk dem allihop utom det första talet 3:

	2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
~~21~~	~~22~~	23	~~24~~	25	~~26~~	~~27~~	~~28~~	29	~~30~~

Nu är 5 det första talet, som står ostruket i listan, och måste därför vara ett primtal — det är ju inte inte en multipel av varken 2 eller 3 (om det vore delbart med ett av de strukna talen, t.ex. 4, skulle det ju även vara delbart med 2). Stryk nu vart femte tal från 5 utom 5 själv (bara 25 blir nu aktuellt):

	2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
~~21~~	~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	29	~~30~~

Fortsätt på samma sätt. Nästa ostrukna tal 7 är ett primtal; stryk alltså vart sjunde tal efter 7 själv, o.s.v., o.s.v. De ostrukna tal som slutligen står kvar är precis primtalen. I vårt lilla exempel blir några fler strykningar inte aktuella, och vi har alltså funnit samtliga primtal upp till 30 ovan. (Från föregående exempel vet vi redan, att det räcker att kontrollera delbarhet med primtalen upp till $\sqrt{30} < 7$, alltså 2, 3 och 5.)

§3. Primtalsfaktoriseringens entydighet

Satsen, att varje heltal kan skrivas som en produkt av primtal, ger upphov till den naturliga frågan: går det att göra detta på olika sätt? Kan det t.ex. vara sant att

$$8521 \cdot 577 = 8513 \cdot 599 \ ?$$

(Alla ingående tal är primtal.) I detta specifika fall kan vi naturligtvis kontrollera att det är falskt, genom att räkna ut produkterna till $4\ 916\ 617 \neq 5\ 099\ 287$, men det finns kanske andra sätt att faktorisera 4 916 617 på?

Redan 2000 år före räknedosans uppfinnande visste de invigda, att primtalsfaktorisering bara kan ske på (i princip) ett enda sätt. Det finns förstås alltid möjligheten, att ändra på ordningen mellan faktorerna. T.ex. så är

$$30 = 2 \cdot 3 \cdot 5 = 2 \cdot 5 \cdot 3 = 3 \cdot 2 \cdot 5 = \cdots$$

(Hur många möjligheter finns det? Se kapitlet om kombinatorik!) Men dessa omordningar är ju bara rent kosmetiska, och vi säger därför, att alla dessa faktoriseringar som lika *upp till ordningen av faktorerna*.

Sats 5.4: Aritmetikens fundamentalsats

Varje heltal $n \geq 2$ kan primtalsfaktoriseras på ett och endast ett sätt, upp till ordningen av faktorerna.

(Alternativt kan det formuleras som: Antag $n \geq 2$ och att $n = p_1 p_2 \cdots p_r = q_1 q_2 \cdots q_s$, där $p_1, \ldots, p_r$ och $q_1, \ldots, q_s$ är primtal. Då är $s = r$ och $p_1, \ldots, p_r$ och $q_1, \ldots, q_r$ är samma primtal, men eventuellt i annan ordning.)

Vi visade ovan existensen av primtalsfaktoriseringar. Beviset för entydigheten är mer komplicerat och får anstå till senare. Men från denna vet vi, *utan att ens snegla på 2:ans multiplikationstabell*, att $3 \cdot 5 \neq 2 \cdot 2 \cdot 2 \cdot 2$, bara på ren intellektuell styrka, för att inte tala om att $8521 \cdot 577 \neq 8513 \cdot 599$. I sanning en sats värd att tatuera på insidan av ögonlocken, så att den, även när man sover, ständigt är framför ens ögon... (Att tatuera in kurslitteratur räknas förresten *inte* som fusk, så vitt vi kan se, då vi genomögnar universitetets regler. Kursledaren kan säkert tipsa om lämpliga satser att lämna till tatueraren.)

Vissa av primtalen i faktoriseringen av ett tal kan förstås vara lika. T.ex. är

$$48 = 2 \cdot 2 \cdot 2 \cdot 2 \cdot 3 = 2^4 \cdot 3.$$

Antalet faktorer som är lika med ett visst primtal p kallar vi **multipliciteten** för p. Alltså är 2 en primtalsfaktor i 48 med multiplicitet 4, medan 3 har multiplicitet 1.

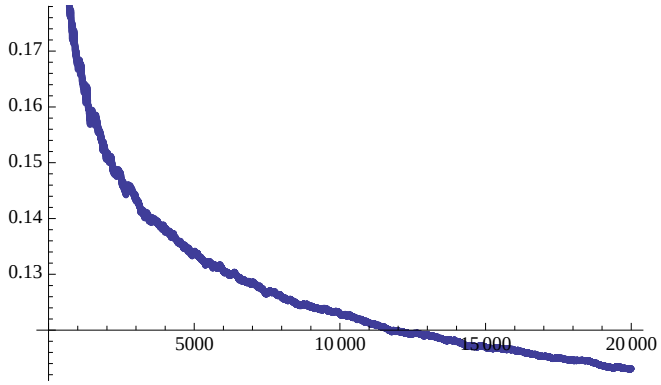
Hur gör vi förresten med negativa tal? Då kan vi förstås skriva t.ex. $-6 = (-1) \cdot 2 \cdot 3$ och får på det sättet en motsvarande faktorisering med en faktor -1 och ett antal positiva primtal. Men -1 är inte ett primtal för det, utan räknas, precis som talet 1, varken till primtalen eller de sammansatta talen.

§4. Hur många primtal finns det?

Kanske vore det praktiskt om det bara funnes några få primtal. Alla tal skulle då vara produkter av dessa. Men så är det inte.

Sats 5.5

Det finns oändligt många primtal.



FIGUR 1: Sannolikheten att ett tal mindre än n är ett primtal.

Bevis

Vi har två ömsesidigt uteslutande möjligheter.

- Det finns oändligt många primtal.
- Det finns inte oändligt många primtal, d.v.s. bara ett ändligt antal.

Vi skall visa att det andra alternativet inte kan inträffa genom att härleda en motsägelse ur det. Därmed måste det första alternativet vara det som gäller.

Antag att $p_1, p_2, \ldots, p_r$ är listan av *alla* primtal som finns och betrakta

$$x = p_1 p_2 \cdots p_r + 1.$$

(Vi kanske inte kan räkna ut x, men vi vet åtminstone att detta heltal finns.) Vi påstår nu, att primtalen $p_1, p_2, \ldots, p_r$ ej kan ingå i primtalsfaktoriseringen av x. För om primtalet p_i vore en faktor i x, skulle det dela både x och produkten $p_1 p_2 \cdots p_r$, och därmed också deras differens, som är 1. Men inget tal större än 1 kan förstås dela 1.

Men x har förstås någon primtalsfaktorisering (enligt Aritmetikens fundamentalsats), och denna måste då bestå av andra primtal än just de angivna $p_1, p_2, \ldots, p_r$. Dessa kan alltså inte vara alla primtal som finns, utan det måste finnas fler. Detta är den sökta motsägelsen. Enligt ovan måste det då vara det första alternativet som är sant.

§5. Fler satser och problem om primtal

Satsen ovan ger det mest grundläggande resultatet om primtalens antal. Man kan säga mycket mer. Om vi tittar på följden av primtal

$$2,\ 3,\ 5,\ 7,\ 11,\ 13,\ 17,\ 19,\ 23,\ 29,\ 31,\ 37,\ 41,\ 43,\ 47,\ 53,\ 61,\ \ldots,$$

ser vi att avståndet mellan primtalen successivt tycks öka. De kommer glesare och glesare. Av de första hundra talen är 25 primtal. Sannolikheten, att ett slumpvis valt tal bland dessa är ett primtal, är alltså $\frac{25}{100} = 25\%$. Av de första tusen talen är 168 primtal, varför sannolikheten, att ett slumpvis valt tal bland dessa skall vara ett primtal, är lägre, $\frac{168}{1000} = 16.8\%$. Ritar man en tabell och låter sin dator jobba kan man se hur sannolikheten sjunker, ju större tal vi betraktar. Se Figur 1.

Trots att kurvan är så skakig, vilket svarar mot svårigheten att beskriva primtalens exakta fördelning bland heltalen, verkar sannolikheten ändå i stort bete sig rätt regelbundet. Man kan visa, med stor ansträngning och mycket teknik, att kurvan ungefär följer grafen till $\frac{1}{\ln n}$. Detta är *Primtalssatsen*:

Sats 5.6: Primtalssatsen; Hadamard & De la Vallée-Poussin 1896

Sannolikheten, att ett slumpvis valt tal mellan 1 och n skall vara primtal, är ungefär[a]

$$\frac{1}{\ln n}.$$

[a]Vi avstår från att precisera den specifika innebörden hos ordet *ungefär*.

Det finns gott om andra svåra satser och än fler olösta problem om primtal. Par av primtal, vilka skiljer sig åt med 2, som t.ex. $3, 5$ eller $17, 19$ kallas för **primtalstvillingar**. Ingen vet om det finns oändligt många sådana tvillingar, fast man arbetat flera hundra år på att försöka avgöra detta. I april 2013 skedde ett mycket oväntat genombrott, då Zhang demonstrerade existensen av oändligt många par av primtal, som visserligen inte (som man så hett åtrår) bara skiljer sig åt med 2, men i alla fall med högst 70 000 000. Gränsen har sedan successivt pressats ned, tack vare världens alla talteoretikers gemensamma ansträngningar, och verkar för närvarande befinna sig på 246. Fortfarande en bra bit från 2 alltså, och nu verkar dessutom gränsen vara nådd för denna metods räckvidd.

Ett annat mycket celebert olöst problem är *Goldbachs förmodan* från 1742, som säger, att varje jämnt tal större än 2 kan skrivas som summan av två primtal.

Teorien för primtal har sedan länge varit intressant bland matematiker för sin egen skull, men även bland amatörer. För något tiotal år sedan åkte t.o.m. en primtalsentusiast, anställd hos ett telebolag i USA, i fängelse sedan han (olovligt) programmerat datorerna till att leta efter stora primtal på ledig tid. Men oturligt nog var det en bug i hans program och datorerna nöjde sig inte med "ledig tid". Efter ett tag togs allt mer av det multinationella bolagets totala datorkapacitet i anspråk till att leta efter sällsynta primtal, med följd att det tog allt längre och längre tid för allt surare kunder att få sina samtal kopplade...

Ett av de sju stora miljondollarpris, som utlysts av Clay Mathematics Institute, rör *Riemann-hypotesen*. Denna är motiverad av problemet, att få ett mer precist grepp om primtalens fördelning. Trots att det förstås, ursprungligen, handlar om heltal, kretsar hypotesen kring egenskaper hos de komplexa nollställena till en viss komplexvärd funktion, *Riemanns zeta-funktion*. Om dessa ligger utplacerade så, som Riemann för 150 år sedan förmodade, så får man nämligen mycket bättre uppskattningar än Primtalssatsen ovan för sannolikheten, att ett givet heltal skall vara primtal.

§6. Divisionsalgoritmen

Vad är division, som vi får lära oss det i skolan? Det är förstås en konkret metod att räkna ut kvoter. För tillfället lever vi dock i en värld befolkad av heltal allenast, vilket betyder att vi tvingas tänka om rörande begreppet division.

Vi är väl exempelvis vana vid att tänka oss kvoten

$$\frac{22}{7} = 3.1428\ldots.$$

Inom talteorien föredrar vi dock, att skriva

$$22 = 3 \cdot 7 + 1,$$

där *(heltals)kvoten* är 3 och *resten* är 1. Vi finner dessa med trappan eller liggande stolen, vad som nu föredras.

Sats 5.7: Divisionsalgoritmen

Till varje heltal n och varje positivt heltal b finns en unik **kvot** k och en unik **rest** r, med egenskaperna att

$$n = kb + r \qquad \text{och} \qquad 0 \leq r < b.$$

Bevis

Tänk (och rita gärna) en tallinje med heltalspunkterna angivna. Ett godtyckligt tal n är antingen en heltalsmultipel kb av b, så att $n = kb + 0$, eller så ligger n mellan två sådana heltalsmultipler kb och $(k+1)b$. Heltalspunkterna mellan dessa två är

$$kb + 1,\ kb + 2,\ \ldots,\ kb + (b-1),$$

och n måste då vara något av dessa tal. I bägge fallen är $n = kb + r$ och $0 \leq r < b$. (Är beviset oklart så tänk igenom det för $b = 5$.)

Vad säger satsen i några speciella exempel?

Exempel 8.
Tag t.ex. $b = 2$. Divisionsalgoritmen säger, att varje tal n kan skrivas som $k \cdot 2 + r$, där r är ett heltal med $0 \leq r < 2$. Det betyder att r är 0 eller 1. Alltså kan varje tal n skrivas som $2k$ eller som $2k + 1$. Vi känner igen uppdelningen av tal i jämna och udda.

Generellt säger villkoret $0 \leq r < b$, att resten r är något av talen 0, 1, ..., $b - 1$.

Exempel 9.
För $b = 5$ säger satsen, att varje tal n kan skrivas på en och endast en av formerna

$$5k,\ 5k + 1,\ 5k + 2,\ 5k + 3,\ 5k + 4.$$

Observera, att vi kan dividera även negativa tal.

Exempel 10.
Division av -12 med 5 ger $-12 = (-3) \cdot 5 + 3$, med kvoten -3 och resten 3. Kom ihåg, att resten, här 3, skall vara något av talen $0, 1, \ldots, 4$. I termer av beviset ovan, har vi först markerat alla, positiva såväl som negativa, multipler av 5 på tallinjen, och noterar sedan, att -12 ligger 3 steg till höger om -15. Jämför detta med att $12 = 2 \cdot 5 + 2$, där resten är 2.

§7. Euklides algoritm

Först ett motiverande problem. Varje rationellt tal kan skrivas på flera sätt som bråk. T.ex. är

$$\frac{1}{2} = \frac{3}{6} = \frac{112}{224} = \frac{k}{2k}$$

för vilket heltal $k \neq 0$ som helst. Givet ett bråk som $\frac{60}{48}$ kan vi förstås förkorta det:

$$\frac{60}{48} = \frac{4 \cdot 15}{4 \cdot 12} = \frac{15}{12} = \frac{3 \cdot 5}{3 \cdot 4} = \frac{5}{4}.$$

Här har vi hela tiden hittat *gemensamma delare* till täljaren och nämnaren och förkortat bort dessa, ända tills vi inte längre kan fortsätta. Vi säger att $\frac{5}{4}$ är slutförkortat.

Men kunde vi kanske hitta andra sätt att faktorisera och komma fram till ett annat slutresultat? Vi skyndar oss att försäkra läsaren, att så inte är fallet. Världen är god, i alla fall dess bråk, och det finns bara en slutförkortad form av ett bråk. Det följer av Aritmetikens fundamentalsats.

Problemet vi önskar studera i detta avsnitt är: *Hur hittar vi gemensamma delare och förkortar ett bråk maximalt?* Det kan synas som ett banalt problem (hur många bråk är vi närmare bekanta med?), men lösningen involverar en teknik som är fundamental och användbar i många andra sammanhang.

Nu skall vi först ge en definition. När vi förkortade $\frac{60}{48}$ ovan, så fann vi, att 4 och 3 var delare till både 60 och 48. Sådana tal kallas *gemensamma delare*. Vi är särskilt intresserade av den största gemensamma delaren.

Definition 5.8

Låt a och b vara två heltal, ej båda 0. **Största gemensamma delaren** $\text{SGD}(a, b)$ är det största heltal, som delar både a och b.

Exempel 11.
Ur primtalsfaktoriseringarna

$$30 = 2 \cdot 3 \cdot 5 \qquad \text{och} \qquad 24 = 2 \cdot 2 \cdot 2 \cdot 3$$

avläser vi de gemensamma delarna för 30 och 24 till ± 1, ± 2, ± 3 och ± 6, så att $\text{SGD}(30, 24) = 6$.

I föregående exempel ser vi med vilken lätthet vi finner största gemensamma delaren ur de ingående talens primtalsfaktoriseringar, eftersom vi då direkt kan se vilka primtal, som är gemensamma. Men primtalsfaktoriseringar är oftast svåra att finna, och det finns en direkt och mycket snabbare metod för beräkning av största gemensamma delaren, kallad *Euklides algoritm*.

Vi visar den först i ett exempel. För att se varför metoden fungerar, behövs ett lemma[1].

Lemma 5.9

Låt a och b vara två heltal. Om $d \mid a$ och $d \mid b$, så gäller att $d \mid (ax + by)$ för godtyckliga heltal x och y.

Bevis

Villkoren $d \mid a$ och $d \mid b$ betyder, enligt definitionen av delbarhet, att det finns heltal k och l, sådana att $a = kd$ och $b = ld$. Då är

$$ax + by = kdx + ldy = (kx + ly)d,$$

så att det finns ett tal c (nämligen $c = kx + ly$) med $ax + by = cd$. Det innebär, enligt definitionen, att $d \mid (ax + by)$.

Lemmat säger således, bland annat, att d delar $a + b$, $a - b$, $a + 2b$, $3a + b$. Dessa svarar ju mot olika värden på x och y. Sådana tal, på formen $ax + by$, kallas för **linjärkombinationer** av a och b. Lemmat utsäger alltså, att om d delar två tal, så delar d också varje linjärkombination av dem.

Exempel 12.
Vi ger en beskrivning av Euklides algoritm genom ett exempel. Vi önskar beräkna $\text{SGD}(216, 66)$. Dividera först det större talet med det mindre (trappan eller liggande stolen):

$$216 = 3 \cdot 66 + 18.$$

Kvoten är 3 och resten är 18. Notera nu två saker.

- Resten $18 = 216 - 3 \cdot 66$ är en linjärkombination av 216 och 66. Varje delare till 216 och 66 är därför också en delare i 18, enligt lemmat.
- Å andra sidan är talet 216 en linjärkombination av 66 och 18, så att varje delare till 66 och 18 också är en delare i 216.

De gemensamma delarna till 216 och 66 är alltså precis de gemensamma delarna till 66 och 18. Speciellt har de båda talparen samma *största* gemensamma delare, alltså

$$\text{SGD}(216, 66) = \text{SGD}(66, 18).$$

Följaktligen har vi reducerat problemet till att söka största gemensam delare till två mindre tal.

[1] Med *lemma* menas en hjälpsats, ofta av mera teknisk karaktär och inte hållen för lika intressant som en fullvärdig sats.

Vi kan förstås fortsätta på samma sätt. Dividera igen det större talet med det mindre:

$$66 = 3 \cdot 18 + 12.$$

Precis som ovan inser vi, att

$$\text{SGD}(216, 66) = \text{SGD}(66, 18) = \text{SGD}(18, 12).$$

Nästa steg i algoritmen blir

$$18 = 1 \cdot 12 + 6,$$

så att

$$\text{SGD}(216, 66) = \text{SGD}(66, 18) = \text{SGD}(18, 12) = \text{SGD}(12, 6).$$

Sista steget blir

$$12 = 2 \cdot 6 + 0,$$

Resten är nu noll; divisionen gick jämnt ut. Största gemensamma delaren till 12 och 6 är alltså 6, och vi har visat att

$$\text{SGD}(216, 66) = \text{SGD}(66, 18) = \text{SGD}(18, 12) = \text{SGD}(12, 6) = 6.$$

Vi kunde ha argumenterat med vilka positiva heltal $a \geq b$ som helst i stället för 216 och 66. Först dividerar vi a med b:

$$a = kb + r, \qquad 0 \leq r \leq b - 1.$$

Vitsen med att göra divisionen är, med argumentet ovan, att

$$\text{SGD}(a, b) = \text{SGD}(b, r).$$

För varje delare till a och b är även en delare i linjärkombinationen $r = a - kb$. Omvänt är varje delare till b och r också en delare i linjärkombinationen $a = kb + r$.

Om nu $r = 0$, så är b själv en delare till både a och b; sålunda är $\text{SGD}(a, b) = b$ och vi är färdiga. Annars är b och r mindre tal än a och b, och vi kan upprepa divisionen. Talen krymper successivt ända tills någon rest blir 0. Då är vi klara. Detta är essensen i Euklides algoritm.

Sats 5.10: Euklides algoritm

Låt $a \geq b$ vara positiva heltal. Successiva divisioner leder till följande situation:

$$\begin{aligned} a &= kb + r \\ b &= k_1 r + r_1 \\ r &= k_2 r_1 + r_2 \\ &\vdots \\ r_{m-2} &= k_m r_{m-1} + r_m \\ r_{m-1} &= k_{m+1} r_m \end{aligned}$$

Då är $\text{SGD}(a, b) = r_m$ den sista icke-försvinnande resten.

Bevis

Enligt Divisionsalgoritmen är

$$b > r > r_1 > r_2 > \cdots \geq 0.$$

Resterna krymper alltså i varje steg, vilket visar, att algoritmen någon gång måste ta slut (efter maximalt b steg). Vi har redan argumenterat för ovan, att

$$\text{SGD}(a, b) = \text{SGD}(b, r) = \text{SGD}(r, r_1) = \cdots = \text{SGD}(r_{m-1}, r_m) = r_m.$$

Vi går nu in närmare på bråkförkortning.

Definition 5.11

Två tal med största gemensam delare 1 säges vara **relativt prima**. Ett heltalsbråk, i vilket täljare och nämnare är relativt prima, säges vara **slutförkortat**.

Exempel 13.
Vi vill förkorta bråket $\frac{66}{216}$. Ovan såg vi att $\text{SGD}(216, 66) = 6$, och alltså kan vi förkorta bråket enligt

$$\frac{66}{216} = \frac{6 \cdot 11}{6 \cdot 36} = \frac{11}{36}.$$

Talen 11 och 36 är relativt prima ($\text{SGD}(11, 36) = 1$). Bråket är slutförkortat.

Sats 5.12

Varje heltalsbråk, med positiv täljare och nämnare, är lika med ett unikt slutförkortat heltalsbråk, med positiv täljare och nämnare.

Bevis

Beviset är inte svårt och använder satsen om unik primtalsfaktorisering. Om vi har två slutförkortade bråk $\frac{a}{b} = \frac{a'}{b'}$, så får vi att $ab' = a'b$. Genom att primtalsfaktorisera bägge sidorna och utnyttja att a är relativt prima med b (så att deras faktoriseringar innehåller *olika* primtal), ser vi att varje primtalsfaktor i a också förekommer i a', med samma multiplicitet, och omvänt. Härav får vi $a = a'$, vilket sedan ger att $b = b'$.

Vill vi att satsen skall gälla även negativa bråk, kan vi kräva att minustecknet skall befinna sig i täljaren. Den allmännare satsen lyder då: *Varje heltalsbråk är lika med ett unikt slutförkortat heltalsbråk med positiv nämnare.* Så till exempel är

$$\frac{-286}{1859} = \frac{-2}{13}.$$

§8. Diofantiska ekvationer

Euklides algoritm visar sig ge ett verktyg för att lösa ekvationer av typen

$$ax + by = c,$$

där a, b och c är givna heltal och vi bara är intresserade av *heltalslösningar* x och y. Sådana ekvationer kallas *diofantiska*, efter den grekiske talteoretikern Diofantos (c:a 250 f.Kr.). Benämningen är inte helt historiskt korrekt; Diofantos själv sökte *rationella* lösningar till sina ekvationer.

Vore vi intresserade av att beskriva alla *reella* lösningar, så kunde vi (förstås) lätt beskriva dem genom att lösa ut y (åtminstone om $b \neq 0$). Alla lösningar (x, y) ges då av $y = \frac{c}{b} - \frac{a}{b}x$ för godtyckligt x. Lösningarna beskriver en linje i xy-planet. Men eftersom $\frac{c}{b} - \frac{a}{b}x$ inte behöver vara ett heltal bara för att x är det, så säger detta inte något om vilka heltalslösningar det finns. Vissa diofantiska ekvationer kan rentav helt sakna heltalslösningar.

Exempel 14.
Ekvationen $35x + 55y = 102$ har inga heltalslösningar. Ty oavsett vilka värden man ger x och y, kommer vänsterledet att vara delbart med 5. Men 102 är inte delbart med 5.

Litet mer tjusigt formulerat: Vänsterledet är en linjärkombination av 35 och 55, och enligt vårt tidigare lemma är det då delbart med $\text{SGD}(35, 55) = 5$. För att en ekvation $35x + 55y = c$ skall vara lösbar, måste således c vara delbart med 5.

Generaliserar vi argumentet i exemplet, har vi omedelbart följande lemma.

Lemma 5.13

Ett nödvändigt villkor för att det skall finnas lösningar till den diofantiska ekvationen $ax + by = c$ är att $\text{SGD}(a, b)$ delar c.

Om detta villkor är uppfyllt, kan vi dela båda leden i ekvationen med $\text{SGD}(a, b)$ och få en ekvivalent ekvation (samma lösningar). Efter divisionen kommer största gemensamma delaren av koefficienterna i vänsterledet att vara 1.

Exempel 15.
Ekvationen $35x + 55y = 100$, där $\text{SGD}(35, 55) = 5$, har samma lösningar som ekvationen $7x + 11y = 20$, för vilken $\text{SGD}(7, 11) = 1$.

I fortsättningen antar vi därför att $\text{SGD}(a, b) = 1$. I detta fall har ekvationen $ax + by = c$ oändligt många lösningar, och det finns ett systematiskt sätt att beskriva dessa, givet *en enda* lösning (x_0, y_0), en *partikulärlösning*, till ekvationen.

Sats 5.14

Den diofantiska ekvationen $ax + by = c$, där $\text{SGD}(a, b) = 1$, har den allmänna lösningen

$$\begin{cases} x = x_0 - bn \\ y = y_0 + an, \end{cases}$$

där n är ett godtyckligt heltal och (x_0, y_0) är en partikulärlösning till ekvationen.

Bevis

Vi verifierar först, medelst insättning, att formeln ovan verkligen ger lösningar till ekvationen:

$$ax + by = a(x_0 - bn) + b(y_0 + an) = ax_0 + by_0 = c,$$

ty $ax_0 + by_0 = c$, emedan (x_0, y_0) ju antogs vara en partikulärlösning till ekvationen.

Vi visar nu, att varje lösning måste vara på denna form. Antag alltså (x, y) vara någon lösning till $ax + by = c$. Då är

$$ax + by = c = ax_0 + by_0 \qquad \Leftrightarrow \qquad a(x - x_0) = -b(y - y_0).$$

Nu måste a dela $y - y_0$. Det kan inses genom att studera primtalsfaktoriseringarna av de fyra ingående talen a, b, $x - x_0$ och $y - y_0$. Eftersom primtalsfaktoriseringen av ett heltal är unik (bortsett från ordningen av faktorerna och eventuella tecken), så måste precis samma faktorer ingå i $a(x - x_0)$ som i $b(y - y_0)$. Men a och b har inga gemensamma primfaktorer enligt antagandet $\text{SGD}(a, b) = 1$, så alla primtalen, som ingår i a, måste också ingå i $y - y_0$ och med minst lika hög multiplicitet. Detta betyder just att a delar $y - y_0$.

Vi kan därför skriva $y - y_0 = an$, för något heltal n. Av

$$a(x - x_0) = -b(y - y_0) = -ban$$

följer sedan $x - x_0 = -bn$, och vår lösning (x, y) är alltså på angiven form. (Vi dividerade här med a, och förutsatte alltså tyst att detta var nollskilt. Om $a = 0$, men $b \neq 0$, kan vi föra ett liknande resonemang. Både a och b kan inte vara 0 på grund av villkoret $\text{SGD}(a, b) = 1$.)

Återstående problemet är alltså att finna en enda ynka partikulärlösning.

Exempel 16.
Ibland kan man hitta en partikulärlösning utan ansträngning. Ekvationen $7x + 34y = 1$ kan man enkelt inse har en lösning $(x_0, y_0) = (5, -1)$, varför dess allmänna lösning, enligt satsen, ges av formeln

$$\begin{cases} x = 5 - 34n \\ y = -1 + 7n. \end{cases}$$

En metod som alltid fungerar, förutsatt att $\text{SGD}(a, b) = 1$, är att hitta en lösning till *hjälpekvationen* $ax + by = 1$ genom Euklides algoritm körd baklänges, varpå denna multipliceras med c.

Exempel 17.
Vi vill lösa den diofantiska ekvationen $34x + 37y = 3$. Som $\text{SGD}(34, 37) = 1$, skall det enligt satsen finnas (oändligt många) lösningar till denna. För att finna en partikulärlösning, räknar vi först ut $\text{SGD}(34, 37)$ med Euklides algoritm.

$$37 = 1 \cdot 34 + 3$$
$$34 = 11 \cdot 3 + 1$$

Lös ut resterna i schemat:

$$3 = 37 - 34$$
$$1 = 34 - 11 \cdot 3$$

Utgå nu från den sista likheten $1 = 34 - 11 \cdot 3$ och ersätt 3 i denna med uttrycket för 3 från den första ekvationen:

$$1 = 34 - 11(37 - 34) = 12 \cdot 34 - 11 \cdot 37.$$

Hjälpekvationen $34x + 37y = 1$ har alltså en partikulärlösning $(12, -11)$. Den ursprungliga ekvationen $34x + 37y = 3$ har då en partikulärlösning

$$(x_0, y_0) = (3 \cdot 12, 3 \cdot (-11)) = (36, -33).$$

Allmänna lösningen är då enligt satsen

$$\begin{cases} x = 36 - 37n \\ y = -33 + 34n. \end{cases}$$

Euklides algoritm tog här slut efter två rader. I bevisskissen för Bézouts identitet i nästa avsnitt visar vi hur det går till att köra algoritmen baklänges när stegen är fler.

§9. Bevis för Aritmetikens fundamentalsats

Vi är nu rustade att bevisa den återstående delen av Aritmetikens fundamentalsats, entydigheten av en primtalsfaktorisering. Två viktiga resultat kommer att hjälpa oss på vägen.

Sats 5.15: Bézouts identitet

Största gemensamma delaren $\text{SGD}(a, b)$ till två heltal a och b kan skrivas som en linjärkombination av a och b, d.v.s. det finns heltal x och y, sådana att $\text{SGD}(a, b) = ax + by$.

Bevis

I stället för att ge ett stringent bevis, väljer vi att förklara själva bevisidén med hjälp av ett belysande exempel. För detta ändamål väljer vi två tal $a = 312$ och $b = 221$. För att

finna $\text{SGD}(a, b)$ tillämpar vi förstås Euklides algoritm:

$$\begin{aligned} 312 &= 1 \cdot 221 + 91 \\ 221 &= 2 \cdot 91 + 39 \\ 91 &= 2 \cdot 39 + 13 \\ 39 &= 3 \cdot 13 + 0 \end{aligned}$$

Då den sista nollskilda resten är 13, är alltså $\text{SGD}(312, 221) = 13$.

Nu vänder vi på stegen (steken) i Euklides algoritm och kör den baklänges. Vi börjar från den nästsista raden och jobbar oss uppåt:

$$\begin{aligned} 13 &= 91 - 2 \cdot 39 \\ &= 91 - 2 \cdot (221 - 2 \cdot 91) = 5 \cdot 91 - 2 \cdot 221 \\ &= 5 \cdot (312 - 1 \cdot 221) - 2 \cdot 221 = 5 \cdot 312 - 7 \cdot 221. \end{aligned}$$

Därmed lyckades vi med att skriva $13 = \text{SGD}(312, 221)$ som en linjärkombination av 312 och 221 (med $x = 5$ och $y = -7$). Notera att det vanligen kan krävas både positiva och negativa tal här.

Bézouts identitet kan också formuleras som så, att den diofantiska ekvationen $ax + by = \text{SGD}(a, b)$ har minst en lösning. Beviset använder precis den lösningsteknik som infördes i föregående avsnitt. För härvarande ändamål behöver vi bara *att* en lösning finns, och ingen vetskap om hur lösningen ser ut. Ur blotta existensen får vi nämligen följande mycket användbara resultat.

Lemma 5.16: Euklides lemma

Om primtalet p delar produkten ab av två heltal a och b, så måste p dela a eller b.

Bevis

Om p delar a, så är vi färdiga. Antag därför, att p inte delar a. Då är uppenbarligen $\text{SGD}(a, p) = 1$, eftersom p bara har de positiva delarna 1 och p. Vi kan då skriva $1 = ax+py$ enligt Bézouts identitet. Multiplikation med b ger

$$b = (ax + py)b = (ab)x + pby.$$

Enligt antagande är produkten ab delbar med p. Båda termerna i högerledet är då delbara med p, vilket innebär att talet b är delbart med p. Vi har alltså visat, att om a inte är delbart med p, så är b delbart med p. Minst ett av talen a och b är därmed delbart med p.

Euklides lemma låter sig elegant generaliseras till en produkt av fler faktorer än två: Om primtalet p delar en produkt $a_1 a_2 \cdots a_k$ av heltal, så måste p dela någon av faktorerna a_i. Det intuitiva beviset bygger på en upprepad tillämpning av föregående lemma. Antag, att p inte delar något av talen $a_1, \ldots, a_{k-1}$. Som p inte delar a_1, men delar $a_1(a_2 \cdots a_k)$, så måste p dela produkten $a_2 \cdots a_k$. Som p inte delar a_2, men delar $a_2(a_3 \cdots a_k)$, så måste p dela produkten $a_3 \cdots a_k$. Upprepar vi detta $k - 1$ gånger, finner vi slutligen, att p måste dela a_k.

Exempel 18.

En snabb illustration: Är 65537^{99} delbart med 17? Visserligen är 65537 ett slags kändisprimtal och alltså absolut inte delbart med 17, men hjärtat sjunker i bröstet, när man inser, att 65537^{99} har 482 siffror, vilket alltså är cirka 474 fler siffror än i alla fall våra Medelsvensson-miniräknare kan hantera. Att besvara frågan med heltalsdivision kommer alltså knappast på fråga. Som tur är räddas vi av Euklides lemma. Talet 65537 är inte ensamt delbart med 17, och därför ingen potens av det heller.

Efter detta reklaminslag är vi redo att bevisa entydigheten av primtalsfaktoriseringar. De första talen 2, 3 och 4 kontrollerar vi enkelt har entydiga faktoriseringar, och vi kan förstås gå ännu fler steg rent numeriskt. Om nu primtalsfaktoriseringar inte alltid är entydiga, måste det finnas ett *minsta* tal n med två olika primtalsfaktoriseringar:

$$p_1 p_2 \cdots p_r = n = q_1 q_2 \cdots q_s.$$

Då primtalet p_1 delar vänstra ledet, så är p_1 också en delare till $q_1 q_2 \cdots q_s$. Enligt Euklides lemma måste p_1 dela något av talen q_1, q_2, …, q_s. Vi kan utan vidare antaga (efter en eventuell omordning av talen q_i), att p_1 delar q_1. Då båda dessa är primtal, följer att $p_1 = q_1$. Division med p_1 reducerar därför likheten till

$$p_2 \cdots p_r = q_2 \cdots q_s,$$

och vi har här ett *mindre* tal än n med två olika primtalsfaktoriseringar. Denna motsägelse bevisar, att entydigheten måste gälla alla tal. Beviset för Aritmetikens fundamentalsats är därmed avslutat.

Övningar

5.1. Primtalsfaktorisera 1110.

5.2. Konstruera med Eratosthenes såll en lista över alla primtal upp till 100.

5.3. Talen 113 och 1117 är primtal. Låt $a = 2^5 \cdot 113^3 \cdot 1117^2$ och $b = 2^2 \cdot 113^1 \cdot 1117^{51}$. Vad är $\text{SGD}(a, b)$?

5.4. Bestäm $\text{SGD}(9563, 6205)$.

5.5. Förkorta så långt som möjligt bråken

(a) $\dfrac{3451}{8381}$;

(b) $\dfrac{196\ 707}{250\ 971}$.

5.6. Lös de diofantiska ekvationerna

(a) $14x - 21y = 333$;

(b) $114x + 303y = 168$;

(c) $43x + 19y = 4$.

5.7. Lille Per har av sin moder fått 250 kr för att gå till konditoriet och köpa lyxsemlor till ett pris av 17 kr stycket och mandelkakor till ett pris av 6 kr stycket. När han är framme i konditoriet har han hunnit glömma, hur många av de två slagen bakverk han skulle köpa. Han minns dock klart, att pengarna skulle räcka precis, och att antalet mandelkakor var udda. Hjälp lille Per!

5.8. Visa, att $k^2 + k$ är delbart med 2 för varje heltal k.

5.9. Bevisa, att $n^2 - 1$ är delbart med 8 för alla udda heltal n.

5.10. Bevisa, att talet $n^3 - n$ alltid är delbart med 6, och till yttermera visso med 24, om n är udda.

5.11. (a) Visa att, om n är sammansatt, så har det en delare a med $1 < a \leq \sqrt{n}$.

(b) Förklara, varför det räcker att kontrollera delbarhet med tal upp till $\sqrt{n}$, då det skall verifieras, att ett givet tal n är ett primtal.

(c) Visa, med hjälp av detta och huvudräkning, att 113 är ett primtal.

(d) Hur många divisioner med en miniräknare behöver du som mest genomföra för att undersöka, om 8521 är ett primtal?

5.12. Ett positivt heltal kallas *perfekt* (eller *fullkomligt*), om det är lika med summan av alla sina delare förutom talet självt. Exempelvis är 6 perfekt, ty $6 = 1 + 2 + 3$. Bestäm primtalet p, så att talet $16p$ blir perfekt.

5.13. Är det möjligt, att något av heltalen x eller y är delbart med 3, om

$$x^2 - y^2 = 1995 \ ?$$

Kapitel 6

Moduliräkning

§1. Matematiknerden som urmakare

En viktig lärdom från småskolan var konsten att avläsa en klocka, och som så mycket annat vardagligt, rymmer detta fenomen förvånansvärt djup matematik. En vanlig tolvtimmarsklocka "nollställs" var tolfte timme, så att klockan 15 är detsamma som klockan 3 (visserligen det ena på eftermiddagen, det andra på natten). Med litet fantasi kan vi även tänka oss att klockslaget "27" också skulle betyda alldeles detsamma som klockan 3. Vi säger, att talen 3, 15 och 27 är *kongruenta modulo* 12, och skriver

$$3 \equiv 15 \equiv 27 \pmod{12}.$$

Vid klockslaget "n" skulle då klockan visa på r, där r är resten vid division av n med 12.

Att vi har 12 timmar på ett halvt dygn beror säkert på någon historisk tillfällighet, t.ex. antalet fingrar på den förste urmakarens händer. Under första republikens tid, med dess besatthet av normering och standardisering, experimenterade fransmännen med ett decimalsystem för tid:

> Le jour, de minuit à minuit, est divisé en dix parties, chaque partie en dix autres, ainsi de suite jusqu'à la plus petite portion commensurable de la durée. La centième partie de l'heure est appelée *minute décimale*; la centième partie de la minute est appelée *seconde décimale*.
> (Officiellt dekret från 1793)

SI-systemet gjorde stor succé och erövrade hela den civiliserade världen. De decimala klockorna slog däremot aldrig igenom och föll snart åter i glömska.

Talet 10 är förstås också en mänsklig artefakt. Lika gärna kunde vi tänka oss en klocka, som går ett varv på b timmar, för valfritt positivt heltal b. Detta är den centrala idén i innevarande kapitel: *moduliräkning*, alias kongruensräkning — för den som tycker tal för det mesta är alldeles för stora.

Definition 6.1

Låt b vara ett positivt heltal. Två heltal m och n sägs vara **kongruenta modulo** b, om de lämnar samma rest vid division med b. Vi skriver detta som

$$m \equiv n \pmod{b}.$$

Exempel 1.
Talen

$$13 = 1 \cdot 12 + 1, \qquad 25 = 2 \cdot 12 + 1, \qquad 37 = 3 \cdot 12 + 1$$

lämnar alla samma rest vid division med 12, nämligen 1, och likaså talen

$$-23 = (-2) \cdot 12 + 1, \qquad -11 = (-1) \cdot 12 + 1, \qquad 1 = 0 \cdot 12 + 1.$$

Alltså är

$$-23 \equiv -11 \equiv 1 \equiv 13 \equiv 25 \equiv 37 \pmod{12}.$$

En omedelbar konsekvens av definitionen är, att de hela talen faller i b stycken **kongruensklasser** eller **restklasser** modulo b.

Exempel 2.
Kongruensklasserna modulo 2 utgöres av

$$\cdots \equiv -4 \equiv -2 \equiv 0 \equiv 2 \equiv 4 \equiv \cdots$$
$$\cdots \equiv -3 \equiv -1 \equiv 1 \equiv 3 \equiv 5 \equiv \cdots,$$

vilket är den vanliga uppdelningen i jämna och udda tal.

Exempel 3.
Låt oss beskriva kongruensklasserna modulo 7. Alla tal som är delbara med 7 har samma rest 0 vid division med 7:

$$\cdots \equiv -7 \equiv 0 \equiv 7 \equiv 14 \equiv \cdots \equiv 7k \equiv \cdots$$

för varje heltal k. På samma sätt består klassen

$$\cdots \equiv -6 \equiv 1 \equiv 8 \equiv 15 \equiv \cdots \equiv 7k+1 \equiv \cdots,$$

av de tal, vilka lämnar resten 1 vid division med 7.

Generellt ser vi att, om vi fixerar en rest $r = 0, 1, \ldots, 6$, så är alla tal av formen $7k + r$ kongruenta med varandra modulo 7.

§2. Kongruensrelationens egenskaper

Man kan notera, hurusom symbolen $\equiv$ äger en mer än flyktig likhet med likhetstecknet $=$. Det är inte orätt, att tänka på kongruens som ett slags generaliserad likhet, eftersom de två relationerna har samma sorts egenskaper. Följande sats ger ett stramare uttryck för denna tanke.

Sats 6.2

Kongruensrelationen är en **ekvivalensrelation**, vilket betyder att den besitter följande tre egenskaper.

(a) Reflexivitet: $m \equiv m \pmod{b}$.

(b) Symmetri: $m \equiv n \pmod{b}$ om och endast om $n \equiv m \pmod{b}$,

(c) Transitivitet: $m \equiv n \pmod{b}$ och $n \equiv k \pmod{b}$ medför $m \equiv k \pmod{b}$.

Bevis

Vi bevisar bara transitiviteten, de övriga lämnas åt läsaren. Påståendena $m \equiv n \pmod{b}$ och $n \equiv k \pmod{b}$ säger, enligt definitionen, att m och n har samma rest, liksom n och k (vid division med b). Men då är det klart, att m och k har samma rest, vilket enligt definitionen innebär, att $m \equiv k \pmod{b}$.

Ett användbart alternativt kriterium för kongruens är följande.

Sats 6.3

Det gäller att $m \equiv n \pmod{b}$ om och endast om $b \mid (m - n)$.

Bevis

Satsen ådagalägger ekvivalensen av två utsagor, och vi måste då bevisa, att endera påståendet medför det andra. Antag först, att $m \equiv n \pmod{b}$. Enligt definitionen ovan innebär det att m och n lämnar samma rest vid division med b. Kalla denna rest r; då är $m = kb + r$ och $n = lb + r$ för några heltal k, l. Det leder till

$$m - n = (kb + r) - (lb + r) = (k - l)b,$$

och alltså (enligt definitionen av delbarhet) att $b \mid (m - n)$.

Vi visar nu implikationen åt andra hållet. Antag alltså omvänt, att $b \mid (m - n)$, så att $m - n = ab$ för något heltal a. Enligt Divisionsalgoritmen kan vi dividera n med b och skriva $n = kb + r$, där resten uppfyller $0 \leq r < b$. Nu är

$$m = (m - n) + n = ab + kb + r = (a + k)b + r.$$

Eftersom $0 \leq r < b$, är detta det entydiga resultatet av division av m med b. Talet r är rest för både n och m vid division med b; sålunda $m \equiv n \pmod{b}$.

§3. Räkneregler för kongruenser

Riktigt roligt (lärarspråk för ...) och i alla fall användbart blir det inte, förrän vi inser att vi faktiskt kan räkna med rester nästan som vanligt. Litet löst, och med en gnutta kreativ bokföring, säger denna sats, att *resten modulo b av ett algebraiskt uttryck inte förändras om vi byter talen i uttrycket mot andra tal med samma rest modulo b.*

Sats 6.4

Antag $m \equiv m' \pmod{b}$ och $n \equiv n' \pmod{b}$. Då är

$$m + n \equiv m' + n' \pmod{b} \qquad \text{och} \qquad mn \equiv m'n' \pmod{b}.$$

Bevis

Enligt förutsättningarna och Sats 6.3 finns det tal k och l, så att $m' - m = kb$ och $n' - n = lb$. Ur detta följer regeln för addition direkt. Enär

$$(m' + n') - (m + n) = (m + kb) + (n + lb) - m - n = (k + l)b$$

är delbart med b, är $m' + n' \equiv m + n$ enligt Sats 6.3. För multiplikation har vi på motsvarande vis, att

$$m'n' - mn = (m + kb)(n + lb) - mn = (lm + kn + klb)b$$

är delbart med b, och således är $m'n' \equiv mn$ enligt Sats 6.3.

Exempel 4.
En välkänd delbarhetsregel säger, att ett positivt tal är delbart med 3 precis då dess siffersumma är det. Vi visar den mer detaljerade utsagan, att ett positivt tal är kongruent med sin siffersumma modulo 3.

Notera att $10 = 3 \cdot 3 + 1$, så $10 \equiv 1 \pmod{3}$. Enligt satsen är då

$$100 = 10 \cdot 10 \equiv 1 \cdot 1 \pmod{3},$$

så att $100 \equiv 1$. Ur det följer, enligt satsen igen, att

$$1000 = 10 \cdot 100 \equiv 1 \cdot 1 = 1.$$

Genom att fortsätta på samma vis inser vi, att alla tiopotenser är kongruenta med 1 modulo 3:

$$10^k \equiv 1 \pmod 3, \qquad k \geq 0.$$

Låt oss nu, för enkelhets skull, betrakta ett fyrsiffrigt tal $\overline{abcd}$ (med *siffrorna* a, b, c och d). Som en ytterligare konsekvens av satsen ovan kan vi, i uttryck med multiplikationer och additioner av tal, utbyta alla 10-potenser mot 1 utan att förändra resten modulo 3, så att

$$\overline{abcd} = 1000a + 100b + 10c + d \equiv a + b + c + d \pmod 3$$

är kongruent med sin siffersumma. Beviset för större tal är helt analogt.

Exempel 5.
Vilken rest fås vid division av 123 456 789 med 3? Enligt föregående exempel är resten

$$123\,456\,789 \equiv 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 = 45 \equiv 0 \pmod 3,$$

alltså är talet delbart med 3.

Exempel 6.
Varje publicerad bok har en *ISBN-kod*, som innehåller tio siffror i fyra grupper, t.ex. 91–7738–647–7. Den första gruppen är en språkkod (91 för svenska, 0 för engelska, etc.), den andra gruppen förlagskoden, den tredje bokens nummer inom förlaget och den sista en kontrollsiffra, med vars hjälp det kontrolleras, om de övriga siffrorna är rätt kopierade (som för personnumren). I ISBN-koden $a_1a_2a_3 \ldots a_{10}$ är kontrollsiffran a_{10} vald enligt formeln

$$a_{10} \equiv a_1 + 2a_2 + 3a_3 + \cdots + 9a_9 \pmod{11}.$$

Addition av $10a_{10}$ till båda leden ger den ekvivalenta ekvationen

$$0 \equiv 11a_{10} \equiv a_1 + 2a_2 + 3a_3 + \cdots + 10a_{10} \pmod{11}$$

(kom ihåg att $11 \equiv 0 \pmod{11}$), som varje giltig kod må uppfylla. Kontrollsiffran är alltså en rest vid division med 11 och kan vara en av siffrorna 0 till 9 eller bokstaven X, kodande resten 10.

Exponentiering är ju upprepad multiplikation, så av Sats 6.4 får vi direkt:

Sats 6.5

Om $m \equiv m' \pmod b$, så är $m^k \equiv (m')^k \pmod b$ för varje naturligt tal k.

Exempel 7.
Om Universum är tillräckligt sparsamt befolkat av materia, kommer tyngdkraften inte att förmå bromsa dess expansion. Det kommer då att fortsätta utvidgas i all evighet, allt glesare, mörkare och tommare. Stjärnorna kommer att slockna. Kvarlevorna kommer att sugas ned i de svarta hålen, som irrar rymden kring som ett slags kosmiska dammsugare. Men dessa är inte eviga, utan det kan inträffa, rent ofattbart sällan visserligen, att en elementarpartikel lyckas fly hålet (det är något kvantmekaniskt). Det förlorar därigenom massa och kommer alltså själv en vacker dag att dö. Så småningom kommer det sista svarta hålet att ha slocknat, och rymden är tömd på materia. Universum är helt tomt och ingenting mer kommer någonsin att hända. Det beräknas inträffa om cirka 10^{100} dagar (eller möjligen år; det är inte så kvistigt på en kosmisk tidsskala).

Den viktiga frågan är nu: *Vilken veckodag inträffar detta?* Det handlar förstås om, att beräkna resten av 10^{100} modulo 7, vilket går galant med moduliräkning och satsen ovan. Vi nyttjar successivt kongruenserna $10 \equiv 3$, $9 \equiv 2$ och $8 \equiv 1$:

$$10^{100} \equiv 3^{100} = 9^{50} \equiv 2^{50} = (2^3)^{16} \cdot 2^2 = 8^{16} \cdot 4 \equiv 1^{16} \cdot 4 = 4 \pmod 7.$$

I dag är det torsdag (det brukar det vara, då denna föreläsning avhålles). Världen går då under fyra veckodagar senare, alltså på en måndag. Förstås.

Njut förslagsvis ett litet tag av vad vi sluppit genom att räkna så här. Heltalet 10^{100} är mycket stort (101 siffror), så vilken tid och hur många döda träd skulle inte ha krävts för att med trappan eller liggande stolen dela det med 7?

Vi kan nå vår domedagsprofetia ännu kvickare genom följande celebra sats.

Sats 6.6: Fermats lilla sats

Låt p vara ett primtal. Om $p \nmid a$, är

$$a^{p-1} \equiv 1 \pmod{p}.$$

Multiplicerar vi kongruensen i satsen med a, får vi $a^p \equiv a \pmod{p}$, och detta håller sträck för alla heltal a utan inskränkning. Beviset för satsen är inte så svårt och kommer i en senare kurs.

Exempel 8.
Med Fermats lilla sats får vi direkt

$$10^{100} = (10^6)^{16} \cdot 10^4 \equiv 1^{16} \cdot 3^4 = 81 \equiv 4 \pmod{7},$$

ty primtalet $7 \nmid 10$.

§4. Varning

Kongruenser kan alltså adderas, subtraheras och multipliceras. Men det finns ju fyra räknesätt? Division fungerar mindre väl. Moduliräkning handlar ju, för det första, om heltal, och division leder i allmänhet till rationella tal. Så det är oklart hur division ens skall definieras. (Vad skulle exempelvis $\frac{2}{3}$ betyda modulo 6?)

Inte ens om resultatet blir ett heltal låter det sig dock göras problemfritt. Till exempel är $2 \equiv 14 \mod 12$, men $1 \not\equiv 7 \pmod{12}$. Kongruensen kan alltså inte divideras med 2.

Förutom avsaknaden av division är teorien för kongruenser behäftad med vissa andra egenheter. Somliga räknelagar fungerar helt enkelt inte för kongruenser. För reella (och komplexa) tal vet vi, att $ab = 0$ medför $a = 0$ eller $b = 0$. Detta är inte sant för kongruenser. Vi har $2 \cdot 3 \equiv 0 \pmod{6}$, men varken 2 eller 3 är $0 \pmod{6}$. Som en konsekvens kan man inte heller ur likheten $ax \equiv ay \pmod{b}$, där $a \not\equiv 0 \pmod{b}$, dra slutsatsen att $x = y$. Detta är relaterat till divisionsproblemet.

Det går att rädda situationen, i synnerhet om vi räknar modulo primtal, men det är en annan historia (avhandlas i högre kurser i talteori eller abstrakt algebra).

§5. Vad moduliräkning betyder för just ditt kriminella nätverk

I många sammanhang vill man, att en text bara skall kunna läsas av en enda mottagare. När man t.ex. knappar in sin bankomatkod vore det katastrofalt, om den kunde läsas av någon som avlyssnar den elektroniska trafiken. Därför *krypteras* informationen innan den skickas på ett sätt som garanterar, att endast banken kan läsa den.

Samma typ av problem uppkommer, när lärare skall skicka tentamina över e-post. De kan inte skickas i klartext, eftersom det inte är alldeles otänkbart, att några duktiga studenter tjuvlyssnar över nätet (så skickliga studenter är ju förstås en lärares dröm).

Det har länge varit en kamp mellan dem, som konstruerar krypteringssystem, och dem, som försöker bryta dem. Datorernas intåg har förstås underlättat kodknäckandet. Fantomens hemliga chiffer tar inte många millisekunder att lösa. Samtidigt har användningen av datorer inneburit,

att mycket mer känslig information kan spridas för missbruk, så att det uppstått ett enormt behov av säker kryptering.

År 1977 utvecklades krypteringssystemet *RSA*, som, troligen för all överskådlig framtid, kommer att kunna bibehållas tillräckligt säkert. Algebra och talteori har skänkt det godas förkämpar, chiffermakarna, seger över de onda anstiftarna, dechiffrerarna.

RSA bygger på de enklaste principer av moduliräkning. Situationen vi föreställer oss är, att Alice önskar skicka hemliga meddelanden till Bob (namnvalen är standard inom denna industri).

1. Hemliga nyckeln. Alice och Bob kommer hemligt överens om två olika primtal p och q. De väljer sedan två heltal d och e, sådana att

$$de \equiv 1 \pmod{(p-1)(q-1)}.$$

Talen pq och e kan vara mer eller mindre offentlig handling, men talen p, q och d vaktar de noga på.

2. Kryptering. Alice vill skicka det hemliga heltalet x (som förstås kan vara kodad text) till Bob. Talet x förväntas ligga i intervallet $0 \leq x \leq pq-1$. (Har hon mycket på hjärtat eller är av den pladdriga sorten, så får hon stycka upp sitt meddelande.) Alice beräknar talet x^e, reducerar till resten modulo pq, och skickar.

3. Dekryptering. Bob tar emot ett tal $y \equiv x^e \pmod{pq}$ i sin ände. För att dekryptera, beräknar Bob nu bara talet y^d och tar resten modulo pq. Då får han $y^d \equiv x^{de} \pmod{pq}$, och detta är samma som $x \pmod{pq}$, alltså Alices ursprungliga meddelande, enligt nedan.

4. Varför fungerar RSA? Villkoret $de \equiv 1 \pmod{(p-1)(q-1)}$ betyder, att vi kan skriva $de - 1 = m(p-1)(q-1)$ för något heltal m. Vi har

$$x^{de-1} = x^{m(p-1)(q-1)} = (x^{p-1})^{m(q-1)} \equiv 1^{m(q-1)} = 1 \pmod{p}$$

enligt Fermats lilla sats, varav talet $x^{de-1}-1$ är delbart med p. På samma sätt är det även delbart med q och därför med produkten pq; således $x^{de-1} \equiv 1 \pmod{pq}$. Bobs metod för dekryptering frambringar då

$$y^d \equiv x^{de} = x^{de-1} \cdot x \equiv 1 \cdot x = x \pmod{pq},$$

som önskat.

5. Varför är RSA säkert? Den ondskefulla Eva eller Eve (namnet anspelar inte så mycket på syndafallet (synd på så rar äppelkaka) som på den engelska glosan *Eavesdropper*) skulle kunna interceptera Alices krypterade meddelande $y \equiv x^e \pmod{pq}$. Hon står dock rådvill inför dess avkodning till klartext, ty för detta ändamål skulle hon ha bruk för talet d, och därom har hon ej den ringaste aning.

Talet e kan däremot vara mer eller mindre offentlig handling, men det hjälper inte Eva det bittersta, såvida hon inte också känner primtalen p och q. Givet dessa och e, skulle hon kunna lista ut d ur formeln $de \equiv 1 \pmod{(p-1)(q-1)}$. Notera att det inte räcker att veta produkten pq för att lista ut $(p-1)(q-1)$. Eva måste veta faktorerna p och q.

Produkten pq och talet e kan förstås också försöka hemlighållas, men de behövs för att skicka de krypterade meddelandena. De måste då vara kända av alla kommunikatörer och är delikata att hemlighålla. Här närmar vi oss RSA-systemets ömma punkt.

Det är för närvarande extremt svårt, både tids- och datorkrävande, att primtalsfaktorisera stora tal. Funne man en snabb algoritm för detta, kunde RSA alltså knäckas. Eva kunde finna faktorerna p och q i pq och därifrån den hemliga nyckeln d, och det vore usla nyheter för världens säkerhet och våra sparpengar. Det har emellertid länge varit en fast trossats hos de invigda, att några sådana (tillräckligt snabba) algoritmer ej existerar. Om det tar tusen år att faktorisera ett tusensiffrigt tal, ombesörjer Alice klokt nog ett tiotusensiffrigt tal pq och skiftar det en gång om året. Eva har inte en chans att hänga med.

Vi har bara surfat litet på ytan här. Mer detaljer finns i t.ex. Wikipedias artiklar om kryptering och RSA-algoritmen.

Övningar

6.1. Visa att $3^{100} - 1$ är jämnt delbart med 16.

6.2. Vilken rest erhålles då 2^{100} divideras med 23?

6.3. Vilken är sista siffran i talen

(a) 37^{100};

(b) $(17^{15} + 8^{25})^{10}$?

6.4. (a) Är 0–467–51402–X en giltig ISBN-kod?

(b) Är 1–56–004151–5 en giltig ISBN-kod?

(c) Vad är kontrollsiffran för 14–2000–076–? ?

6.5. Visa med kongruensräkning, att talet $n^3 + 2n$ är delbart med 3 för alla heltal n.

6.6. Antag n ej vara en multipel av 7. Visa, att n^6 lämnar resten 1 vid division med 7

(a) med Fermats lilla sats;

(b) utan Fermats lilla sats.

6.7. Vilka heltal $g = 0, 1, 2, \ldots, 6$ har egenskapen att resterna modulo 7 av potenserna $g^1, g^2, g^3, \ldots$ täcker in samtliga sex nollskilda kongruensklasser modulo 7 (alltså att varje heltal, ej delbart med 7, är kongruent med någon potens g^k)?

6.8. Vilka sexsiffriga tal av formen *abccba* är jämnt delbara med 33?

6.9. Är

$$\frac{19^{92} - 91^{29}}{90}$$

ett naturligt tal?

Del III

Komplex algebra

Kapitel 7

Komplexa tal på rektangulär form

§1. De komplexa talens historia

Att något är ett tal innebär, löst sagt, att det skall gå att räkna med det, ungefär som vi räknar med vanliga heltal. Det innebär, att samma räknelagar, t.ex. den distributiva lagen, bör gälla. Genom skolan har vi successivt utökat vår repertoar från heltalen $\mathbb{Z}$ till de rationella talen $\mathbb{Q}$ till att även omfatta de reella talen $\mathbb{R}$ (oändliga decimalbråk). Men nu, då livet i $\mathbb{R}$ har mist sin oskuldsfulla tjusning, skall ytterligare en utvidgning äga rum, då vi tar steget till de komplexa talen $\mathbb{C}$.

De komplexa talen har två ursprung. Ett är algebraiskt, och idéhistoriskt tankeväckande. Då man under 1500-talet härledde formler för lösningarna till tredje- och fjärdegradsekvationer, dök det upp kvadratrötter ur negativa tal som ett nödvändigt mellanled. Sådana "fanns" ju inte (och läsaren kanske fortfarande tycker, att de inte riktigt "finns"), och det höll man hårt på. Men det visade sig tekniskt bekvämt att räkna med dessa och ändå behandla dem som ett slags tal, fast *imaginära*. De fanns bara på låtsas, till skillnad mot de verkliga, *reella* talen. Slutresultatet blev ju ändå lösningar, som var riktiga reella tal, och det kunde verifieras, genom prövning av de erhållna rötterna, att denna utflykt i fantasien gav korrekta resultat.

Efter några hundra års komplext slavarbete och i ett mer filosofiskt sofistikerat tankeklimat, bestämde sig matematikerna för att skänka de komplexa talen samma status som de reella åtnjöt. Deras existens stadfästes.

Detta skedde under påverkan av de komplexa talens andra ursprung — det geometriska. Visst är det så, att de reella talen ter sig som mest övertygande, och till på köpet användbara, när de tänkes såsom punkter längs den *reella tallinjen*? Då uppstår naturligt frågan, varför samma förfarande ej skulle låta sig genomföras i planet? (Eller rummet?) Vi skall se, att de komplexa talen svarar precis mot punkter i ett talplan. Den praktiska användbarheten indikeras klart och tydligt därav, att en av pionjärena för detta geometriska synsätt var en lantmätare, dansken Caspar Wessel (kring år 1800). Siktet var inställt på att förenkla en del geometri och trigonometri, t.ex. vinkelberäkningar.

§2. Räkning med komplexa tal på riktigt

Först skall vi se hur enkelt det ändå är, att räkna med komplexa tal i praktiken. Vi inför den *imaginära enheten* i, med egenskapen $i^2 = -1$. Något sådant *reellt* tal finns förstås inte, men vi fixar fiffigt till en kvadratrot till -1. Om läsaren känner, att det ligger något av bilskojeri över detta och vägrar låta sig imponeras, så är det inte helt fel. Varför det inte är ett fall för Allmänna reklamationsnämnden, återkommer vi till om ett litet tag.

Ett komplext tal definieras nu som ett tal på formen $a+bi$, där a och b är reella. Dessa räknar vi sedan med "som vanligt", där vi tänker på i som en variabel eller obekant. Den praktiska tumregeln är: *Räkna med komplexa tal som vore de polynom, med den extra regeln att* i^2 *kan ersättas med* -1*, närhelst den dyker upp.*

Exempel 1.
Vi har exempelvis

$$(1+2i)+(3+4i)+(5+6i) = (1+3+5)+(2+4+6)i = 9+12i$$

och

$$(2+3i)(1-i) = 2\cdot 1 + 2\cdot(-i) + (3i)\cdot 1 + 3i\cdot(-i) = 2-2i+3i+3 = 5+i.$$

Exempel 2.
Ekvationen $z+2i = 1+3i$ löses på vanligt sätt genom att dra bort $2i$ från bägge sidor. Lösningen är $z = (1+3i)-2i = 1+i$.

Addition, subtraktion och multiplikation är alltså oproblematiska. Divisionen $\frac{a+bi}{c+di}$ utföres genom förlängning med $c-di$, *konjugatet* av nämnaren $c+di$:

$$\frac{a+bi}{c+di} = \frac{(a+bi)(c-di)}{(c+di)(c-di)} = \frac{ac-adi+bci-bdi^2}{c^2-d^2i^2}$$
$$= \frac{(ac+bd)+(bc-ad)i}{c^2+d^2} = \frac{ac+bd}{c^2+d^2} + \frac{bc-ad}{c^2+d^2}i.$$

Nämnaren c^2+d^2 är reell och vållar inte längre några bekymmer. Formeln behöver (bör) inte själlöst inbankas. Man minns *metoden*, förlängning med nämnarens konjugat.

Exempel 3.
Vi har

$$\frac{2+i}{1-i} = \frac{(2+i)(1+i)}{(1-i)(1+i)} = \frac{1+3i}{2} = \frac{1}{2}+\frac{3}{2}i.$$

Genom förlängningen har vi fått ett reellt tal i nämnaren, samt räknat ut täljaren med vanlig multiplikation.

Exempel 4.
Vi löser ekvationen $(2+3i)z = 1+i$. Vi kan naturligtvis dividera båda led med $2+3i$. Vi kan också multiplicera bägge led med $2-3i$, vilket ger

$$(2-3i)(2+3i)z = (2-3i)(1+i) \quad\Leftrightarrow\quad (2^2+3^2)z = 5-i \quad\Leftrightarrow\quad z = \frac{5}{13}-\frac{1}{13}i.$$

Här använde vi att $(2-3i)(2+3i) = 2^2+3^2 = 13$.

I skolan introduceras vanligen komplexa tal så som vi gjorde ovan. Först frambesvärjer vi ett nytt tal i med den besynnerliga egenskapen, att $i^2 = -1$. Sedan definierar vi komplexa tal som "tal" på formen $a+bi$, där a och b är reella. Visst blev det lätt att räkna med dem, men filosofiskt är detta otillfredsställande. Får vi "hitta på" nya tal eller matematiska objekt hur som helst? Varför skulle våra gamla räknelagar bara fortsätta att gälla? Tydligen krävde vi, att de komplexa talen skulle fungera precis som de reella. Men hur vet vi, att vi inte råkar in i ett träsk av självmotsägelser och paradoxer?

Kritiken är befogad. De komplexa talen fungerar naturligtvis, men det är inte uppenbart ur ovanbemälda tillvägagångssätt. Mer korrekt är den formella procedur vi nu skall beskriva. De komplexa talen "hittas inte på" eller trollas fram ur ett intet; de *konstrueras* i termer av de tal vi redan känner, de reella. Sedan *bevisar* vi, att de verkligen lyder under samma räknelagar som de reella talen.

§3. Konstruktion av de komplexa talen

Nu kommer en formell definition, som inte gör någon glad, följd av strängt bevisade räknelagar.

Definition 7.1

Ett **komplext tal** är ett par (a, b) av reella tal. Addition och multiplikation definieras av formlerna

$$(a, b) + (c, d) = (a + c, b + d), \qquad (a, b) \cdot (c, d) = (ac - bd, ad + bc).$$

Mängden av komplexa tal betecknas $\mathbb{C}$.

Addition sker koordinatvis, enligt samma princip som för addition av vektorer i $\mathbb{R}^2$. Redan här kan vi skönja en koppling till geometrien i planet. Multiplikationen verkar däremot uppsatt på en höft, och visst ser det mer än osannolikt ut, att den skall lyda vanliga hederliga räknelagar?

Sats 7.2: Räknelagar för addition

Addition av komplexa tal lyder under följande räknelagar.

(a) Associativa lagen: $(x + y) + z = x + (y + z)$.

(b) Kommutativa lagen: $x + y = y + x$.

(c) Nolla för addition: $x + (0, 0) = x$.

(d) Additiva inverser: Varje $x = (a, b)$ har en unik additiv invers $-x = (-a, -b)$, sådan att $x + (-x) = (0, 0)$.

Bevis

Alla räknelagarna följer direkt ur definitionen av addition. Till exempel visas Kommutativa lagen genom

$$(a, b) + (c, d) = (a + c, b + d) = (c + a, d + b) = (c, d) + (a, b).$$

Att $(-a, -b)$ är additiv invers till (a, b) visas genom

$$(a, b) + (-a, -b) = (a + (-a), b + (-b)) = (0, 0),$$

och omvänt är det klart, att enda talet (c, d) med egenskapen

$$(0, 0) = (a, b) + (c, d) = (a + c, b + d)$$

är just $(c, d) = (-a, -b)$.

Precis som för reella tal, låter de additiva inverserna oss definiera subtraktion enligt

$$x - y = x + (-y),$$

således

$$(a, b) - (c, d) = (a, b) + (-c, -d) = (a - c, b - d).$$

Sats 7.3: Distributiva lagarna

Det gäller att

$$(x + y)z = xz + yz \qquad \text{och} \qquad z(x + y) = zx + zy$$

för alla komplexa tal x, y, z.

Bevis

Den vänstra följer av räkningen

$$\begin{aligned}\Big((a,b)+(c,d)\Big)\cdot(e,f) &= (a+c,b+d)(e,f)\\ &= \Big((a+c)e-(b+d)f,\ (a+c)f+(b+d)e\Big)\\ &= (ae-bf,af+be)+(ce-df,cf+de) = (a,b)(e,f)+(c,d)(e,f),\end{aligned}$$

den högra på samma sätt.

Sats 7.4: Räknelagar för multiplikation

Multiplikation av komplexa tal lyder under följande räknelagar.

(a) Associativa lagen: $(xy)z = x(yz)$.

(b) Kommutativa lagen: $xy = yx$.

(c) Etta för multiplikation: $x \cdot (1,0) = x$.

(d) Multiplikativa inverser: Varje $x = (a,b) \neq (0,0)$ har en unik multiplikativ invers $x^{-1} = \Big(\frac{a}{a^2+b^2},\ -\frac{b}{a^2+b^2}\Big)$, sådan att $x \cdot x^{-1} = (1,0)$.

Bevis

Åter är dessa direkta konsekvenser av definitionen av multiplikation. Kommutativa lagen följer exempelvis från

$$(a,b)(c,d) = (ac-bd, ad+bc) = (ca-db, cb+da) = (c,d)(a,b).$$

Att $(1,0)$ är etta för multiplikation, följer från räkningen

$$(a,b)(1,0) = (a\cdot 1 - b\cdot 0,\ a\cdot 0 + b\cdot 1) = (a,b).$$

För de multiplikativa inverserna kan vi kontrollera, att

$$\begin{aligned}(a,b)\cdot\left(\frac{a}{a^2+b^2},\ -\frac{b}{a^2+b^2}\right)&\\ = \left(a\cdot\frac{a}{a^2+b^2}+b\cdot\frac{b}{a^2+b^2},\ -a\cdot\frac{b}{a^2+b^2}+b\cdot\frac{a}{a^2+b^2}\right)&\\ = \left(\frac{a^2+b^2}{a^2+b^2},\ 0\right) = (1,0).&\end{aligned}$$

För att visa entydighet för inversen, låt oss pröva om det kan finnas fler tal y med egenskapen $xy = (1,0)$. Multiplicera båda led med x^{-1} (som vi redan vet finns):

$$x^{-1}(xy) = x^{-1}(1,0).$$

Högerledet förenklas till x^{-1} enligt lagen om ettan, medan vänsterledet förenklas till $(x^{-1}x)y = (1,0)y = y$ enligt Associativa lagen. Alltså är $y = x^{-1}$. Den multiplikativa inversen är unik.

Precis som för reella tal, låter oss de multiplikativa inverserna definiera division, enligt formeln $\frac{x}{y} = xy^{-1}$.

Sammanfattningsvis uppfyller de komplexa talen *exakt samma räknelagar*, som gäller för de reella talen. För de reella talen togs de som axiom, för de komplexa bevisades de som satser. Alla

de räknelagar vi kunde bevisa i kapitel 1 med utgångspunkt i axiomen *måste nu nödvändigt äga beständighet även för komplexa tal*. Det gäller Kvadrerings- och Konjugatreglerna, reglerna för bråkräkning, och även formlerna för aritmetiska och geometriska summor. Alla dessa bevisades ju ur axiomen, de fundamentala räknelagarna, vilka gäller med samma styrka, som vi precis bevisat, för de komplexa talen. Bevisen kan därför bara upprepas, med den enda skillnaden, att bokstäverna (variablerna) nu tillåts beteckna komplexa tal.

Däremot gäller inte ordningsaxiomen för komplexa tal. De reella talen bildar en linje, de komplexa ett plan. Det finns inget naturligt eller bra sätt att ordna dem på.

De reella talen ligger inneslutna i de komplexa talen, under pseudonym. Det komplexa talet $(a, 0)$ beter sig nämligen som det vanliga reella talet a under båda räkneoperationerna.

Sats 7.5

Komplexa tal på formen $(a, 0)$ beter sig som reella tal, i meningen att

$$(a, 0) + (b, 0) = (a + b, 0) \qquad \text{och} \qquad (a, 0)(b, 0) = (ab, 0).$$

Bevis

Omedelbar konsekvens av definitionerna.

Vi kommer nu överens om, att det reella talet a och det komplexa talet $(a, 0)$ bara är två beteckningar för samma tal. De reella talen är därmed innefattade i de komplexa. Den algebraiska strukturen är densamma, för enligt satsen beter ju sig addition och multiplikation likadant, varför detta inte vållar några problem i räkningar. (Detta sker med samma självklara moraliska rätt, som vi låter heltalen vara en del av de rationella talen. Då handlar det om att identifiera heltalet m med bråket $\frac{m}{1}$.)

Nu kommer ett avgörande steg. Ett godtyckligt komplext tal $z = (a, b)$ kan skrivas som

$$z = (a, b) = (a, 0) + (0, b) = (a, 0) + (b, 0)(0, 1) = a + b(0, 1),$$

där vi i sista likheten identifierat $a = (a, 0)$ och $b = (b, 0)$. (Kontrollera med hjälp av definitionen, att $(b, 0)(0, 1) = (0, b)$.)

Definition 7.6

Den **imaginära enheten** i är det komplexa talet $(0, 1)$.

Av räkningen ovan följer nu:

Sats 7.7

Varje komplext tal kan skrivas som

$$z = (a, b) = a + bi.$$

Detta är det vanliga sättet att skriva komplexa tal. Under den moderniserade beteckningen förvandlas definitionerna av addition och multiplikation till de litet mer välbekanta

$$(a + bi) + (c + di) = (a + c) + (b + d)i \quad \text{och} \quad (a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

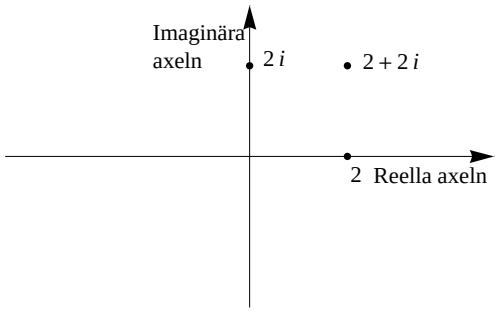
Vad är då så speciellt med talet i? Enligt definitionen på multiplikation är

$$i^2 = i \cdot i = (0, 1) \cdot (0, 1) = (-1, 0) = -1.$$

Vi har alltså skapat ett nytt (komplext) tal, vars kvadrat är -1! Men denna gång hittade vi inte bara på det, utan vi konstruerade det rigoröst.

Låt oss nu betrakta additionen

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



FIGUR 1: Komplexa talplanet.

litet närmare. Vid närmare påseende sker ju intet märkligt eller övernaturligt här. Det är helt vanlig bokstavsräkning, med i som ett slags variabel (liksom x eller y). Vid närmare eftertanke gäller detta även för multiplikationen, ty multipliceras $a+bi$ och $c+di$ enligt distributiva lagarna, fås

$$(a+bi)(c+di) = ac + adi + bci + bdi^2 = (ac-bd) + (ad+bc)i,$$

där vi utnyttjat identiteten $i^2 = -1$.

Den fundamentala identiteten $i^2 = -1$ är därför den enda man faktiskt behöver belasta minnet med, vad komplex aritmetik anbelangar (och någon övermänsklig belastning lär den ju knappast vara). Vi har därmed rättfärdigat vårt diktum ovan, att man räknar med komplexa tal som med polynom och ersätter i^2 med -1, när den behagar dyka upp.

Ovan anmärkte vi, att de komplexa talen lyder precis samma räkneregler som de reella. Det betyder bland annat, att bråkräkning löper på som vanligt. Förlängning och förkortning fungerar precis som de skall. Därmed har vi rättfärdigat principen för division av komplexa tal, att förlänga med konjugatet till nämnaren. Sats 7.4 ovan gav visserligen en explicit formel för inversen

$$(a+bi)^{-1} = (a,b)^{-1} = \left(\frac{a}{a^2+b^2},\ -\frac{b}{a^2+b^2}\right) = \frac{a}{a^2+b^2} - \frac{bi}{a^2+b^2},$$

men denna gör man klokt i, att inte lära sig utantill. Det är precis samma sak vi skulle få, om vi förlängde $\frac{1}{a+bi}$ med $a-bi$:

$$\frac{1}{a+bi} = \frac{a-bi}{(a+bi)(a-bi)} = \frac{a-bi}{a^2-b^2i^2} = \frac{a-bi}{a^2+b^2}.$$

Nämnaren är ett reellt tal. Rent konkret kan alltså division med $a+bi$ utföras genom förlängning med $a-bi$.

§4. KOMPLEXA TALPLANET

Varje komplext tal $a+bi = (a,b)$ svarar mot en punkt i planet (givet ett rätvinkligt koordinatsystem). Se Figur 1. I detta sammanhang, då vi ser planet som mängden $\mathbb{C}$ av komplexa tal, så går planet under benämningen det **komplexa talplanet**.

Horisontella axeln består av alla komplexa tal av formen $(a,0) = a$, d.v.s. de reella talen. Den kallas därför **reella axeln**. Vertikala axeln består av alla komplexa tal på formen $(0,b) = bi$, de **rent imaginära** talen. Denna axel kallas den **imaginära axeln**.

Definition 7.8

Låt $z = a+bi$ vara ett komplext tal. Dess **realdel** är $\operatorname{Re} z = a$ och dess **imaginärdel** är $\operatorname{Im} z = b$.

Observera särskilt, att imaginärdelen $\operatorname{Im} z = b$, namnet till trots, är ett *reellt* tal (den definieras *inte* som bi).

Realdelen och imaginärdelen anger alltså koordinaterna i det komplexa talplanet. För att få koordinaterna för en punkt i ett rätvinkligt koordinatsystem, skall punkten projiceras på axlarna. Då erhålles en rektangel. Formen $a + bi$ kallas därför det komplexa talets **rektangulära form**.

Exempel 5.
Den imaginära axeln kan beskrivas med ekvationen $\operatorname{Re} z = 0$, medan den reella axeln har ekvationen $\operatorname{Im} z = 0$.

Exempel 6.
Vilka komplexa tal $z = a + bi$ uppfyller ekvationen

$$\operatorname{Re}(1 + i)z = 0 \ ?$$

De sökta talen beskrivs algebraiskt av

$$0 = \operatorname{Re}(1 + i)z = \operatorname{Re}(1 + i)(a + bi) = \operatorname{Re}(a - b) + (a + b)i = a - b.$$

Geometriskt betyder $a = b$ diagonalen i det komplexa talplanet.

§5. Konjugat

Division med det komplexa talet $a + bi$ utförs, som vi såg ovan, genom förlängning med $a - bi$. Genom multiplikation med konjugatet kan vi alltså få ett läbbigt komplext tal att bli slätslickat reellt. Detta begrepp är så viktigt, att det förtjänar ett eget namn.

Definition 7.9

Låt $z = a + bi$ vara ett komplext tal. Dess **konjugat** är

$$\overline{z} = a - bi.$$

Sats 7.10

Låt z och w vara komplexa tal. Då gäller följande.

(a) $\overline{z + w} = \overline{z} + \overline{w}$

(b) $\overline{z - w} = \overline{z} - \overline{w}$.

(c) $\overline{zw} = \overline{z}\,\overline{w}$.

(d) $\overline{\left(\frac{z}{w}\right)} = \frac{\overline{z}}{\overline{w}}$.

(e) $\overline{(\overline{z})} = z$.

Bevis

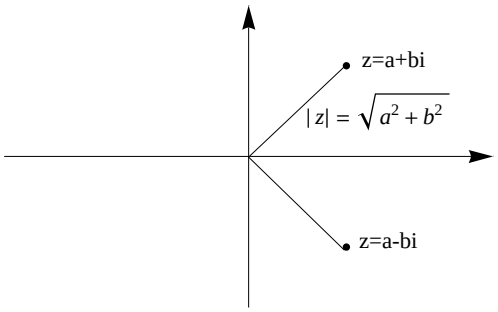
Egenskaperna (a), (b) och (e) är uppenbara. Vi visar (c) och (d).

(c) Låt $z = a + bi$ och $w = c + di$. Då ger en enkel kalkyl

$$\begin{aligned}\overline{zw} = (a - bi)(c - di) = (ac - bd) - (ad + bc)i \\ = \overline{(ac - bd) + (ad + bc)i} = \overline{(a + bi)(c + di)} = \overline{z}\,\overline{w}.\end{aligned}$$

(d) Konjugera ekvationen $z = \frac{z}{w} \cdot w$ till

$$\overline{z} = \overline{\frac{z}{w} \cdot w} = \overline{\left(\frac{z}{w}\right)} \cdot \overline{w}$$



FIGUR 2: Konjugat och absolutbelopp.

enligt (c). Dividera båda led med $\overline{w}$, och den önskade likheten faller ut.

Exempel 7.
Vi har att $z = \overline{z}$ om och endast om z är ett reellt tal, ty

$$a + bi = a - bi \quad \Leftrightarrow \quad 2bi = 0 \quad \Leftrightarrow \quad b = 0.$$

Följande samband mellan konjugat och real- respektive imaginärdel är ofta användbart.

Sats 7.11

Om z är ett komplext tal så är

$$\operatorname{Re} z = \frac{z + \overline{z}}{2} \qquad \text{och} \qquad \operatorname{Im} z = \frac{z - \overline{z}}{2i}.$$

Bevis

Låt $z = a + bi$. Då är

$$z + \overline{z} = (a + bi) + (a - bi) = 2a = 2 \operatorname{Re} z.$$

Den andra likheten följer på samma sätt ur

$$z - \overline{z} = (a + bi) - (a - bi) = 2bi = 2i \operatorname{Im} z.$$

Den viktigaste egenskapen hos konjugatet, som vi utnyttjade vid division, är följande.

Sats 7.12

Låt $z = a + bi$. Det gäller att

$$z\overline{z} = a^2 + b^2 \geq 0$$

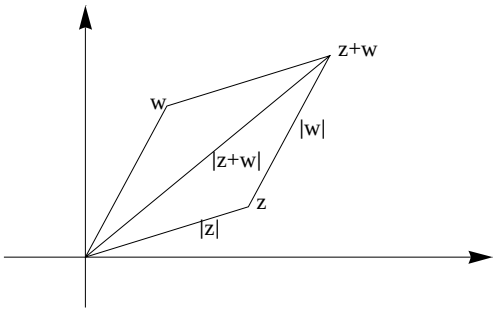
är ett icke-negativt reellt tal. Likhet $z\overline{z} = 0$ råder precis när $z = 0$.

Bevis

Vi har

$$z\overline{z} = (a + bi)(a - bi) = a^2 + b^2.$$

En summa av reella kvadrater är uppenbarligen reell och icke-negativ. Den kan bara vara noll om bägge kvadraterna är noll, d.v.s. $a = b = 0$.



Figur 3: Triangelolikheten.

§6. Absolutbelopp

Som vi kan se i Figur 2, svarar konjugering geometriskt mot spegling i den reella talaxeln. Vi skall nu införa absolutbeloppet, som också har en direkt geometrisk tolkning. Det mäter avståndet från origo till $z = a + bi$ i komplexa talplanet (eller längden av motsvarande vektor). Se Figur 2. Pythagoras sats beräknar detta avstånd till $\sqrt{a^2 + b^2}$.

Definition 7.13

Absolutbeloppet av det komplexa talet $z = a + bi$ är

$$|z| = \sqrt{a^2 + b^2} = \sqrt{z\overline{z}}.$$

Sats 7.14

Låt z och w vara komplexa tal. Då gäller följande.

(a) $|zw| = |z||w|$.

(b) $\left|\frac{z}{w}\right| = \frac{|z|}{|w|}$.

Bevis

Dessa egenskaper följer av de algebraiska egenskaperna för konjugatet. Vi nöjer oss med att verifiera (a). Den följer direkt ur definitionen:

$$|zw| = \sqrt{(zw)\overline{(zw)}} = \sqrt{(z\overline{z})(w\overline{w})} = \sqrt{z\overline{z}}\sqrt{w\overline{w}} = |z||w|.$$

För absolutbeloppet av en summa finns ingen enkel formel, däremot följande viktiga olikhet.

Sats 7.15: Triangelolikheten

Låt z och w vara komplexa tal. Då gäller olikheten

$$|z + w| \leq |z| + |w|.$$

Bevis

Detta inses geometriskt av Figur 3. Talen 0, z, w, $z + w$ bildar en parallellogram. Halva denna är en triangel med sidlängder $|z|, |w|, |z+w|$. Triangelolikheten följer nu omedelbart av en välkänd sats ur den euklidiska geometrien, att två sidor i en triangel tillsammans är längre än den tredje.

Tror vi inte omedelbart på detta, kan vi argumentera för det. (Som Groucho Marx

påstås ha yttrat (belägg tycks saknas): "These are my principles. If you don't like them, I have others.") Den långa sidan i triangeln, diagonalen i parallellogrammen, har längd $|z+w|$. Men denna beskriver ju den kortaste vägen från 0 till punkten $z+w$ och måste alltså vara kortare än summan av de två andra sidornas längder $|z|+|w|$, som är längden av omvägen över z.

Exempel 8.
Avståndet mellan de komplexa talen u och z är $|u-z|$. Sätt nämligen $w = u - z$. Ur Figur 3 framgår, att det sökta avståndet mellan $u = z + w$ och z är $|w| = |u - z|$.

Geometriskt är livet med absolutbelopp bara värt att leva i $\mathbb{C}$.

Exempel 9.
Låt oss bestämma vilka z, som uppfyller $|z+1+i| = 2$. Avståndet mellan två komplexa tal ges av absolutbeloppet av deras skillnad (föregående exempel). Ekvationen

$$2 = |z+1+i| = |z-(-1-i)|$$

uttrycker villkoret, att avståndet från z till $-1-i$ må vara precis 2. Svaret är alltså en *cirkel* i komplexa talplanet med radie 2 och centrum $-1-i$.

En variation:

Exempel 10.
Ekvationen $|13z+1+i| = 2$ kan skrivas om till

$$2 = |13z+1+i| = \left|13\left(z - \frac{-1-i}{13}\right)\right| = 13\left|z - \frac{-1-i}{13}\right|,$$

sålunda $\left|z - \frac{(-1-i)}{13}\right| = \frac{2}{13}$, som uttrycker villkoret, att z ligger på en cirkel med radie $\frac{2}{13}$ och centrum $-\frac{1}{13} - \frac{i}{13}$.

Här kommer ett svårare exempel.

Exempel 11.
Vi studerar ekvationen

$$\left|\frac{z+1}{z+i}\right| = \sqrt{2}.$$

Omskrivning leder till ekvationen

$$|z+1| = \sqrt{2}|z+i|.$$

Eftersom detta är positiva tal, är denna ekvation ekvivalent med

$$|z+1|^2 = 2|z+i|^2.$$

Sätter vi $z = x + yi$, så får vi

$$(x+1)^2 + y^2 = 2(x^2 + (y+1)^2) \quad \Leftrightarrow \quad x^2 - 2x + y^2 + 4y + 1 = 0.$$

Denna ekvation kvadratkompletteras till

$$4 = (x^2 - 2x + 1) + (y^2 + 4y + 4) = (x-1)^2 + (y+2)^2.$$

Men det sista uttrycket är precis $|z-(1-2i)|^2$, så ekvationen svarar mot cirkeln $|z-(1-2i)| = 2$ med radie 2 och centrum $1-2i$.

Ibland tröttnar man på cirklar och vill ha linjer.

Exempel 12.
Vi tolkar geometriskt ekvationen

$$|z+1| = |z+i|.$$

Eftersom detta är positiva tal, är ekvationen ekvivalent med

$$|z+1|^2 = |z+i|^2.$$

Ansätter vi $z = x + iy$, får vi

$$(x+1)^2 + y^2 = x^2 + (y+1)^2 \quad \Leftrightarrow \quad x = y.$$

Svaret är alltså linjen $\operatorname{Re} z = x = y = \operatorname{Im} z$.

§7. Komplexa andragradsekvationer

Nu skall vi se hur livet, i princip, blir enklare med komplexa tal, i alla fall så länge man slipper räkna med dem, utan kan hålla sig på ett teoretiskt plan. Som bekant kan vissa andragradsekvationer sakna reella rötter. Ekvationen $z^2 + 1 = 0$ har inte några reella rötter; däremot så finns det ju en komplex rot $z = i$, till och med två stycken, eftersom $z = -i$ också är en rot.

När vi löser ekvationer inom de komplexa talen, så kan det förstås rimligtvis finnas fler lösningar, eftersom vi nu arbetar i en större talmängd. Men det som är förvånande, det är att det *alltid finns lösningar*. Vi kan rentav tillåta ekvationens koefficienter att själva vara komplexa tal. Genom att på halvt taskspelarvis lägga till en rot (i) till en enda ekvation ($z^2 + 1 = 0$), har vi alltså som någon sorts magisk bonus fått rötter till *alla* möjliga andragradsekvationer. (Och magien slutar inte där. Faktum är, att *alla algebraiska ekvationer kan lösas fullständigt* inom de komplexa talen. Se Algebrans fundamentalsats.)

Vi visar först exempel på andragradsekvationer med reella koefficienter, men irreella rötter.

Exempel 13.
Andragradsekvationen $z^2 = -3$ saknar förstås reella rötter, men vi kan finna de irreella rötterna genom ansatsen $z = x + yi$. Vi får

$$-3 = (x+yi)^2 = x^2 - y^2 + 2xyi.$$

Imaginärdelen av högerledet är $2xy$, medan imaginärdelen av vänsterledet är noll. Således måste någon av x eller y vara noll. Vi vet redan att reella lösningar saknas, alltså måste det vara $x = 0$. Härav $-3 = -y^2$, så att $y = \pm\sqrt{3}$. Lösningarna är $z = \pm\sqrt{3}i$. Observera, att detta är samma sak som vi skulle ha fått, om vi naivt hade "dragit kvadratroten" ur båda leden i $z^2 = -3$.

Exempel 14.
Vi vill lösa ekvationen

$$z^2 + 4z + 7 = 0.$$

Via kvadratkomplettering ser vi att

$$0 = z^2 + 4z + 7 = (z+2)^2 + 3,$$

så att ekvationen är ekvivalent med $(z+2)^2 = -3$. Från föregående exempel vet vi då, att $z + 2 = \pm\sqrt{3}i$. Ekvationens lösningar är därför $z = -2 \pm \sqrt{3}i$. Observera, att detta är samma sak som vi skulle ha fått, om vi naivt hade applicerat pq-formeln:

$$z = -\frac{4}{2} \pm \sqrt{\left(\frac{4}{2}\right)^2 - 7} = -2 \pm \sqrt{-3}.$$

Vi visar nu den allmänna metoden för lösning av komplexa andragradsekvationer.

Exempel 15.
Vi löser

$$z^2 + 4iz - (7 + 4i) = 0.$$

Proceduren ovan för att lösa reella andragradsekvationer innefattade två steg: först kvadratkomplettering och sedan rotutdragning. Dessa två steg kan vi fortfarande använda, med den enda skillnaden, att rotutdragningen är mer komplicerad.

Först gör vi kvadratkompletteringen:

$$0 = z^2 + 4iz - (7 + 4i) = z^2 + 4iz + (2i)^2 - (3 + 4i) = (z + 2i)^2 - (3 + 4i).$$

Variabelbytet $w = z + 2i$ transformerar ekvationen till den enkla formen

$$w^2 = 3 + 4i. \tag{1}$$

Precis som i lösningen av den reella andragradsekvationen återstår nu bara en rotutdragning. Emellertid har vi nu ingen rotfunktion att ta till, utan nödgas göra detta för hand. Ansätt $w = x + yi$. Substitution av detta i (1) ger

$$3 + 4i = w^2 = (x + yi)^2 = x^2 - y^2 + 2xyi,$$

varpå identifikation av real- och imaginärdelar leder till ekvationssystemet

$$\begin{cases} x^2 - y^2 = 3 \\ 2xy = 4. \end{cases} \tag{2}$$

Vi kan få en ekvation till genom att ta absolutbeloppet av bägge sidor i (1):

$$x^2 + y^2 = |w|^2 = |w^2| = |3 + 4i| = 5. \tag{3}$$

Nu har vi alltså hela tre ekvationer, som ger egenskaper hos futtiga två obekanta, och det vore väl skrutt om vi inte kunde bestämma x och y utifrån dessa! Hur skall vi då göra? Om vi lägger ihop översta ekvationen i (2) med (3), så blir vi av med y:

$$2x^2 = (x^2 - y^2) + (x^2 + y^2) = 3 + 5 = 8 \quad \Leftrightarrow \quad x = \pm 2.$$

Tar vi istället skillnaden så blir vi av med x:

$$2y^2 = -(x^2 - y^2) + (x^2 + y^2) = -3 + 5 = 2 \quad \Leftrightarrow \quad y = \pm 1.$$

Detta ger oss skenbart fyra lösningar, men bara $w = x + yi = 2 + i$ och $w = x + yi = -2 - i$ uppfyller villkoret $2xy = 4$. Dessa är alltså lösningarna till (2)–(3) och därmed till (1).

Slutligen får vi från $z = w - 2i$ lösningarna $z = 2 - i$ och $z = -2 - 3i$ till vår ursprungliga ekvation.

Vi bör inte frestas "lösa" ekvationen $w^2 = 3 + 4i$ genom att skriva

$$w = \pm\sqrt{3 + 4i}.$$

För det första ger ju detta inte så särdeles med information. Rimligen vill vi ha svaret på formen $a + bi$. För det andra går det nämligen inte att definiera en entydig kvadratrot ur irreella tal. Det visar sig vara omöjligt till och med för negativa tal.

En del människor skriver gärna $i = \sqrt{-1}$. Det gör inte vi, utan håller detta skrivsätt för direkt olämpligt. Eftersom både $i^2 = -1$ och $(-i)^2 = -1$, kan ju både i och $-i$ ses som kvadratrötter till -1. Vi kunde förstås bestämma oss för den ena och högtidligt deklarera $i = \sqrt{-1}$, men vi kan inte samtidigt förvänta oss, att dessa "kvadratrötter" snällt skall uppfylla samma räkneregler som vi är vana vid. Mycket riktigt leder det till absurda konsekvenser som

$$-1 = \sqrt{-1}\sqrt{-1} = \sqrt{(-1)\cdot(-1)} = \sqrt{1} = 1.$$

Alltså uppfyller våra "kvadratrötter" inte längre räknelagen $\sqrt{a}\sqrt{b} = \sqrt{ab}$, och man kan undra, vad det då är för vits med att använda kvadratrotssymbolen.

När vi definierar kvadratroten ur ett (reellt) icke-negativt tal $a \geq 0$, är situationen annorlunda. Det finns visserligen en positiv och en negativ lösning till $x^2 = a$, men det vållar inga svårigheter, att utvälja den ena av dem, *nämligen den positiva.* Denna kallar vi så för $\sqrt{a}$. Den andra lösningen är då $-\sqrt{a}$, och bägge lösningarna kan beskrivas som $\pm\sqrt{a}$. Vi använder, i denna definition, *ordningen* på de reella talen, då vi kräver $\sqrt{a} \geq 0$. För komplexa tal finns det ingen liknande naturlig ordning, och det finns inget skäl att anse i vara en "bättre" kvadratrot än $-i$ till -1.

Vi påstår nu, att proceduren i exemplet alltid fungerar.

Sats 7.16

Andragradsekvationen

$$z^2 + pz + q = 0$$

har precis två komplexa lösningar för godtyckliga komplexa koefficienter p och q.

Bevis

Låt oss sammanfatta lösningen i exemplet nyss, i mer abstrakta termer. Först skrev vi, via kvadratkomplettering, om ekvationen som

$$0 = z^2 + pz + q = \left(z + \frac{p}{2}\right)^2 + \left(q - \frac{p^2}{4}\right),$$

vilken är på den enklare formen $w^2 = a + bi$. Det gäller alltså att visa, att denna ekvation alltid har precis två lösningar.

Vi ansätter $w = x + yi$, vilket ger

$$a + bi = w^2 = (x + yi)^2 = x^2 - y^2 + 2xyi,$$

varpå identifikation av real- och imaginärdelar, kombinerat med

$$x^2 + y^2 = |w|^2 = |w^2| = |a + bi| = \sqrt{a^2 + b^2},$$

leder till ekvationssystemet

$$\begin{cases} x^2 - y^2 = a \\ 2xy = b \\ x^2 + y^2 = \sqrt{a^2 + b^2}. \end{cases}$$

Från den första och tredje ekvationen finner vi

$$x^2 = \frac{1}{2}\left(\sqrt{a^2 + b^2} + a\right), \qquad y^2 = \frac{1}{2}\left(\sqrt{a^2 + b^2} - a\right).$$

Detta är två icke-negativa tal, ty

$$\sqrt{a^2 + b^2} \geq \sqrt{a^2} = |a| \geq \pm a.$$

Alltså kan vi utdraga kvadratrötter och få

$$x = \pm\sqrt{\frac{1}{2}\left(\sqrt{a^2 + b^2} + a\right)}, \qquad y = \pm\sqrt{\frac{1}{2}\left(\sqrt{a^2 + b^2} - a\right)}.$$

Det ger fyra möjligheter för x och y. Exakt två av dessa har rätt tecken för att även uppfylla ekvationen $2xy = b$, vilket betyder att vår andragradsekvation har exakt två lösningar.

I nästa kapitel skall vi ge ett annat bevis med hjälp av den polära formen.

Övningar

7.1. Beräkna

(a) $(2+3i)+(1+2i)$;
(b) $(2+3i)+(1-2i)$;
(c) $(2+3i)-(1-2i)$;
(d) $(2+3i)\cdot(1+2i)$;
(e) $(2+3i)^2$;
(f) $(5+i)^3$;
(g) $(1-i)^4$.

7.2. Beräkna

(a) $\frac{1}{1+i}$;
(b) $\frac{1}{2-5i}$;
(c) $\frac{1+i}{1-i}$;
(d) $\frac{3+4i}{1-5i}$;
(e) $\frac{1}{i}$;
(f) $(1+i)^{-2}$.

7.3. Bestäm $\operatorname{Re} z$ och $\operatorname{Im} z$, om z är följande tal:

(a) $2+3i$;
(b) $2-3i$;
(c) 2;
(d) $5i$;
(e) $-i$.

7.4. Beräkna

(a) $\overline{2+3i}$;
(b) $\overline{1-3i}$;
(c) $\overline{3i}$;
(d) $\overline{(2+3i)}(2+3i)$;
(e) $|3+4i|$;
(f) $|-3-4i|$;
(g) $|4i|$;
(h) $|-7i|$;
(i) $|i|$.

7.5. Beräkna absolutbeloppet av

(a) $(1+3i)(4-5i)(1+i)$;
(b) $\frac{(1+2i)(1+\sqrt{3}i)}{(1+i)^3}$.

7.6. Lös ekvationerna

(a) $z+(1+i)\overline{z}=1-i$;

(b) $2z + i\overline{z} = 3 + 3i$;

(c) $\overline{z} \cdot 2z = 1 + i$.

7.7. Tolka geometriskt i komplexa talplanet ekvationerna

(a) $\operatorname{Re} z = 2$;

(b) $\operatorname{Im} z = 2$;

(c) $\operatorname{Im} z \geq 0$;

(d) $\operatorname{Re} z + \operatorname{Im} z = 2$;

(e) $\overline{z} + z = 0$;

(f) $\bar{z} = z$.

7.8. Lös ekvationerna

(a) $z^2 + (3 + 4i)z - 1 + 5i = 0$;

(b) $z^2 + (1 + 4i)z - 3 + 3i = 0$;

(c) $z^2 + (2 - 2i)z + 4 - 2i = 0$.

7.9. Sök de komplexa tal z, för vilka talet

$$z + \frac{1}{z}$$

är reellt. Sök också de z, för vilka talet är rent imaginärt.

7.10. De komplexa talen z och p är sådana, att z är rent imaginärt och $z \neq p$. Visa, att

$$\left| \frac{z - p}{z + \overline{p}} \right| = 1.$$

Kapitel 8

Komplexa tal på polär form

Komplexa tal är punkter i det komplexa talplanet. Punkter i ett plan kan beskrivas, inte bara med rätvinkliga koordinatsystem, utan också i termer av vinklar och avstånd till origo, s.k. *polära koordinater*. Detta är intimt förknippat med multiplikation av komplexa tal, ett av naturens stora under… tillsammans med SVT och DNA (som för övrigt är högervridet).

§1. Polära koordinater

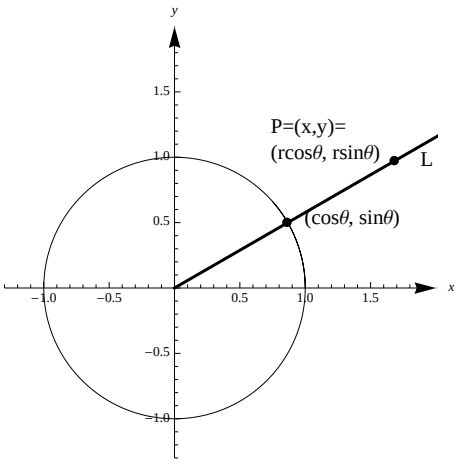
I Figur 1 betraktar vi en stråle L från origo, bildande en vinkel $0 \leq \theta < 2\pi$ med positiva x-axeln, räknat i radianer moturs. Denna skär enhetscirkeln $x^2+y^2 = 1$ i en punkt, som definitionsmässigt har koordinaterna $(\cos\theta, \sin\theta)$ (definitionen av cosinus och sinus). Detta kan utvidgas till vinklar $\theta \geq 2\pi$, som då betyder att vi snurrat ett antal varv moturs kring origo, och även negativa vinklar, som betyder att vi snurrat medurs.

Betrakta nu en annan punkt P på denna stråle, situerad på avståndet r från origo. Dess **rektangulära koordinater** är

$$P = (x, y) = (r\cos\theta, r\sin\theta).$$

Det är klart, att då strålen L snurrar ett varv kring origo, så sveper den samtidigt över hela planet. Varje punkt i planet kan därför beskrivas på denna form. Talparet (r, θ) kallas de **polära koordinaterna** för punkten P.

Exempel 1.
Positiva x-axeln är mängden av punkter med polära koordinater $(r, 0)$, där $r \geq 0$. Negativa x-axeln har polära koordinater (r, π), där $r \geq 0$.



Figur 1: Polära koordinater.

Den **radiella koordinaten** $r \geq 0$ är entydigt given, ty denna mäter avståndet mellan P och origo. Vinkeln θ är däremot inte entydig. Vi kan alltid snurra strålen L ett antal varv framåt eller bakåt och landa i samma position. Den **azimutala koordinaten** θ är bara bestämd upp till addition av en heltalsmultipel av 2π. Det finns således oändligt många uppsättningar polära koordinater för samma punkt, såvida vi inte lägger på kravet $0 \leq \theta < 2\pi$.

Exempel 2.
Punkten $(1, 0)$ i rektangulära koordinater har de polära koordinaterna

$$\ldots,\ (1, -2\pi),\ (1, 0),\ (1, 2\pi),\ (1, 4\pi),\ \ldots$$

alla lika giltiga.

I origo går det riktigt snett. Här är den radiella koordinaten $r = 0$, men den azimutala koordinaten θ är fullkomligt obestämd och helt valfri. Dessa märkligheter vållar dock i praktiken sällan bekymmer.

Vi har formlerna

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

för koordinatbyte. Av detta följer

$$r = \sqrt{(r\cos\theta)^2 + (r\sin\theta)^2} = \sqrt{x^2 + y^2}.$$

Någon explicit formel, som uttrycker θ i de rektangulära koordinaterna x och y, saknas däremot. (Det *går* att vaska fram, men formeln blir sönderstyckad i en massa fall och rent otroligt klafsig.)

Exempel 3.
Punkten P med de rektangulära koordinaterna

$$P = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

har den radiella koordinaten

$$r = \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2} = 1.$$

Vi kan lätt hitta den azimutala koordinaten. Ritar vi en figur är det nämligen uppenbart, att vinkeln mot positiva x-axeln är $\theta = 45^\circ = \frac{\pi}{4}$. Eller så kan vi använda våra trigonometriska kunskaper och direkt skriva

$$P = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \left(\cos\frac{\pi}{4}, \sin\frac{\pi}{4}\right).$$

Alla möjliga azimutala koordinater för P är på formen $\theta = \frac{\pi}{4} + 2n\pi$, där $n \in \mathbb{Z}$ (är ett heltal), alltså

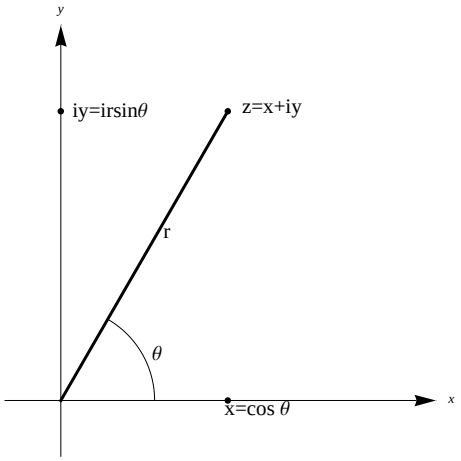
$$\theta = \ldots,\ -\frac{7\pi}{4},\ \frac{\pi}{4},\ \frac{9\pi}{4},\ \frac{17\pi}{4},\ \ldots.$$

§2. Polär representation av komplexa tal

Eftersom komplexa tal $z = x + iy$ svarar mot punkter (x, y) i det komplexa talplanet, kan vi beskriva dem med polära koordinater. Se Figur 2. Om $(x, y) = (r\cos\theta, r\sin\theta)$, så är

$$z = x + iy = r\cos\theta + ir\sin\theta = r(\cos\theta + i\sin\theta).$$

Radiella koordinaten $r = \sqrt{x^2 + y^2}$ känner vi igen som absolutbeloppet $|z|$.



FIGUR 2: Polär form.

Definition 8.1

Azimutala koordinaten θ kallas **argumentet** av z och betecknas med $\arg z$.

Argumentet är, som vi såg i föregående avsnitt, inte entydigt bestämt, men två olika argument för samma tal skiljer sig åt med en heltalsmultipel av 2π.

De komplexa talen $\cos\theta + i\sin\theta$ ligger på enhetscirkeln. När θ varierar mellan 0 och 2π, genomlöper dessa hela enhetscirkeln, och vi introducerar följande exponentiella skrivsätt. Motiveringen — i form av potenslagar — kommer sedan.

Definition 8.2

Den **komplexa exponentialfunktionen** definieras av

$$e^{\theta i} = \cos\theta + i\sin\theta \qquad \text{för reellt } \theta. \tag{1}$$

Därmed kan vi skriva ett godtyckligt komplext tal som

$$z = x + iy = r\cos\theta + ir\sin\theta = r(\cos\theta + i\sin\theta) = re^{\theta i},$$

där det sistnämnda är den **polära formen**, i motsats till den rektangulära formen $x + yi$, som behandlades i förra kapitlet.

Exempel 4.
Vi önskar uttrycka $z = 1 + i$ och $w = -1 - i$ på polär form. Se Figur 3. De två talen har samma absolutbelopp

$$|z| = |w| = \sqrt{1^2 + 1^2} = \sqrt{2}.$$

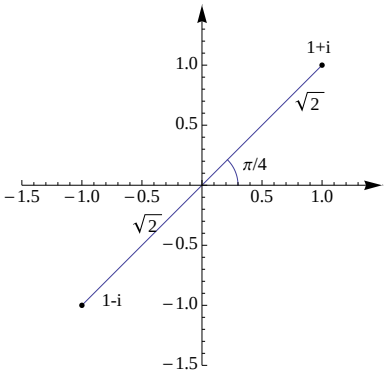
Vidare ser vi från figuren, att $\arg z = \frac{\pi}{4}$ och $\arg w = \frac{5\pi}{4}$. Alltså är

$$1 + i = \sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right) = \sqrt{2}e^{\frac{\pi}{4}i}$$
$$-1 - i = \sqrt{2}\left(\cos\frac{5\pi}{4} + i\sin\frac{5\pi}{4}\right) = \sqrt{2}e^{\frac{5\pi}{4}i}.$$

I rektangulära koordinater (x, y) beskriver olikheterna

$$a \le x \le b, \qquad c \le y \le d$$

en *rektangel*. I polära koordinater kan vi, på motsvarande sätt, enkelt beskriva cirkulära kurvor eller områden:



FIGUR 3: Talen $\pm(1+i)$ i komplexa talplanet.

Exempel 5.
Ekvationen $|z| = 3$ beskriver i komplexa talplanet cirkeln med centrum origo och radie 3. Vi kan också säga, att den består av alla punkter på formen $3e^{\theta i}$. Cirkelskivan (cirkeln samman med dess innanmäte) med samma centrum och radie tecknas av olikheten $|z| \leq 3$, alternativt ges av alla punkter $re^{\theta i}$, där $0 \leq r \leq 3$. Första kvadranten skär ut en cirkelsektor av denna skiva, tecknad av olikheterna

$$0 \leq |z| \leq 3, \qquad 0 \leq \arg z \leq \frac{\pi}{2}.$$

§3. Den komplexa exponentialfunktionen

Låt $z = e^{\theta i}$ vara ett komplext tal på enhetscirkeln. Tillämpar vi definitionen av $e^{-\theta i}$ får vi

$$e^{-\theta i} = \cos(-\theta) + i\sin(-\theta) = \cos\theta - i\sin\theta = \overline{z}, \tag{2}$$

alltså konjugatet. Detta kommer att utnyttjas flera gånger framöver.

Sats 8.3

För reella argument α, β har vi

(a) $e^{\alpha i}e^{\beta i} = e^{(\alpha+\beta)i}$;

(b) $\dfrac{e^{\alpha i}}{e^{\beta i}} = e^{(\alpha-\beta)i}$.

Bevis

(a) Vänsterledet är

$$e^{\alpha i}e^{\beta i} = (\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta) = \\ = (\cos\alpha\cos\beta - \sin\alpha\sin\beta) + i(\cos\alpha\sin\beta + \sin\alpha\cos\beta),$$

medan högerledet är

$$e^{(\alpha+\beta)i} = \cos(\alpha+\beta) + i\sin(\alpha+\beta).$$

Additionsformlerna för cosinus och sinus har precis innebörden, att realdelarna respektive imaginärdelarna är lika.

(b) Från (a) har vi $e^{(\alpha-\beta)i}e^{\beta i} = e^{\alpha i}$; dividera denna med $e^{\beta i}$.

Beviset för denna sats är för resten en god hjälp för att memorera de trigonometriska additionsformlerna. Notera hur de kan rekonstrueras genom att utgå från potenslagen $e^{\alpha i}e^{\beta i} = e^{(\alpha+\beta)i}$ och definitionen av $e^{\theta i}$, och sedan marschera baklänges genom beviset.

Exempel 6.
Inversen z^{-1} till ett komplext tal z på enhetscirkeln är lika med konjugatet $\overline{z}$. Detta kan visas på mångahanda sätt.

Ett sätt går via den polära formen. Låt $z = e^{\theta i} = \cos\theta + i\sin\theta$. Då är, enligt satsen ovan,

$$z^{-1} = \frac{1}{z} = \frac{e^{0i}}{e^{\theta i}} = e^{-\theta i} = \overline{z}.$$

Ett annat sätt går via den rektangulära formen $z = a + bi$. Om $1 = |z| = \sqrt{a^2+b^2}$, är nämligen

$$z^{-1} = \frac{1}{a+bi} = \frac{a-bi}{(a+bi)(a-bi)} = \frac{a-bi}{a^2+b^2} = a - bi = \overline{z}.$$

Vi hade inte ens behövt ansätta $z = a + bi$. Det hade räckt gott och väl med räkningen

$$z^{-1} = \frac{1}{z} = \frac{\overline{z}}{z\overline{z}} = \frac{\overline{z}}{|z|^2} = \overline{z},$$

eftersom $|z| = 1$.

Den komplexa exponentialfunktionen definieras med hjälp av de trigonometriska funktionerna. Detta skall vi nu vända på, och lösa ut sinus och cosinus i termer av komplexa exponentialer. Dessa formler kommer, så småningom, att förmå läsaren betrakta trigonometri som en matematikens fotgängarzon, med onödiga fartrestriktioner.

Sats 8.4: Eulers formler

För varje reellt argument θ gäller det att

$$\cos\theta = \frac{e^{\theta i} + e^{-\theta i}}{2} \qquad \text{och} \qquad \sin\theta = \frac{e^{\theta i} - e^{-\theta i}}{2i}.$$

Bevis

Ekvationerna (1) och (2) adderas till

$$e^{\theta i} + e^{-\theta i} = (\cos\theta + i\sin\theta) + (\cos\theta - i\sin\theta) = 2\cos\theta,$$

vilket ger uttrycket för $\cos\theta$. Sinusformeln följer på samma sätt genom att subtrahera (2) från (1).

Exempel 7.
Med $\theta = \pi$ i exponentialfunktionen erhålles *Eulers identitet*

$$e^{\pi i} = -1.$$

Detta eleganta samband knyter ihop tre fundamentala tal, som kommer från till synes helt olika delar av matematiken (e från analysen, π från geometrien och i från algebran). För drivna konspirationsteoretiker är det inte svårt, att dra höga växlar på detta yppersta prov på matematisk skönhet. Möjligen kan vi antyda, att den månne utgör ett bevis på Guds existens.

§4. Multiplikation och division av komplexa tal på polär form

Det finns ju, förstås, en intelligent orsak till införandet av polär representation. Den kommer vi till nu. Multiplikation och division av komplexa tal visar sig vara särskilt smidig i denna form, och dessutom äga en åskådlig geometrisk tolkning. Från Sats 8.3 skördar vi utan dröjsmål följande formler för multiplikation och division av komplexa tal på polär form.

Sats 8.5

Två komplexa tal

$$z = r(\cos\alpha + i\sin\alpha) = re^{\alpha i} \qquad \text{och} \qquad w = s(\cos\beta + i\sin\beta) = se^{\beta i}$$

på polär form har produkten

$$zw = rse^{(\alpha+\beta)i} = rs\Big(\cos(\alpha+\beta) + i\sin(\alpha+\beta)\Big)$$

och kvoten

$$\frac{z}{w} = \frac{r}{s}e^{(\alpha-\beta)i} = \frac{r}{s}\Big(\cos(\alpha-\beta) + i\sin(\alpha-\beta)\Big).$$

Kontentan av satsen är följande geometriska tolkning av multiplikation: *Absolutbeloppen multipliceras, medan argumenten adderas.* Analogt gäller för division: *Absolutbeloppen divideras, medan argumenten subtraheras.*

Exempel 8.
Vad är den geometriska tolkningen av multiplikation med i? Multipliceras $re^{\theta i}$ med $i = e^{\frac{\pi}{2}i}$ erhålles

$$re^{\theta i} \cdot e^{\frac{\pi}{2}i} = re^{(\theta+\frac{\pi}{2})i}.$$

Absolutbeloppet förblir oförändrat, medan argumentet växer med $\frac{\pi}{2}$. Talet $re^{\theta i}$ har vridits ett kvarts varv moturs.

Allmänt betyder multiplikation med $e^{\theta i}$ vridning vinkeln θ, ett utmärkt exempel på den geometriska användbarheten av komplexa tal.

Exempel 9.
På samma sätt betyder multiplikation med $1+i$ en multiplikation av absolutbeloppet med $\sqrt{2}$ samt vridning vinkeln $\frac{\pi}{4}$ moturs:

$$re^{\theta i} \cdot (1+i) = re^{\theta i} \cdot \sqrt{2}e^{\frac{\pi}{4}i} = \sqrt{2}re^{(\theta+\frac{\pi}{4})i}.$$

Satsen fungerar analogt vid multiplikation av fler tal än två. Låt oss alltså taga potenser:

Sats 8.6: De Moivres formel

För heltalsexponenter n är

$$\Big(r(\cos\theta + i\sin\theta)\Big)^n = r^n(\cos n\theta + i\sin n\theta).$$

Bevis

För positivt n har vi

$$\Big(r(\cos\theta + i\sin\theta)\Big)^n = \big(re^{\theta i}\big)^n = \underbrace{re^{\theta i} \cdot re^{\theta i} \cdot \cdots \cdot re^{\theta i}}_{n}$$
$$= r^n e^{n\theta i} = r^n(\cos n\theta + i\sin n\theta)$$

enligt satsen för multiplikation. Negativa exponenter n behandlas analogt.

Exempel 10.
Vi vill beräkna $(1-i)^{99}$ på rektangulär form. Det är klart att de 99 parenteserna kan multiplieras ihop, givet ett obegränsat förråd av både tid och papper, men smidigare blir det förstås i polär form.

Absolutbeloppet av basen är $|1-i| = \sqrt{2}$ och argumentet är $\arg(1-i) = \frac{7\pi}{4}$ (rita bild!). Polära formen är

$$1 - i = \sqrt{2}e^{\frac{7\pi}{4}i},$$

och de Moivres formel ger oss

$$(1-i)^{99} = \left(\sqrt{2}e^{\frac{7\pi}{4}i}\right)^{99} = (\sqrt{2})^{99}e^{99\cdot\frac{7\pi}{4}i}.$$

Vi kan förenkla. Absolutbeloppet av talet är

$$(\sqrt{2})^{99} = 2^{49}\sqrt{2}$$

och argumentet är

$$99 \cdot \frac{7\pi}{4} = \left(172 + \frac{5}{4}\right)\pi.$$

Alla tal som skiljer sig från detta med en heltalsmultipel av 2π är också argument (och lika goda!). Det svarar ju mot att snurra ett antal varv. Snurra nu alltså 86 varv medurs till argumentet $\frac{5}{4}\pi$, så fås

$$(1-i)^{99} = 2^{49}\sqrt{2}e^{\frac{5\pi}{4}i}.$$

Talet ligger i tredje kvadranten. (Minns, att kvadranterna numreras moturs med start i den första, där bägge koordinaterna är positiva.)

Med lätthet återgår vi nu till rektangulär form:

$$(1-i)^{99} = 2^{49}\sqrt{2}\left(\cos\frac{5\pi}{4} + i\sin\frac{5\pi}{4}\right) = 2^{49}\sqrt{2}\left(-\frac{1}{\sqrt{2}} - \frac{i}{\sqrt{2}}\right) = -2^{49}(1+i).$$

Mycket räkning med trigonometriska funktioner förenklas med användning av komplexa tal.

Exempel 11.
De Moivres formel säger för $n = 2$, att

$$(\cos\theta + i\sin\theta)^2 = \cos 2\theta + i\sin 2\theta.$$

Kvadrerar vi vänsterledet på traditionellt sätt får vi

$$(\cos^2\theta - \sin^2\theta) + i(2\sin\theta\cos\theta).$$

Dessa två uttryck skall överensstämma, och alltså har vi att

$$\cos 2\theta = \cos^2\theta - \sin^2\theta \qquad \text{och} \qquad \sin 2\theta = 2\sin\theta\cos\theta.$$

Detta är ju formlerna för fördubbling av vinkeln, och de är alltså en rätt omedelbar konsekvens av potenslagarna för komplexa tal.

Läsaren observerar förstås, att logiken är cirkulär, eftersom vi använde additionsformlerna för att få fram potenslagarna. Men här önskar vi mest visa vilken trigonometrisk information, som ligger fördold i potenslagarna. För övrigt kan man ge direkta bevis av dessa, utan att gå omvägen via trigonometri, men då krävs en större apparat.

Exempel 12.
Formler för trippla vinkeln följer från de Moivres formel för $n = 3$, parat med litet kreativ användning av Trigonometriska ettan $\cos^2\theta + \sin^2\theta = 1$:

$$\cos 3\theta = \operatorname{Re} e^{3\theta i} = \operatorname{Re}(\cos\theta + i\sin\theta)^3 = \cos^3\theta - 3\cos\theta\sin^2\theta = 4\cos^3\theta - 3\cos\theta.$$

Exempel 13.
Givet ett positivt heltal n, söker vi en formel för

$$S(x) = 1 + 2\cos x + \cdots + 2\cos nx.$$

Om vi tänker lite suddigt men komplext här, ser vi, att detta nästan är som en geometrisk serie. Om vi nämligen, i stället för $\cos kx$, hade haft e^{ikx}, så vore alla termerna potenser av en och samma bas e^{ix}. Och geometriska serier kan vi förvisso addera. Försöker vi precisera detta, kan vi utnyttja Eulers formel $2\cos x = e^{ix} + e^{-ix}$ och skriva

$$\begin{aligned} S(x) = 1 + 2\cos x + \cdots + 2\cos nx &= 1 + (e^{ix} + e^{-ix}) + \cdots + (e^{inx} + e^{-inx}) = \\ &= e^{-inx} + e^{-i(n-1)x} + \cdots + 1 + \cdots + e^{i(n-1)x} + e^{inx} = \frac{e^{i(n+1)x} - e^{-inx}}{e^{ix} - 1} \end{aligned}$$

I den sista likheten har vi använt formeln för en geometrisk summa. Eftersom summan av reella tal vanligen är reell, borde detta väl synas på svaret. Vi får fortsätta leta efter en lämplig omformulering. Suck! Vi skulle nu kunna övergå till rektangulär form, utföra divisionen och slutligen nå ett uttryck för $S(x)$ i termer av cosinus och sinus av $(n+1)x$, nx och x. Därmed vore förstås problemet löst.

Men, skam till sägandes i en kurs, som aktar sig för att lära ut tjusiga engångstrick, utan bara systematiska och rejäla metoder, finns det en ännu enklare formel. Då använder man den befängda idén, att förlänga uttrycket med $e^{-i\frac{x}{2}}$:

$$\begin{aligned} S(x) = \frac{e^{i(n+1)x} - e^{-inx}}{e^{ix} - 1} = \frac{(e^{i(n+1)x} - e^{-inx})e^{-i\frac{x}{2}}}{(e^{ix} - 1)e^{-i\frac{x}{2}}} \\ = \frac{e^{i(n+\frac{1}{2})x} - e^{-i(n+\frac{1}{2})x}}{e^{i\frac{1}{2}x} - e^{-i\frac{1}{2}x}} = \frac{\sin(n+\frac{1}{2})x}{\sin\frac{x}{2}}. \end{aligned}$$

Den sista likheten följer ur Eulers formel $\sin x = \frac{e^{ix} - e^{-ix}}{2i}$. Formeln för $S(x)$ är användbar i t.ex. signalbehandling, men här platsar den som ett exempel på samspelet mellan trigonometri och komplexa tal.

§5. Binomiska ekvationer

Betrakta, för komplext c och ett positivt heltal n, ekvationen

$$z^n = c.$$

Den innehåller två termer allenast, och är därför känd som en **binomisk** ekvation.

Hade det handlat om reella tal, så att c vore reellt och vi sökte reella lösningar z, skulle vi ha löst den genom att rotutdragning ur bägge sidor. Åtskiljande de två fallen, då n är jämnt eller udda, finge vi för $c \geq 0$ lösningarna

$$z = \begin{cases} \pm\sqrt[n]{c} & \text{om } n \text{ jämnt} \\ \sqrt[n]{c} & \text{om } n \text{ udda.} \end{cases}$$

Ekvationen $z^4 = 1$ har sålunda de två reella lösningarna $z = \pm 1$, medan $z^3 = 1$ har den enda reella lösningen $z = 1$.

Situationen för komplexa lösningar är totalt annorlunda. Speciellt så finns det alltid precis lika många lösningar som ekvationens gradtal (Algebrans fundamentalsats), och vi kan ange dem explicit. Idén är att skriva $c = r(\cos\theta + i\sin\theta)$ på polär form och utnyttja de Moivres formel.

Sats 8.7

Den binomiska ekvationen

$$z^n = r(\cos\theta + i\sin\theta), \qquad r \geq 0,$$

har precis de n lösningarna

$$z = \sqrt[n]{r}\left(\cos\left(\frac{\theta}{n} + k\cdot\frac{2\pi}{n}\right) + i\sin\left(\frac{\theta}{n} + k\cdot\frac{2\pi}{n}\right)\right) = \sqrt[n]{r}e^{\left(\frac{\theta}{n}+k\cdot\frac{2\pi}{n}\right)i}, \tag{3}$$

för $k = 0, 1, \ldots, n-1$.

Bevis

Vi ansätter en lösning i polära koordinater

$$z = s(\cos\zeta + i\sin\zeta),$$

så att, enligt de Moivres formel,

$$z^n = s^n(\cos n\zeta + i\sin n\zeta).$$

Två tal har samma polära koordinater precis när deras absolutbelopp är lika och deras argument skiljer sig åt med en multipel av 2π. Alltså är

$$s^n(\cos n\zeta + i\sin n\zeta) = z^n = r(\cos\theta + i\sin\theta)$$

om och endast om

$$s^n = r \qquad \text{och} \qquad n\zeta = \theta + a\cdot 2\pi$$

för något heltal a. Ur detta ser vi att

$$s = \sqrt[n]{r} \qquad \text{och} \qquad \zeta = \frac{\theta}{n} + a\cdot\frac{2\pi}{n}.$$

Detta verkar ju som detta ger oändligt många lösningar, en för varje heltal a. Men enligt Divisionsalgoritmen kan varje a skrivas entydigt som $a = k + bn$, där k och b är heltal och resten $k = 0, 1, 2, \ldots, n-1$, så att

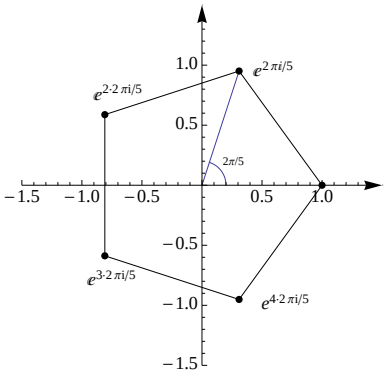
$$\zeta = \frac{\theta}{n} + a\cdot\frac{2\pi}{n} = \frac{\theta}{n} + (k+bn)\frac{2\pi}{n} = \frac{\theta}{n} + k\cdot\frac{2\pi}{n} + b\cdot 2\pi$$

skiljer sig med en heltalsmultipel av 2π från en av lösningarna (3) och ger alltså samma komplexa tal.

Satsen säger först, att alla lösningarna har samma absolutbelopp $\sqrt[n]{r}$ (som för den reella varianten), och sedan, att argumentet för en lösning är $\frac{\theta}{n}$ adderat med en multipel av ett helt varv dividerat på n. Ritas lösningarna upp i det komplexa talplanet, ligger de alltså som hörnen i en regelbunden n-hörning. (Se figuren nedan.)

Exempel 14.
Lösningarna till $z^5 = 1$ fås direkt ur satsen. Vi har att $1 = 1\cdot(\cos 0 + i\sin 0) = e^0$, och de fem lösningarna är

$$z = \cos\frac{2k\pi}{5} + i\sin\frac{2k\pi}{5} = e^{\frac{2k\pi}{5}i}$$



FIGUR 4: Lösningarna till $z^5 = 1$ bildar en regelbunden pentagon.

för $k = 0, 1, 2, 3, 4$. Se Figur 4.

Komplexa andragradsekvationer kan också lösas genom övergång till polär form.

Exempel 15.
Vi önskar lösa ekvationen

$$z^2 + 2iz - 1 - 2i = 0.$$

Kvadratkomplettering ger

$$0 = z^2 + 2iz - 1 - 2i = (z + i)^2 - 2i,$$

varefter substitutionen $w = z + i$ transformerar ekvationen till

$$w^2 = 2i = 2e^{\frac{\pi}{2}i},$$

med lösningarna

$$w = \sqrt{2}e^{\frac{\pi}{4}i} = \sqrt{2}\left(\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\right) = 1 + i$$

och

$$w = \sqrt{2}e^{\frac{5\pi}{4}i} = \sqrt{2}\left(-\frac{1}{\sqrt{2}} - \frac{i}{\sqrt{2}}\right) = -1 - i.$$

Lösningarna till den ursprungliga ekvationen är

$$z = -i \pm (1 + i) = \begin{cases} -1 - 2i \\ 1. \end{cases}$$

ÖVNINGAR

8.1. Skriv på rektangulär form de komplexa tal, vars absolutbelopp och argument är

(a) $\sqrt{2}$, resp. $\frac{\pi}{4}$;
(b) 1 resp. π;
(c) $\sqrt{2}$ resp. $\frac{9\pi}{4}$;
(d) 1 resp. $\frac{\pi}{2}$;
(e) 1 resp. 2π;
(f) $\frac{1}{\sqrt{2}}$ resp. $-\frac{\pi}{4}$;
(g) 1 resp. -100π.

8.2. Rita följande komplexa tal i ett talplan och ange dem på polär form.

(a) 17;
(b) -11;
(c) i;
(d) $-1+i$;
(e) $i\sqrt{3}-1$;
(f) $\sqrt{3}+3i$.

8.3. Bestäm absolutbelopp och argument av

(a) $e^{i\frac{\pi}{8}} = \cos\frac{\pi}{8} + i\sin\frac{\pi}{8}$;
(b) $e^{\theta i} = \cos\theta + i\sin\theta$.

8.4. Beräkna $\left(\frac{1}{2} + i\frac{\sqrt{3}}{2}\right)^{100}$.

8.5. Använd de Moivres formel för att härleda formeln för trippla vinkeln för sinus:

$$\sin 3\theta = 3\sin\theta - 4\sin^3\theta.$$

8.6. Använd Eulers formler för att härleda ett uttryck för $\sin\alpha\cos\beta$, (i termer av andra sinus- och cosinusvärden.)

8.7. Punkterna i det komplexa talplanet vrids vinkeln $\frac{\pi}{2}$ moturs kring origo. Beskriv avbildningen som multiplikation med ett komplext tal. I vilka tal övergår 1 respektive $-3+2i$? Och $a+bi$?

8.8. Punkterna i det komplexa talplanet vrids vinkeln $\frac{5\pi}{6}$ moturs kring origo och multipliceras med 2. Beskriv avbildningen som multiplikation med ett komplext tal. I vilka tal övergår 1 respektive $-1+i$?

8.9. I texten definierade vi $e^{\theta i} = \cos\theta + i\sin\theta$. Mer allmänt sätter vi

$$e^{x+yi} = e^x e^{yi} = e^x(\cos y + i\sin y).$$

Beräkna e^z, om z är

(a) 0;
(b) $i\frac{\pi}{2}$;
(c) $\frac{1}{2}\ln 2 + i\frac{\pi}{4}$;
(d) $i\pi$;
(e) $3-i$.

8.10. Visa att, om $z = x+iy$, så är $|e^z| = e^x$. Vad är $\arg e^z$?

8.11. Lös ekvationerna

(a) $z^2 = 5+12i$;
(b) $z^2 - (2+2i)z - 5 - 10i = 0$.

8.12. Lös ekvationerna (observera att du har tillgång till två metoder!)

(a) $z^2 = -i$;
(b) $z^2 = 1+i$.

8.13. Lös följande ekvationer och rita ut rötterna i det komplexa talplanet.

(a) $z^3 = i$;
(b) $z^3 = 1+i$;
(c) $z^5 = 4i$.

Kapitel 9

Polynom

§1. Varför då?

Låt oss, för omväxlings skull, ge något slags vidare motivering till varför man skall studera ett begrepp, i det här fallet polynom. Idén bakom matematik är att bygga upp enkla modeller av en komplicerad verklighet. Våra matematiska beskrivningar av verkligheten är faktiskt nästan överdrivet enkla.

Tag t.ex. passmyndighetens matematiska modell att ange en människas längd som ett precist heltal i centimeter, och jämför med vad som skulle hända, om de hade konsulterat en petig algebralektor. En människas längd kan ju variera med ett par centimeter mellan morgon och kväll, och det åtminstone måste man väl taga hänsyn till? Således finns det inte ett fixt tal, som beskriver avståndet mellan hjässa och fotsula, utan många, ett för varje tidpunkt på dagen. I stället för ett enda banalt tal får vi då en varierande längd, en härligt komplicerad funktion av tiden på dagen, för att inte tala om månens tidvattenskraft och den valda frisyrens höjd eller eventuella Buffalo-skor för mätobjektet. Det finns säkert också intressanta kvantmekaniska effekter, som kan ställa till det. Det hade krävts dagar av intensivt studium bara för att fylla i en enda rad i passet.

I tillämpningar som denna förekommer funktioner, som vi inte har en chans att beskriva eller beräkna explicit, men som vi, i en förenklad modell av verkligheten, ändå mirakulöst kan komma åt. Vi såg, att en människas längd är en komplicerad funktion, som vi rått och hjärtlöst approximerar med ett enda tal, alltså med en konstant funktion.

Mer generellt behöver vi ett förråd av enkla funktioner, så enkla att man kan göra något intressant med dem, men tillräckligt komplicerade för att ändå komma i närheten av verkligheten. De enklaste funktionerna i matematiken är *polynomen* eller *polynomfunktionerna*, vars värden låter sig beräknas genom tre av de fyra elementära räknesätten: addition, subtraktion och multiplikation. Exempelvis är

$$f(x) = 2x + 3, \quad g(x) = -2x^2 + 3x + 5 \quad \text{och} \quad h(x) = 1.3x^{17} + 2.313x^2 + 36.1x$$

polynom av grad 1, 2 respektive 17.

Mer komplicerade funktioner beräknas ofta genom att de approximeras med polynom. T.ex. så ger

$$e^x \approx 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3$$

ett gott närmevärde till exponentialfunktionen i ett litet intervall kring 0. I den endimensionella analysen möter oss en systematisk teori för dessa *Taylor-polynom*. Miniräknaren, som ju bara behöver beräkna några futtiga, säg åtta, decimaler av oändligt många, fuskar genom att ha inprogrammerat approximerande Taylor-polynom för alla vanliga funktioner.

Polynom utgör alltså en omistlig del av matematikerns verktygslåda, och det är detta som, ur ett statligt ekonomiskt perspektiv, månande om de lönsamma tillämpningarna, motiverar vårt studium av dem. (Sedan tycker ju förstås vi matematiker, att de är intressanta, ja rentav "skitkul", men det är en annan sak.)

§2. Definitioner

Definition 9.1

Med ett **polynom** i variabeln x menas ett uttryck av typen

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

Koefficienterna $a_0, a_1, \ldots, a_n$ kan vara godtyckliga komplexa tal. Förutsatt $a_n \neq 0$, så kallas n för polynomets **grad** och vi skriver $\deg f = n$ (efter engelskans *degree*).

Exempel 1.
Om $n = 0$, så har vi ett *konstant* polynom $f(x) = a_0$.

Exempel 2.
Polynomet $x^3 + 2x + 17$ har grad 3 och koefficienterna $a_3 = 1$, $a_2 = 0$, $a_1 = 2$ och $a_0 = 17$.

Två polynom

$$p(x) = a_n x^n + \cdots + a_1 x + a_0 \qquad \text{och} \qquad q(x) = b_m x^m + \cdots + b_1 x + b_0$$

anses vara lika, om de har samma grad $n = m$, samt motsvarande koefficienter är lika, så att $a_i = b_i$ för varje $i = 0, 1, 2, \ldots, n$.

I föregående upplaga av denna bok fanns insprängd en "övning", där läsaren ombads försöka reta upp sin föreläsare genom att påpeka, att inte alla polynom tilldelas ett gradtal enligt definitionen ovan. Det har varit vår strävan, att i denna nya upplaga söka nedtona dessa mer aggressiva övningar, och vi avslöjar i stället själva, att **nollpolynomet** 0, som är *identiskt* noll, kommer att sakna gradtal enligt vår definition.

Gradtalet definierades som den högsta potensen av x med en nollskild koefficient, så är det klart, att det konstanta polynomet med 7 har graden 0, ty $7 = 7 \cdot x^0$. Kravet på en nollskild koefficient rimmar illa med nollpolynomet 0, som helt saknar sådana. Detta lämnar fältet öppet för hur man vill se nollpolynomet.

Läsaren uppmanades nu tillspörja föreläsaren, vad vederbörande egentligen ansåg nollpolynomet ha för gradtal. Sista delproblemet bestod i att ringa en kvällstidning. Nåväl, det *finns* två konventioner här, och matematikerna är inte överens.

Somliga anser nollpolynomet helt sakna gradtal. Emellertid visar det sig ofta vara bekvämt, att definiera dess grad till

$$\deg 0 = -\infty.$$

Vi väljer denna utväg och skall strax motivera saken.

För att eliminera eventuella missförstånd, när vi vill uttrycka med symboler, att $p(x)$ är nollpolynomet, så undviker vi att skriva $p(x) = 0$, vilket kan tolkas som en ekvation, och skriver istället $p(x) \equiv 0$. Detta utläses som att polynomet $p(x)$ är *identiskt* lika med 0.

§3. Operationer på polynom

Polynom är algebraiska uttryck och de kan därför adderas, subtraheras och multipliceras.

Låt

$$p(x) = a_n x^n + \text{ termer av lägre grad}, \qquad a_n \neq 0, \tag{1}$$

och

$$q(x) = b_m x^m + \text{ termer av lägre grad}, \qquad b_m \neq 0, \tag{2}$$

vara två godtyckliga polynom, av grad n och m, respektive. Deras produkt är

$$p(x)q(x) = a_n b_m x^n x^m + \text{ termer av lägre grad}, \tag{3}$$

Vi ser, att produkten $p(x)q(x)$ av två nollskilda polynom inte kan vara nollpolynomet.

Högstagradstermen i produkten $p(x)q(x)$ är produkten av högstagradstermerna i $p(x)$ och $q(x)$. Speciellt har produkten har grad $m+n$. Även om någon av faktorerna, till exempel $p(x)$, är nollpolynomet (och därmed även produkten är nollpolynomet) så gäller det att $-\infty = -\infty + \deg q(x)$. (Denna additionsformel bör uppfattas intuitivt, då oändligheten inte är något tal i vanlig mening, och aritmetiska operationer inte automatiskt kan överföras till mängden av tal utökad med oändligheten.) Vi har alltså:

Sats 9.2

Produkten av två polynom $p(x)$ och $q(x)$ uppfyller

$$\deg p(x)q(x) = \deg p(x) + \deg q(x).$$

Exempel 3.
För $p(x) = 17x^2 + x + 1$ och $q(x) = 10x^{100} + 1$ är

$$p(x)q(x) = (17x^2 + x + 1)(10x^{100} + 1) = 170x^{102} + 10x^{101} + 10x^{100} + 17x^2 + x + 1,$$

med grad $102 = 100 + 2$.

§4. Delbarhet

Två polynom $p(x)$ och $q(x)$ kan normalt inte divideras, i varje fall inte om vi kräver att kvoten $\frac{p(x)}{q(x)}$ skall vara ett nytt polynom (t.ex. är ju $\frac{1}{x}$ och $\frac{1}{x^2+1}$ inte polynom). Kvoten $\frac{p(x)}{q(x)}$ av två polynom kallas för en **rationell funktion**.

Definition 9.3

Ett polynom $q(x)$ kallas **delare** till polynomet $p(x)$, om det finns ett polynom $k(x)$, så att $p(x) = q(x)k(x)$. Vi säger också att $p(x)$ **delas** av $q(x)$. Vi skriver detta som $q(x) \mid p(x)$.

Ett annat sätt att uttrycka detta är att säga, att kvoten $\frac{p(x)}{q(x)}$ är ett nytt polynom, i varje fall om $q(x)$ inte är nollpolynomet. I så fall måste även $p(x)$ vara nollpolynomet (som delar sig själv), och kvoten $\frac{p(x)}{q(x)}$ är meningslös. Från $\deg p(x) = \deg q(x) + \deg k(x)$ ser vi speciellt att $\deg q(x) \leq \deg p(x)$.

Exempel 4.
Från identiteten

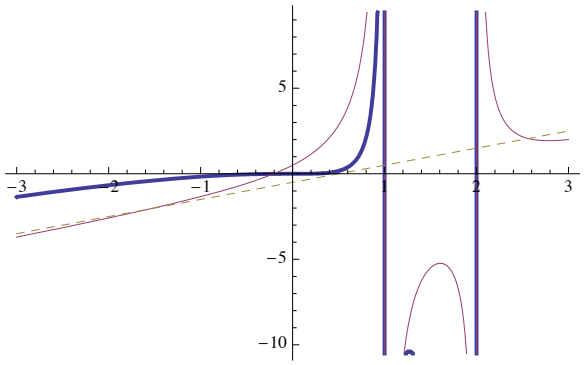
$$x^3 - x = x(x-1)(x+1)$$

ser vi, hurusom $x^3 - x$ delas av x, $x-1$ och $x+1$.

Exempel 5.
Vi har att $x^7 \mid x^n$ om och endast om $n \geq 7$.

Exempel 6.
Det konstanta polynomet $c \neq 0$ delar varje polynom $p(x)$, ty $p(x) = c \cdot \frac{p(x)}{c}$. Omvänt gäller, att det konstanta polynomet c (grad 0) själv bara delas av konstanter.

Exempel 7.
Nollpolynomet 0 delas av alla polynom $q(x)$, ty $0 = 0 \cdot q(x)$. Nollpolynomet är det enda polynom, som delas av alla polynom.



FIGUR 1: Tre rationella funktioner.

Exempel 8.
Vi kan bestämma alla delare $q(x)$ i $x+1$ så här. Antag $x+1 = q(x)k(x)$. Ur $1 = \deg q(x) + \deg k(x)$ springer två möjligheter. Antingen är $\deg q(x) = 1$ och $\deg k(x) = 0$, eller tvärtom. Det betyder att en av faktorerna $q(x)$ eller $k(x)$ måste vara ett konstant polynom $c \neq 0$ och den andra faktorn följaktligen $\frac{1}{c}x + \frac{1}{c}$.

Exempel 9.
Varför intresserar vi oss för begreppet delbarhet? Här är en tillämpning. En rationell funktion $r(x)$ är enligt definitionen en kvot av två polynom

$$r(x) = \frac{p(x)}{q(x)}.$$

Den är definierad för alla x, sådana att $q(x) \neq 0$. I Figur 1 har skisserats graferna till tre rationella funktioner,

$$r_1(x) = \frac{x^3}{(x-1)(x-2)}, \qquad r_2(x) = \frac{-2+7x-7x^2+2x^3}{2(x-1)(x-2)}$$

$$\text{och} \qquad r_3(x) = \frac{1+4x-4x^2+x^3}{(x-1)(x-2)},$$

som alla har samma nämnare $(x-1)(x-2)$ och alltså precis samma definitionsmängder $x \neq 1, 2$. Deras respektive grafer är markerade med tjock, streckad respektive tunn linje.

I likhet med tjuren Ferdinand, är det en av kurvorna, som *inte är som de andra*, nämligen den streckade grafen till $r_2(x)$. I de två andra graferna avtecknar sig nollställena $x = 1$ och $x = 2$ till nämnaren med förfärande skärpa — när x närmar sig dessa värden, går funktionerna mot $\pm\infty$. Men grafen till $r_2(x)$ är en trist rät linje, som inte när någon längtan efter det oändliga i det ändliga.

Förklaringen till denna skillnad är, att *täljaren i $r_2(x)$ är delbar med nämnaren*:

$$r_2(x) = \frac{-2+7x-7x^2+2x^3}{(x-1)(x-2)} = \frac{(2x-1)(x-1)(x-2)}{2(x-1)(x-2)} = x - \frac{1}{2}.$$

Den rationella funktionen $r_2(x)$ överensstämmer därmed med polynomet $x - \frac{1}{2}$ för alla x, utom för $x = 1$ och $x = 2$, där $r_2(x)$ är odefinierad. Dess graf är alltså en linje med två hål i.

Exemplet visar hur viktigt det kan vara, att kunna avgöra, om ett polynom är delbart med ett annat. Vi behöver en systematisk metod för division, och denna kommer nu.

§5. Divisionsalgoritmen

Att dividera polynomet $p(x)$ med polynomet $q(x)$ innebär, precis som för heltal, att söka polynom $k(x)$ och $r(x)$, så att $p(x) = q(x)k(x) + r(x)$. Är $r(x) \equiv 0$, har divisionen gått jämnt ut och $q(x) \mid p(x)$, men detta sker blott alltför sällan.

Liksom för heltal visar sig både kvoten $k(x)$ och resten $r(x)$ vara unika, så fort vi ställer kravet på $r(x)$, att den skall vara "liten". I fallet med polynom skall detta tolkas som $\deg r(x) < \deg q(x)$; resten skall ha mindre grad än divisorn.

Skolan har delat upp svenska folket i två antagonistiska läger vad division anbelangar: tillskyndarna av trappan respektive liggande stolen. Vi föredrar trappan, men det går lika bra med liggande stolen om någon nu skulle föredra denna variant. S.k. "kort division", så ömt omhuldad av den nya generationens pedagoger, är däremot otjänlig för ändamålet.

Exempel 10.
Som illustration skall vi dividera polynomet $p(x) = x^3 + x + 2$ med polynomet $q(x) = x - 1$:

$$\begin{array}{r|l} x-1 & x^3 \qquad +x+2 \end{array}$$

I första steget jämför vi högstagradstermen x^3 i dividenden $x^3 + x + 2$ med högstagradstermen x i divisorn $x - 1$. Vi har $x^3 = x^2 \cdot x$ och skriver därför x^2 ovanför strecket:

$$\begin{array}{r|l} & x^2 \\ \hline x-1 & x^3 \qquad +x+2 \end{array}$$

Sedan multiplicerar vi *hela* divisorn $x - 1$ med x^2 och subtraherar detta från dividenden $x^3 + x + 2$:

$$\begin{array}{r|l} & x^2 \\ \hline x-1 & x^3 \qquad +x+2 \\ & \underline{-x^3+x^2} \end{array}$$

Utför subtraktionen, nu förvandlad till en addition, och flytta ned nästa term x från dividenden:

$$\begin{array}{r|l} & x^2 \\ \hline x-1 & x^3 \qquad +x+2 \\ & \underline{-x^3+x^2} \\ & \qquad x^2 \quad +x \end{array}$$

Vad vi åstadkommit hittills är väsentligen omskrivningen

$$(x^3 + x + 2) - x^2(x - 1) = x^2 + x + 2.$$

(Den sista termen 2 har ännu inte flyttats ned från dividenden. Det går utmärkt att flytta ned den samtidigt med x, om det föredras.) Nu vill vi upprepa förfarandet och subtrahera ännu en multipel av $x - 1$ från återstoden $x^2 + x + 2$.

Jämför alltså högstagradstermen x^2 med högstagradstermen i $x^2 + x$ med högstagradstermen x i divisorn $x - 1$. Vi ser att $x^2 = x \cdot x$, och vi skriver därför x på raden för kvoten:

$$\begin{array}{r|l} & x^2 \quad +x \\ \hline x-1 & x^3 \qquad +x+2 \\ & \underline{-x^3+x^2} \\ & \qquad x^2 \quad +x \end{array}$$

Multiplicera x med divisorn $x-1$ och subtrahera detta från x^2+x+2:

$$\begin{array}{rr|l} & x^2 & +x \\ \hline x-1) & x^3 & +x+2 \\ & -x^3+x^2 & \\ \hline & x^2 & +x \\ & -x^2 & +x \\ \hline \end{array}$$

Utför additionen och flytta ned den sista termen 2 i dividenden:

$$\begin{array}{rr|l} & x^2 & +x \\ \hline x-1) & x^3 & +x+2 \\ & -x^3+x^2 & \\ \hline & x^2 & +x \\ & -x^2 & +x \\ \hline & & 2x+2 \end{array}$$

Denna räkning svarar mot omskrivningen

$$(x^3+x+2)-x^2(x-1)-x(x-1)=2x+2.$$

För det sista steget upprepar vi proceduren än en gång. Jämför högstagradstermen $2x$ i $2x+2$ med högstagradstermen x i divisorn $x-1$. Vi ser att $2x=2\cdot x$, och 2 är alltså sista termen i kvoten:

$$\begin{array}{rr|l} & x^2 & +x+2 \\ \hline x-1) & x^3 & +x+2 \\ & -x^3+x^2 & \\ \hline & x^2 & +x \\ & -x^2 & +x \\ \hline & & 2x+2 \end{array}$$

Multiplicera 2 med $x-1$ och subtrahera:

$$\begin{array}{rr|l} & x^2 & +x+2 \\ \hline x-1) & x^3 & +x+2 \\ & -x^3+x^2 & \\ \hline & x^2 & +x \\ & -x^2 & +x \\ \hline & & 2x+2 \\ & & -2x+2 \\ \hline \end{array}$$

Utför additionen:

$$\begin{array}{rr|l} & x^2 & +x+2 \\ \hline x-1) & x^3 & +x+2 \\ & -x^3+x^2 & \\ \hline & x^2 & +x \\ & -x^2 & +x \\ \hline & & 2x+2 \\ & & -2x+2 \\ \hline & & 4 \end{array}$$

Nu är inte längre högstagradstermen i resten 4 delbar med högstagradstermen x i divisorn $x-1$, vilket är detsamma som att säga att $0=\deg 4<\deg q(x)=1$. Det är följaktligen omöjligt att fortsätta trappan, och vi är färdiga. I och med det sista steget åstadkom vi

$$(x^3+x+2)-x^2(x-1)-x(x-1)-2(x-1)=4,$$

det vill säga

$$x^3+x+2=(x^2+x+2)(x-1)+4.$$

Förhoppningsvis är läsaren övertygad om att räkningen i exemplet fungerar i allmänhet. En viktig egenskap var att graderna av resterna sjönk i varje steg — det betyder att räkningarna garanterat tar slut så småningom. Om $p(x)$ har grad n så slutar divisionen efter maximalt $n+1$ steg. När tar processen slut? Det sker när resten som producerats i ett steg har för liten grad, så att vi inte kan börja på nästa steg. Det sker precis när $\deg r(x) < \deg q(x)$. Observera att, om $r(x) \equiv 0$, så är fortfarande $\deg r(x) < \deg q(x)$. Här är ett exempel till:

Exempel 11.
Division av $x^5 + 2x^4 + 5x + 2$ med $x^2 - x + 1$ producerar följande räkning:

$$\begin{array}{rl|l} & & x^3 + 3x^2 + 2x - 1 \\ \cline{3-3} x^2 - x + 1 \big) & x^5 + 2x^4 \qquad\quad + 5x + 2 & \\ & -x^5 + x^4 - x^3 & \\ \cline{2-2} & \qquad 3x^4 - x^3 & \\ & -3x^4 + 3x^3 - 3x^2 & \\ \cline{2-2} & \qquad\quad 2x^3 - 3x^2 + 5x & \\ & -2x^3 + 2x^2 - 2x & \\ \cline{2-2} & \qquad\quad -x^2 + 3x + 2 & \\ & \qquad\quad x^2 - x + 1 & \\ \cline{2-2} & \qquad\quad 2x + 3 & \end{array}$$

Vi har delat ett polynom av grad 5 med ett polynom av grad 2 och kvoten har grad 3. Den sista resten $r(x) = 2x + 3$ har grad 1, och det innebär att processen är slut. Vi kan inte multiplicera divisorns högstagradsterm x^2 med något x^n (där $n \geq 0$, förstås) och få högstagradstermen $2x$ i $r(x)$.

Vi kan nu formulera ett liknande resultat om polynomdivision, som vi tidigare avhandlade för heltal, under formen av ett allmänt, abstrakt resultat om existensen av en unik kvot och en unik rest.

Sats 9.4: Divisionsalgoritmen

Till varje polynom $p(x)$ och varje nollskilt polynom $q(x)$ finns en unik **kvot** $k(x)$ och en unik **rest** $r(x)$, med egenskaperna att

$$p(x) = k(x)q(x) + r(x) \qquad \text{och} \qquad \deg r(x) < \deg q(x).$$

Bevis

Den specifika divisionsalgoritmen, trappan, som beskrevs ovan, visar alltså hur man kan hitta $k(x)$ och $r(x)$, och speciellt att de existerar. Det återstår att visa, att dessa är unika, så att vi inte kan hitta en annan kvot $\tilde{k}$ och rest $\tilde{r}$ med någon annan metod. Vi prövar alltså, om det månne kan gälla, att

$$p(x) = q(x)k(x) + r(x) = q(x)\tilde{k}(x) + \tilde{r}(x),$$

där även $\deg \tilde{r} < \deg q(x)$. Algebraisk omskrivning ger

$$q(x)(k(x) - \tilde{k}(x)) = \tilde{r}(x) - r(x). \tag{4}$$

Men $\deg(\tilde{r}(x) - r(x)) < \deg q(x)$ och, enligt vad vi ovan observerat, så kan inte $q(x)$ då dela $\tilde{r}(x) - r(x)$ med mindre att det är nollpolynomet. Alltså är $\tilde{r}(x) - r(x) \equiv 0$, således $\tilde{r}(x) = r(x)$. Givet detta, säger ekvation (4) att $q(x)(k(x) - \tilde{k}(x)) \equiv 0$, vilket ger $k(x) - \tilde{k}(x) \equiv 0$. Det är alltså samma kvot $\tilde{k}(x) = k(x)$ och rest $\tilde{r}(x) = r(x)$ det handlar om.

En omformulering är att vi har

$$\frac{p(x)}{q(x)} = \frac{k(x)q(x)}{q(x)} + \frac{r(x)}{q(x)} = k(x) + \frac{r(x)}{q(x)}.$$

§6. Restsatsen och faktorsatsen

Exempel 12.
Vi gör ett tankeexperiment. I det experimentet får vi (av en excentrisk matematiklektor) en resa på en vecka till en grekisk ö, med uppgift att dividera

$$p(x) = (x^2 + 2x + 3)(x - 2)^{1000} + 24$$

med $x - 1$. Sent sista dagen sätter vi äntligen igång med arbetet och inser, med stigande fasa, att detta kommer att kräva ungefär 1003 subtraktioner, två per återstående minut. Dessutom kräver det en hel del papper, som vi naturligtvis glömt ta med. Panik.

Finns det något sätt att fuska? När vi reser hem skall vi alltså ha ett konkret resultat

$$p(x) = (x - 1)k(x) + r(x)$$

att uppvisa, där vi redan från början, enligt Divisionsalgoritmen nyss, vet att $\deg r(x) < \deg(x - 1) = 1$, d.v.s. vi vet att $r(x) = r$ är en reell konstant (grad 0 om $r \neq 0$ och grad $-\infty$ om $r = 0$).

Nå, vi kan åtminstone kan komma åt resten r. Stoppa in $x = 1$ i likheten ovan:

$$p(1) = (1 - 1)k(1) + r = r.$$

Men samtidigt är

$$p(1) = (1 + 2 \cdot 1 + 3)(1 - 2)^{1000} + 24 = 6 \cdot (-1)^{1000} + 24 = 30,$$

så i alla fall har vi hittat $r = 30$, trots baksmälla, och utan överdrivet mycket arbete eller papper. Sedan är det förstås frågan om, hur vi kan övertyga vår uppdragsgivare, att nöja sig med halva det utlovade resultatet. Kanske borde vi läsa psykologi i stället? Resttermer *är* väl mer väsentliga än kvoter...

Exemplet kan vi generalisera till följande sats.

Sats 9.5: Restsatsen

Låt $p(x)$ vara ett polynom, och låt α vara ett komplext tal. Division av $p(x)$ med $x - \alpha$ ger då

$$p(x) = (x - \alpha)k(x) + r,$$

varvid resten är det konstanta polynomet $r = p(\alpha)$.

Bevis

Vi upprepar argumentet i exemplet. Ur Divisionsalgoritmen vet vi att

$$p(x) = (x - \alpha)k(x) + r(x)$$

för några polynom $k(x)$ och $r(x)$, sådana att $\deg r(x) < \deg(x - \alpha) = 1$. Det följer att $r(x)$ är en konstant r. Sättes $x = \alpha$ i denna identitet, så fås

$$p(\alpha) = (\alpha - \alpha)k(\alpha) + r = r,$$

och vi har identifierat resten.

Följande specialfall är synnerligen viktigt och brukar formuleras som en egen sats.

Sats 9.6: Faktorsatsen

Polynomet $p(x)$ är delbart med $x-\alpha$ om och endast om α är ett nollställe till polynomet, så att $p(\alpha)=0$.

Bevis

Att $x-\alpha$ delar $p(x)$ är detsamma som att säga att resten vid division är $r=0$, och detta är alltså sant om och endast om $p(\alpha)=0$.

Att *faktorisera ett polynom* är enligt Faktorsatsen ekvivalent med att *lösa motsvarande polynomekvation*. Vi kommer att skamlöst utnyttja detta i vad som följer.

§7. Partialbråksuppdelning

Exempel 13.
Titta på följande två uttryck för samma funktion:

$$\frac{2}{x^2-1}=\frac{1}{x-1}-\frac{1}{x+1}.$$

En primitiv funktion till vänsterledet syns inte uppenbart, men till högerledet kan vi direkt skriva ned en primitiv funktion:

$$\int \frac{1}{x-1}\,dx-\int \frac{1}{x+1}\,dx=\ln|x-1|-\ln|x+1|+C=\ln\left|\frac{x-1}{x+1}\right|+C.$$

Exemplet visar, hur användbart det är, att kunna omskriva rationella funktioner $\frac{p(x)}{q(x)}$, och indikerar också, att faktoriseringen av nämnaren (i vårt fall $q(x)=x^2-1=(x+1)(x-1)$) spelar en stor roll. En uppdelning av en rationell funktion som i exemplet kallas en **partialbråksuppdelning**. Ett första steg till en allmän metod fås redan av Divisionsalgoritmen. Division av täljaren med nämnaren ger

$$p(x)=k(x)q(x)+r(x),$$

där resten uppfyller $\deg r(x)<\deg q(x)$, så att

$$\frac{p(x)}{q(x)}=k(x)+\frac{r(x)}{q(x)}.$$

Sålunda kan varje rationell funktion skrivas som summan av ett polynom och en ny rationell funktion, i vilken graden av täljaren är mindre än graden av nämnaren. Vi kan då nöja oss med att titta på rationella funktioner av den sistnämnda typen.

Exempel 14.
Vi önskar partialbråksuppdela

$$s(x)=\frac{x+2}{(x+1)(x+5)},$$

och söker konstanter A och B, sådana att

$$s(x)=\frac{x+2}{(x+1)(x+5)}=\frac{A}{x+1}+\frac{B}{x+5}.$$

Vi kan bestämma A och B genom att skriva om högerledet på samma nämnare:

$$\frac{A}{x+1}+\frac{B}{x+5}=\frac{A(x+5)+B(x+1)}{(x+1)(x+5)}=\frac{(A+B)x+(5A+B)}{(x+1)(x+5)}.$$

Om detta skall bli $s(x)$, måste täljarna överensstämma, d.v.s.

$$(A+B)x + 5A + B = x + 2.$$

Identifikation av koefficienterna ger ekvationssystemet

$$\begin{cases} A+B=1 \\ 5A+B=2 \end{cases} \quad \Leftrightarrow \quad \begin{cases} A=\frac{1}{4} \\ B=\frac{3}{4}; \end{cases}$$

alltså är

$$s(x) = \frac{1}{4(x+1)} + \frac{3}{4(x+5)}$$

den sökta partialbråksuppdelningen.

Att detta alltid låter sig göras, garanteras av följande allmänna sats.

Sats 9.7

En rationell funktion

$$s(x) = \frac{p(x)}{(x-\alpha_1)\cdots(x-\alpha_n)},$$

där talen α_k är olika och $\deg p(x) < n$, kan alltid partialbråksuppdelas som

$$s(x) = \frac{A_1}{x-\alpha_1} + \cdots + \frac{A_n}{x-\alpha_n}.$$

Om $p(x)$ har reella koefficienter och talen α_k är reella, gäller detsamma för konstanterna A_k.

Bevis

Vi nöjer oss med att betrakta det enklaste fallet, med två faktorer i nämnaren:

$$s(x) = \frac{cx+d}{(x-\alpha)(x-\beta)}.$$

Vi gör ansatsen

$$s(x) = \frac{A}{x-\alpha} + \frac{B}{x-\beta} = \frac{A(x-\beta)+B(x-\alpha)}{(x-\alpha)(x-\beta)} = \frac{(A+B)x-(\beta A+\alpha B)}{(x-\alpha)(x-\beta)}$$

och identifierar koefficienterna till ekvationssystemet

$$\begin{cases} A+B=c \\ -\beta A-\alpha B=d \end{cases} \quad \Leftrightarrow \quad \begin{cases} A+B=c \\ (\beta-\alpha)B=\beta c+d. \end{cases}$$

Här har vi adderat β gånger den första ekvationen till den andra, vilket ger ett ekvivalent system. Härur fås enkelt den entydiga lösningen

$$\begin{cases} A=\frac{\alpha c+d}{\alpha-\beta} \\ B=\frac{\beta c+d}{\beta-\alpha}. \end{cases}$$

För multipla faktorer i nämnaren blir det mer komplicerat. Metoden framgår bäst av ett antal exempel.

Exempel 15.
Den rationella funktionen

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

har en dubbel faktor i nämnaren. Vi prövar att partialbråksuppdela genom

$$r(x) = \frac{A}{(x+1)} + \frac{B}{(x+1)^2}.$$

Vi kan hitta A och B på samma sätt som i föregående exempel, genom ett ekvationssystem, eller så kan vi dividera täljaren med $x+1$. Från

$$x + 2 = 1 \cdot (x+1) + 1$$

får vi nämligen att

$$\frac{x+2}{(x+1)^2} = \frac{1\cdot(x+1)+1}{(x+1)^2} = \frac{1}{(x+1)} + \frac{1}{(x+1)^2}.$$

Exempel 16.
Den rationella funktionen

$$r(x) = \frac{x+2}{(x+1)^2(x+5)}$$

har både en dubbel och en enkel faktor i nämnaren. Vi ansätter denna gång

$$r(x) = \frac{A}{x+1} + \frac{B}{(x+1)^2} + \frac{C}{x+5}.$$

Konstanterna A, B och C bestäms genom att göra liknämnigt i högerledet:

$$\begin{aligned} r(x) = \frac{A}{x+1} + \frac{B}{(x+1)^2} + \frac{C}{x+5} &= \frac{A(x+5)(x+1) + B(x+5) + C(x+1)^2}{(x+1)^2(x+5)} \\ &= \frac{(A+C)x^2 + (6A+B+2C)x + (5A+5B+C)}{(x+1)^2(x+5)}. \end{aligned}$$

Om detta skall vara $r(x)$, måste täljarna överensstämma, vilket ger systemet

$$\begin{cases} A + C &= 0 \\ 6A + B + 2C &= 1 \\ 5A + 5B + C &= 2. \end{cases} \quad \Leftrightarrow \quad \begin{cases} A = \frac{3}{16} \\ B = \frac{1}{4} \\ C = -\frac{3}{16}. \end{cases}$$

(Det finns systematiska tekniker för att behärska dessa ekvationssystem, som undervisas i linjär algebra.) Alltså är den sökta partialbråksuppdelningen

$$r(x) = \frac{3}{16(x+1)} + \frac{1}{4(x+1)^2} - \frac{3}{16(x+5)}.$$

Det allmänna resultatet blir ganska avancerat att beskriva, så vi nöjer oss med att formulera det för två olika faktorer i nämnaren.

Sats 9.8

En rationell funktion

$$s(x) = \frac{p(x)}{(x-\alpha)^m(x-\beta)^n},$$

där $\alpha \neq \beta$ och $\deg p(x) < m+n$, kan alltid partialbråksuppdelas som

$$s(x) = \frac{A_1}{x-\alpha} + \frac{A_2}{(x-\alpha)^2} + \cdots + \frac{A_m}{(x-\alpha)^m} + \frac{B_1}{x-\beta} + \frac{B_2}{(x-\beta)^2} + \cdots + \frac{B_n}{(x-\beta)^n}.$$

Om $p(x)$ har reella koefficienter och talen α_k är reella, gäller detsamma för konstanterna A_k och B_k.

Satserna vi hittills sett, om partialbråksuppdelning av en rationell funktion $s(x) = \frac{p(x)}{q(x)}$, gäller under villkoret, att nämnaren $q(x)$ kan faktoriseras som en produkt av linjära faktorer. Över de komplexa talen kan detta alltid göras, och ibland går det också över de reella talen, men hur skall vi göra i allmänhet, om vi nu tvunget vill hålla på med reella tal, t.ex. för att vi vill tillämpa resultatet för att leta primitiva funktioner?

I den situationen visar det sig, enligt Sats 10.5 i nästa kapitel, att en reell nämnare $q(x)$ åtminstone kan faktoriseras till en produkt av linjära och kvadratiska polynom med reella koefficienter. Det är denna sort vi nu behandlar. För våra syften räcker det, som förut, att formulera fallet för två olika faktorer i nämnaren.

Sats 9.9

En rationell funktion

$$s(x) = \frac{p(x)}{(x^2+ax+b)^m(x-\alpha)^n},$$

där $\alpha \neq \beta$ och $\deg p(x) < 2m+n$, kan alltid partialbråksuppdelas som

$$s(x) = \frac{A_1x+B_1}{x^2+ax+b} + \frac{A_2x+B_2}{(x^2+ax+b)^2} + \cdots + \frac{A_mx+B_m}{(x^2+ax+b)^m} + \frac{C_1}{x-\alpha} + \frac{C_2}{(x-\alpha)^2} + \cdots + \frac{C_n}{(x-\alpha)^n}.$$

Om alla koefficienter är reella, gäller detsamma för konstanterna A_k, B_k, C_k.

Exempel 17.
Enligt satsen är

$$s(x) = \frac{1}{(x^2+1)x} = \frac{Ax+B}{x^2+1} + \frac{C}{x}.$$

Vi löser detta som förut genom att först göra liknämnigt till

$$\frac{Ax+B}{x^2+1} + \frac{C}{x} = \frac{Ax^2+Bx}{x(x^2+1)} + \frac{C(x^2+1)}{x(x^2+1)},$$

och sedan jämföra täljarna. För att få likhet måste

$$1 = (A+C)x^2 + Bx + C,$$

vilket direkt ger $A = -1$, $B = 0$ och $C = 1$. Vår partialbråksuppdelning är

$$s(x) = \frac{1}{(x^2+1)x} = -\frac{x}{x^2+1} + \frac{1}{x}.$$

Exempel 18.
Den rationella funktionen

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

hanteras som ovan. Efter att ha löst ett ekvationssystem med sex obekanta och sex ekvationer, får vi partialbråksuppdelningen

$$s(x) = -\frac{1}{2(x+1)} + \frac{1}{4(x+1)^2} + \frac{2x+3}{4\,(x^2+x+2)} + \frac{2-7x}{4\,(x^2+x+2)^2}.$$

Övningar

9.1. Bestäm kvot och rest då

(a) $(x^5 + 3x^4 - 2x^3 + 2x - 1)$ divideras med $(x^3 + x + 1)$;

(b) $(x^6 - 1)$ divideras med $(x - 1)$;

(c) $(x^4 + 2x^3 + 25)$ divideras med $(x^2 + 4x + 5)$;

(d) $(x^2 + 4x + 5)$ divideras med $(x^4 + 2x^3 + 25)$.

9.2. När polynomet $p(x)$ divideras med $x - 1$ blir resten 1, och när det divideras med $x - 2$ blir resten 2. Vilken blir resten, då $p(x)$ divideras med $(x - 1)(x - 2)$?

9.3. Partialbråksuppdela

(a) $\frac{1}{x^2-3x+2}$;

(b) $\frac{x+3}{x^2-3x+2}$;

(c) $\frac{x+3}{x^2-2}$;

(d) $\frac{1}{x^3-3x^2+2x}$.

9.4. Partialbråksuppdela

(a) $\frac{1}{x^3+x}$;

(b) $\frac{1}{(x^2+4)(x^2+5)}$.

9.5. Finn alla nollskilda polynom $p(x)$ med egenskapen att

$$p(x^2) = p(x)^2$$

för alla x.

Kapitel 10

Polynomekvationer

§1. Rationella rötter

Att hitta nollställen till polynom är svårt. Därför uppskattar man desto mer de fall, där det går enkelt, t.ex. för andragradsekvationer eller binomiska ekvationer. I detta avsnitt skall vi se ytterligare ett sådant fall. När en polynomekvation med heltalskoefficienter verkligen har heltalsrötter, något som tyvärr nästan aldrig inträffar i verkliga livet, så går det att finna dem genom prövning av ett ganska begränsat antal möjligheter. Denna typ av problem illustrerar också hur man kan använda Faktorsatsen, och åtnjuter därför en intellektuellt suspekt popularitet bland all världens tentamenskonstruktörer.

Exempel 1.
Om ekvationen $z^3+2z^2-2z-1=0$ avslöjar ett snabbt telefonsamtal (se annons under lämplig rubrik i Gula sidorna) till ett algebramedium, att den har en heltalsrot. Lös ekvationen!

Vi går till väga på följande vis. Kalla den okända heltalsroten r. Då är alltså

$$r^3+2r^2-2r-1=0 \qquad \Leftrightarrow \qquad r^3+2r^2-2r=1.$$

Vänsterledet i den sista ekvationen är delbart med r, så det är högerledet 1 också. Men de enda heltal som delar 1 är ± 1. Om det nu finns en heltalsrot (källan kanske kan betvivlas), så är det alltså med nödvändighet en av dessa två. Vi undersöker saken:

$$p(-1)=-1+2+2-1=2, \qquad \text{men} \qquad p(1)=1+2-2-1=0.$$

Eftersom $z=1$ nu är etablerad som en rot är $p(z)=z^3+2z^2-2z-1$, enligt Faktorsatsen, delbart med $z-1$. En snabb uppställning med trappan ger vid handen $p(z)=(z-1)(z^2+3z+1)$. Lösningarna till $p(z)=0$ är alltså dels $z=1$, dels nollställena till $z^2+3z+1=0$. Men den sista är en kvadratisk ekvation, som enkelt löses på vanligt vis. Lösningarna är $z=\frac{-3\pm\sqrt{5}}{2}$. Alltså har vi hittat alla tre rötter till ekvationen.

Det är förstås inte så att polynom speciellt ofta har heltalsrötter, ens givet, att koefficienterna är heltal. Vi såg ju i exemplet ovan, att $z^2+3z+1=0$ inte hade heltalsrötter, utan lösningarna uttrycktes med tuffa irrationella kvadratrötter. I någon mening kan man säga, att sannolikheten för att en ekvation med heltalskoefficienter skall ha heltalsrötter, om koefficienterna slumpas ut, är noll. Vad argumentet om delbarhet i exemplet sade, var endast att, *om* ekvationen har heltalsrötter, så måste dessa dela den konstanta koefficienten a_0, och att det därför bara kan finnas några få möjligheter för dessa heltal. Men man kan ju vinna på Lotto också, så därför kan det löna sig att leta efter heltalsrötter. Man kan till och med få information om rötter som är brutna tal också.

Sats 10.1: Rationella rotsatsen

Låt

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

vara ett polynom med *heltalskoefficienter*. Om $\frac{s}{t}$ är en rationell rot till ekvationen $p(x) = 0$ på slutförkortad form (d.v.s. $\text{SGD}(s,t) = 1$), så gäller att

$$s \mid a_0 \qquad \text{och} \qquad t \mid a_n.$$

Om speciellt $p(x)$ är ett **moniskt** polynom, med högstagradskoefficienten $a_n = 1$, så är alla eventuella rationella lösningar heltal.

Bevis

Vi gör som i exemplet. Om $r = \frac{s}{t}$ är en rot till $p(x) = 0$, så är

$$\begin{aligned} & a_n \left(\frac{s}{t}\right)^n + a_{n-1}\left(\frac{s}{t}\right)^{n-1} + \cdots + a_1\left(\frac{s}{t}\right) + a_0 = 0 \\ \Leftrightarrow \quad & a_n s^n + a_{n-1}s^{n-1}t + \cdots + a_1 s t^{n-1} + a_0 t^n = 0 \end{aligned}$$

genom multiplikation med t^n. Vi kan skriva om den sista ekvationen till

$$a_n s^n + a_{n-1}s^{n-1}t + \cdots + a_1 s t^{n-1} = -a_0 t^n.$$

Vänsterledet är delbart med s, eftersom alla termerna är delbara med s. Därför delar s också högerledet, $s \mid -a_0 t^n$. Men s och t antogs vara relativt prima och har inga primfaktorer gemensamma. Därför delar s den konstanta termen a_0, vilket vi ville visa.

Om vi sedan i stället skriver om ekvationen till

$$a_{n-1}s^{n-1}t + \cdots + a_1 s t^{n-1} + a_0 t^n = -a_n s^n,$$

så är nu högerledet delbart med t, och på samma sätt som nyss ger detta $t \mid a_n$.

Slutligen, om polynomet är moniskt, innebär $t \mid a_n = 1$ att $t = \pm 1$, och då är $\frac{s}{t} = \pm s$ ett heltal.

Exempel 2.
På en tentamen förekom[a] ekvationen

$$2x^3 + x^2 + 2x + 1 = 0.$$

Elementär tentamenspsykologi säger så här: Ekvationen är varken binomisk eller av andra graden, de enda två sorternas ekvationer vi kan systematiskt lösa. Fastmer rör det sig om en tredjegradsekvation, vartill vi inte har fått lära oss den allmänna lösningen. Då kan vi sluta (om inte examinator vill sätta dit oss, och mot det hjälper bara voodoofetischer), att enda möjligheten är en rationell rot.

Styrkta av denna insikt (och en klunk ur fickpluntan), applicerar vi Rationella rotsatsen. En rationell rot $\frac{s}{t}$ (på slutförkortad form) uppfyller $s \mid 1$ och $t \mid 2$. Alltså är de enda möjligheterna $\frac{s}{t} = \pm 1, \pm\frac{1}{2}$. Vi prövar dem i tur och ordning, och upptäcker $-\dfrac{1}{2}$ vara en rot. Det betyder, enligt Faktorsatsen, att $x + \frac{1}{2}$ delar $2x^3 + x^2 + 2x + 1$. Exekvera trappan eller dylikt för att finna faktoriseringen

$$2x^3 + x^2 + 2x + 1 = \left(x + \frac{1}{2}\right)(2x^2 + 2).$$

De återstående rötterna är alltså lösningar till $2x^2 + 2 = 0$, således $x = \pm i$. Därmed är alla lösningar funna.

[a] Detta är en sann historia.

Exempel 3.
Betrakta ekvationen

$$9x^3 + 6x^2 + 15x + 10 = 0.$$

En eventuell rationell rot måste vara något av talen

$$\pm 1,\ \pm 2,\ \pm 5,\ \pm 10,\ \pm\frac{1}{3},\ \pm\frac{2}{3},\ \pm\frac{5}{3},\ \pm\frac{10}{3},\ \pm\frac{1}{9},\ \pm\frac{2}{9},\ \pm\frac{5}{9},\ \pm\frac{10}{9}.$$

Ännu en klunk ur fickpluntan verkar stimulerande och vi kastar oss över uppgiften, att pröva alla dessa möjligheter. Vi uppdagar roten $x = -\frac{2}{3}$.

Men utan den extra klunken kanske vi tänker klarare. De positiva alternativen diskvalificerar sig direkt, ty $9x^3 + 6x^2 + 15x + 10 > 0$ för positiva x. Det räcker följaktligen, att prova de negativa möjligheterna ovan, halva jobbet.

§2. Algebrans fundamentalsats

Vi har sett, att en andragradsekvation har högst två olika rötter. Ibland har den en dubbelrot, t.ex. $x^2 = 0$ med lösningen $x = 0$, och ibland måste vi tillåta irreella tal för att få våra två rötter, t.ex. för $x^2 + 1 = 0$ med lösningarna $x = \pm i$.

I själva verket är det sant, och inte svårt att visa, att en n:tegradsekvation har högst n rötter. Det gäller t o m att den har precis n rötter om vi tillåter komplexa tal och räknar rötter med *multiplicitet*, se nedan. Detta är svårare att visa.

Exempel 4.
Titta en gång till på ekvationen $p(x) = 2x^3 + x^2 + 2x + 1 = 0$ från exemplet ovan. Där kom vi fram till rötterna $x = -\frac{1}{2}$ och $x = \pm i$. Enligt Faktorsatsen vet vi då, att $p(x)$ har de linjära faktorerna $(x + \frac{1}{2})$, $(x + i)$ och $x - i$, vilket leder till faktoriseringen

$$p(x) = 2x^3 + x^2 + 2x + 1 = \left(x + \frac{1}{2}\right)(2x^2 + 2) = 2\left(x + \frac{1}{2}\right)(x + i)(x - i).$$

I denna faktorisering framträder de olika nollställena för $p(x)$. Detta är visdomen från Faktorsatsen: *Att faktorisera ett polynom i linjära faktorer är alldeles detsamma problem som att finna alla dess nollställen.*

Exempel 5.
På samma sätt kan vi faktorisera

$$x^2 + 2x + 1 = (x + 1)(x + 1) = (x + 1)^2,$$

och även här ser vi nollställena till polynomet. Nu är det emellertid bara ett nollställe, men det förekommer, så att säga, två gånger. Vi kallar detta en *dubbelrot* och anser nollstället $x = -1$ ha multiplicitet 2.

Vi vet, att polynom med reella koefficienter inte behöver ha reella nollställen. Det finns många exempel, ett är $x^8 + 13x^4 + 7x^2 + 113$, som ju alltid är strikt positivt, när x är reellt. Komplexa nollställen måste däremot alltid finnas. Någon gång har vi hävdat, att det komplexa livet är enklare än det reella, och det kan man se rätt påtagligt här:

Sats 10.2: Algebrans fundamentalsats

Varje polynom $p(x)$ av positiv grad med komplexa koefficienter, har minst ett komplext nollställe α, d.v.s. $p(\alpha) = 0$.

Bevis

Eftersom man saknar formler för ekvationslösning i allmänhet sker beviset inte genom att hitta rötterna som för första- och andragradsekvationer, utan existensen visas indirekt. Gauss älskade satsen så mycket, att han producerade flera olika bevis i olika skeden av livet. Det kortaste nyttjar ett avancerat maskineri från komplex analys och tar en enda rad.

Beviset är helt klart överkurs, men det kan törhända roa en och annan läsare att det utskrivet. Inget bevis är helt elementärt, men det enklaste vi känner nyttjar endast några satser från flerdimensionell analys. Skriv först

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n,$$

och betrakta

$$|p(x)| = |a_0 + a_1 x + \cdots + a_n x^n| = |x|^n \left| \frac{a_0}{x^n} + \frac{a_1}{x^{n-1}} + \cdots + a_n \right| \to \infty \quad \text{då } |x| \to \infty,$$

ty $\frac{a_0}{x^n} + \frac{a_1}{x^{n-1}} + \cdots + a_n$ går mot $a_n \neq 0$. Som funktionen $|p(x)|$ är kontinuerlig, följer det av en sats i analysen, att den har ett minimivärde, vilket antages i någon punkt $x = \alpha$. Vi skall visa, att detta är vårt sökta nollställe, alltså att $p(\alpha) = 0$.

Funktionen $|p(x + \alpha)|$ har ett minimivärde i punkten $x = 0$. Skriv

$$q(x) = p(x + \alpha) = b_0 + b_k x^k + b_{k+1} x^{k+1} + \cdots + b_n x^n,$$

där b_k är den första nollskilda koefficienten efter b_0. Vi har sedan tidigare en procedur för att lösa binomiska ekvationer i komplexa tal. Låt därför e vara en lösning till ekvationen $e^k = -\frac{b_0}{b_k}$. Låt vidare t vara positivt och reellt. Då är

$$\begin{aligned} q(te) &= b_0 + b_k t^k e^k + b_{k+1} t^{k+1} e^{k+1} + \cdots + b_n t^n e^n \\ &= b_0 + b_k t^k \left(-\frac{b_0}{b_k} \right) + b_{k+1} t^{k+1} e^{k+1} + \cdots + b_n t^n e^n \\ &= b_0(1 - t^k) + b_{k+1} t^{k+1} e^{k+1} + \cdots + b_n t^n e^n. \end{aligned}$$

Antag nu $0 < t < 1$. Triangelolikheten ger

$$\begin{aligned} |q(te)| &= \left| b_0(1 - t^k) + b_{k+1} t^{k+1} e^{k+1} + \cdots + b_n t^n e^n \right| \\ &\leq \left| b_0(1 - t^k) \right| + \left| b_{k+1} t^{k+1} e^{k+1} + \cdots + b_n t^n e^n \right| \\ &= |b_0|(1 - t^k) + t^{k+1} \left| b_{k+1} e^{k+1} + \cdots + b_n t^{n-k-1} e^n \right|. \end{aligned}$$

Uttrycket

$$\left| b_{k+1} e^{k+1} + \cdots + b_n t^{n-k-1} e^n \right|$$

är uppåt begränsat för $0 < t < 1$, säg av konstanten M. Vi får då

$$|q(te)| \leq |b_0|(1 - t^k) + M t^{k+1} = |b_0| + t^k(Mt - |b_0|).$$

Om nu $|b_0| > 0$, så skulle uttrycket $Mt - |b_0|$ bli < 0, då t är ett tillräckligt litet positivt tal. Det skulle medföra $|q(te)| < |b_0| = |q(0)|$, vilket skulle strida mot det faktum, att $x = 0$ är en minimipunkt för $|q(x)|$. Motsägelsen visar, att

$$0 = |b_0| = |q(0)| = |p(\alpha)|,$$

så att $p(\alpha) = 0$.

Man bör stanna upp och låta sig imponeras av vad fundamentalsatsen säger. Ett veritabelt mirakel har inträffat vid övergången från reella till komplexa tal. De komplexa talen skapades

från de reella genom att lägga till *ett enda nytt tal*, nämligen i, ett irreellt nollställe till den extremt simpla ekvationen $x^2+1=0$. Detta enda nya tal räcker för att vinna nollställen till *alla* polynom, t.o.m. de med komplexa koefficienter. Ekvationen $x^8+13x^4+7x^2+113=0$ har alltså minst en lösning. Vi kan inte skriva ned den (den är säkert gräslig), men vi vet att den finns.

Om vi nu har ett polynom $p(x)$ av t.ex. grad 3, så ser vi alltså, från Algebrans fundamentalsats, att det finns ett nollställe $p(\alpha)=0$. Med Faktorsatsen vet vi då, att

$$p(x)=(x-\alpha)k(x),$$

där $k(x)$ är ett kvadratiskt polynom. Även $k(x)$ har ett nollställe $k(\beta)=0$, vilken kan faktoriseras ut till

$$k(x)=(x-\beta)l(x),$$

där $l(x)$ är ett linjärt polynom. Slutligen har även $l(x)$ ett nollställe $l(\gamma)=0$, varav

$$l(x)=(x-\gamma)m(x),$$

där $m(x)=c$ är en konstant. Sätter vi ihop allt detta får vi faktoriseringen

$$p(x)=(x-\alpha)k(x)=(x-\alpha)(x-\beta)l(x)=(x-\alpha)(x-\beta)(x-\gamma)c.$$

Notera att graden sjönk med 1 för varje division, och att vi har lika många linjära faktorer i produkten (tre) som graden av polynomet vi startade med. På samma sätt får vi i allmänhet följande sats.

Sats 10.3

Varje (komplext) polynom $p(x)$ av grad n kan faktoriseras som

$$p(x)=c(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n),$$

där c är högstagradskoefficienten och de komplexa talen $\alpha_1,\alpha_2,\ldots,\alpha_n$ är polynomets nollställen, eventuellt med upprepning.

Som vi ser av faktoriseringen $x^2-2x+1=(x-1)^2$, kan vissa av talen α_i vara lika. Antalet identiska faktorer $x-\alpha_i$ kallas **multipliciteten** för nollstället α_i.

Exempel 6.
Polynomet

$$p(x)=(x-1)^2(x-2)^{30}(x-3),$$

av grad 33, har nollställena $x=1$ av multiplicitet 2 (dubbelrot), $x=2$ av multiplicitet 30 och $x=3$ av multiplicitet 1 (enkelrot).

Generellt kan vi skriva ett polynom $p(x)$ som

$$p(x)=c\prod_{k=1}^{m}(x-\alpha_k)^{d_k},$$

där $c\neq 0$, $\alpha_k\neq\alpha_l$ för $k\neq l$, och multipliciteten för nollstället α_k är $d_k\geq 1$.

§3. Konjugerade rötter

Nu skall vi se, vad teorien i föregående avsnitt har för konsekvenser för polynom med enbart reella koefficienter.

Sats 10.4: Konjugerade rotsatsen

Låt $p(x)$ vara ett polynom med *reella* koefficienter. Om det irreella talet α ($\operatorname{Im}\alpha\neq 0$) är ett nollställe, så är även dess konjugat $\overline{\alpha}$ ett nollställe. Irreella nollställen till reella polynom kommer alltså i konjugerade par.

Bevis

Vi visar hur det fungerar för ett tredjegradspolynom

$$p(x) = a_3x^3 + a_2x^2 + a_1x + a_0,$$

där koefficienterna a_i är reella (för $i = 0, 1, 2, 3$). Det betyder att $\overline{a}_i = a_i$ (reella tal är sina egna konjugat). Antag α vara ett nollställe, så att $p(\alpha) = 0$. Med tillämpning av räknereglerna för konjugering fås

$$\begin{aligned} 0 = \overline{0} = \overline{p(\alpha)} &= \overline{a_3\alpha^3 + a_2\alpha^2 + a_1\alpha + a_0} = \overline{a_3\alpha^3} + \overline{a_2\alpha^2} + \overline{a_1\alpha} + \overline{a_0} \\ &= \overline{a_3} \cdot \overline{\alpha^3} + \overline{a_2} \cdot \overline{\alpha^2} + \overline{a_1} \cdot \overline{\alpha} + \overline{a_0} = a_3\overline{\alpha}^3 + a_2\overline{\alpha}^2 + a_1\overline{\alpha} + a_0 = p(\overline{\alpha}). \end{aligned}$$

Detta säger precis att $\overline{\alpha}$ också är ett nollställe.

Exempel 7.
Polynomet

$$p(x) = x^4 + 2x^3 + 3x^2 + 2x + 2$$

har nollställena i och $-1 + i$ (telefonsamtal till ett medium!). Vi vill bestämma samtliga nollställen och faktorisera polynomet i en produkt av reella faktorer av så små gradtal som möjligt.

Polynomet är reellt, och enligt Konjugerade rotsatsen är då också konjugaten $-i$ och $-1 - i$ nollställen till polynomet. Nu har vi alla fyra möjliga nollställen (graden är ju 4), och polynomet faktoriseras till

$$p(x) = (x - i)(x + i)(x + 1 - i)(x + 1 + i).$$

Polynomet saknar reella nollställen, varför vi heller inte kan ha någon linjär faktor med reella koefficienter. Däremot kan de irreella förstagradsfaktorerna ovan paras ihop till reella kvadratiska faktorer, ty

$$(x - i)(x + i) = x^2 + 1$$

och

$$(x + 1 - i)(x + 1 + i) = (x + 1)^2 - i^2 = x^2 + 2x + 2.$$

I dessa har vi multiplicerat ihop de linjära faktorer, som hör till ett nollställe jämte dess konjugat. Dessa reella andragradspolynom ger oss en reell faktorisering

$$p(x) = (x^2 + 1)(x^2 + 2x + 2).$$

Det är förvisso sant, även för ett reellt nollställe α till ett reellt polynom $p(x)$, att också $\overline{\alpha}$ är ett nollställe. Då får vi ju inte ett nytt nollställe, när vi konjugerar, utan återfår bara det vi redan hade. Men de irreella nollställena förekommer enligt satsen parvis, α samtidigt med $\overline{\alpha}$. Som i exemplet ovan kan vi alltid faktorisera ett reellt polynom i en produkt av linjära faktorer, svarande till de reella nollställena, och kvadratiska faktorer, svarande till de irreella nollställena. Alla faktorer kommer härvidlag att få reella koefficienter. Det reella faktorer som inte går att faktorisera över de reella talen kallas **irreducibla**.

Sats 10.5

Låt $p(x)$ vara ett polynom med reella koefficienter. Antag de reella nollställena vara $\alpha_1, \ldots, \alpha_m$ och de irreella vara $a_1 \pm b_1 i, \ldots, a_n \pm b_n i$ (där varje $b_k \neq 0$). Då kan polynomet faktoriseras i irreducibla reella faktorer enligt

$$p(x) = c(x-\alpha_1)\cdots(x-\alpha_m)\Big(x^2 - 2a_1x + (a_1^2+b_1^2)\Big)\cdots\Big(x^2 - 2a_nx + (a_n^2+b_n^2)\Big),$$

där c är (reella) högstagradskoefficienten i $p(x)$.

Ett polynom med reella koefficienter kan alltid faktoriseras som en produkt av linjära faktorer, men i allmänhet bara om vi sträcker oss till komplexa tal (Algebrans fundamentalsats). En faktorisering i reella faktorer är möjlig, men till priset av, att vi måste acceptera även kvadratiska faktorer.

Bevis

Från Sats 10.3 har vi en faktorisering i linjära faktorer, vari de konjugerade rötternas faktorer paras ihop till

$$\begin{aligned}\Big(x-(a_k+b_ki)\Big)\Big(x-(a_k-b_ki)\Big) &= \Big((x-a_k)-b_ki\Big)\Big((x-a_k)+b_ki\Big)\\ &= (x-a_k)^2 - b_k^2i^2 = x^2 - 2a_kx + (a_k^2+b_k^2).\end{aligned}$$

§4. Samband mellan rötter och koefficienter

Ett kvadratiskt polynom x^2+ax+b kan, enligt Algebrans fundamentalsats, faktoriseras som

$$x^2+ax+b = (x-\alpha)(x-\beta),$$

där α och β är polynomets nollställen. Om vi multiplicerar ut högerledet här, får vi

$$x^2+ax+b = x^2 - (\alpha+\beta)x + \alpha\beta,$$

och alltså har vi att

$$a = -(\alpha+\beta) \qquad \text{och} \qquad b = \alpha\beta.$$

Direkt från polynomet kan vi alltså utläsa, vad rötternas summa är, liksom deras produkt!

Ett kubiskt polynom x^3+ax^2+bx+c kan, på samma sätt, faktoriseras som

$$\begin{aligned}x^3+ax^2+bx+c &= (x-\alpha)(x-\beta)(x-\gamma)\\ &= x^3 - (\alpha+\beta+\gamma)x^2 + (\alpha\beta+\beta\gamma+\gamma\alpha)x - \alpha\beta\gamma,\end{aligned}$$

varav

$$a = -(\alpha+\beta+\gamma), \qquad b = \alpha\beta+\beta\gamma+\gamma\alpha \qquad \text{och} \qquad c = -\alpha\beta\gamma.$$

Dessa formler är specialfall av en familj mer omfattande samband mellan rötter och koefficienter.

Sats 10.6: Vietas formler

Nollställena $\alpha_1, \ldots, \alpha_n$ (räknade med multiplicitet) till polynomet

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1} + x^n$$

uppfyller identiteterna

$$\begin{aligned} \alpha_1 + \cdots + \alpha_n &= -a_{n-1} \\ \alpha_1\alpha_2 + \cdots + \alpha_{n-1}\alpha_n &= a_{n-2} \\ \alpha_1\alpha_2\alpha_3 + \cdots + \alpha_{n-2}\alpha_{n-1}\alpha_n &= -a_{n-3} \\ &\cdots \\ \alpha_1\alpha_2\cdots\alpha_n &= (-1)^n a_0. \end{aligned}$$

Notera speciellt den första och sista formeln, som ger rötternas summa respektive produkt. Poängen är att vi, utan att anstränga oss med att *hitta* nollställena, direkt från polynomets koefficienter kan läsa av deras summa och produkt.

Exempel 8.
Ekvationen

$$x^7 + (1+i)x^5 - 173x^3 + 723 = 0$$

ser jobbig ut (det är den). Utan att lösa den vet vi, från Vietas formler, att dess sju rötter, räknade med multiplicitet, har summan 0 och produkten -723.

Övningar

10.1. Bestäm konstanten a så att polynomet $p(x) = x^3 - 2x^2 - 19x + a$ blir delbart med $x - 1$. Faktorisera sedan polynomet för detta värde på a.

10.2. Polynomet $p(x) = x^5 - 10x^2 + 15x - 6$ har nollstället $x = 1$. Bestäm nollställets multiplicitet och faktorisera polynomet i irreducibla reella faktorer.

10.3. Ange samtliga rötter till $z^6 - 2z^3 + 2 = 0$.

10.4. Ekvationen

$$z^4 - z^3 + 7z^2 - 9z - 18 = 0$$

har en rent imaginär rot. Lös ekvationen.

10.5. Ekvationen

$$z^4 - 2z^3 + 9z^2 - 14z + 14 = 0$$

har en rot med realdelen 1. Lös ekvationen fullständigt.

10.6. Visa, att man kan bestämma konstanten a, så att polynomet $p(x) = x^4 + 2x^3 + 3x^2 + ax + 2$ får faktorn $x^2 + 2x + 2$. Lös därefter för detta värde på a ekvationen $p(x) = 0$ fullständigt.

10.7. Ekvationen $z^4 - 2z^3 + 2z^2 - 10z + 25 = 0$ har rötterna $z = 2 + i$ och $z = -1 - 2i$. Lös ekvationen fullständigt.

10.8. Ange ett sjättegradspolynom med reella koefficienter, som har enkelt nollställe i $z = 2 - i$ och dubbelt nollställe i i.

10.9. Faktorisera i irreducibla reella faktorer:

(a) $x^2 - 4$;

(b) $x^2 + 2x + 1$;

(c) $x^3 - x$;

(d) $x^3 - 3x + 2$;

(e) $2 - x - x^2$;

(f) $x^4 + 27x$;

(g) $x^6 - 64$.

10.10. Faktorisera $p(x) = x^5 - x^4 + 4x - 4$, som har ett rationellt nollställe, i reella faktorer av så lågt gradtal som möjligt.

10.11. Faktorisera $p(x) = x^6 - 8$ i irreducibla reella faktorer, dvs faktorer av så lågt gradtal som möjligt.

10.12. Bestäm alla rötter, samt deras summa och produkt, för var och en av ekvationerna

(a) $z^2 - 7z + 10 = 0$;

(b) $3z^2 - 21z + 30 = 0$;

(c) $z^3 - 7z^2 + 10z = 0$.

Jämför med Vietas formler.

10.13. Bestäm summan och produkten av samtliga rötter till ekvationen

$$z^7 + (3 - i)z^6 + \pi z^3 + e = 0.$$

10.14. Hur många rötter (komplexa och räknade med multiplicitet) har ekvationen

$$(1 + z^3)^8 = (1 + z^4)^6?$$

10.15. Finn samtliga rötter till ekvationen

$$z^4 - z^3 - (6 + i)z^2 + iz + 6i = 0.$$

10.16. Visa, att polynomet $px^5 + x^3 + q$, där p och q är udda primtal, saknar rationella nollställen.

10.17. Undersök, för alla möjliga primtal p och q, huruvida ekvationen

$$x^3 + px = q$$

har några rationella rötter.

Del IV

Diskret matematik

Kapitel 11

Induktion

§1. Ett motiverande exempel

Betrakta talföljden a_n, definierad genom rekursionsformeln

$$a_n = \begin{cases} \dfrac{1}{2 - a_{n-1}} & \text{om } n \geq 1 \\ 0 & \text{om } n = 0. \end{cases}$$

För att få en uppfattning om denna, beräknar vi de första elementen:

$$\begin{aligned} a_0 &= 0 \\ a_1 &= \frac{1}{2 - a_0} = \frac{1}{2 - 0} = \frac{1}{2} \\ a_2 &= \frac{1}{2 - a_1} = \frac{1}{2 - \frac{1}{2}} = \frac{2}{3} \\ a_3 &= \frac{1}{2 - a_2} = \frac{1}{2 - \frac{2}{3}} = \frac{3}{4} \\ a_4 &= \frac{1}{2 - a_3} = \frac{1}{2 - \frac{3}{4}} = \frac{4}{5}. \end{aligned}$$

Ett mönster framträder. Det förefaller som om en allmän formel för talföljden vore $a_n = \frac{n}{n+1}$. Men hur bevisar vi den? Det är klart, att vi skulle kunna fortsätta beräkna

$$\begin{aligned} a_5 &= \frac{1}{2 - a_4} = \frac{1}{2 - \frac{4}{5}} = \frac{5}{6} \\ a_6 &= \frac{1}{2 - a_5} = \frac{1}{2 - \frac{5}{6}} = \frac{6}{7} \ldots, \end{aligned}$$

men färdiga blir vi aldrig. Det finns ju oändligt många tal n. Vi kan inte härifrån deducera formelns giltighet ens för nästa tal $n = 7$. Och även om vi räknar en gång till och verifierar att $a_7 = \frac{7}{8}$, eller ens räknar så långt som datorn orkar med, är dessa beräkningar inget slutgiltigt bevis för att formeln alltid är sann. Varför skulle det inte kunna finnas något riktigt stort n, för vilket den är falsk?

Ett sätt är förstås att försöka förstå *varför* räkningarna ovan ser ut som de gör. Vi satte ovan in 0 i rekursionsformeln och fick ut $\frac{1}{2}$, satte in $\frac{1}{2}$ och fick ut $\frac{2}{3}$, och så vidare. Blir det alltid så? Vad skulle ske, om vi faktiskt provade med $a_n = \frac{n}{n+1}$? Kommer då nästa tal i talföljden alltid att bli $\frac{n+1}{n+2}$ (som vi alltså tror är formeln för a_{n+1})? Enkel räkning visar att så verkligen är fallet:

$$a_{n+1} = \frac{1}{2 - a_n} = \frac{1}{2 - \frac{n}{n+1}} = \frac{n+1}{2(n+1) - n} = \frac{n+1}{n+2}.$$

Känner vi oss övertygade? Vi har här ett allmänt resonemang, som visar, att om formeln $a_n = \frac{n}{n+1}$ stämmer i ett steg, så stämmer den även i nästa. Samma uträkning, som ovan gav oss $a_1 = \frac{1}{2}$, $a_2 = \frac{2}{3}$, $a_3 = \frac{3}{4}$, …, kan därför fortsättas i all evighet och ge $a_n = \frac{n}{n+1}$ för alla n. Världen är inte så ond ändå.

§2. Dominobrickorna

Grunden för resonemanget ovan är en matematisk idé vid namn *induktion*. Vi visar den nu i sin mest renodlade form. Låt oss placera ett antal dominobrickor D_1, D_2, D_3, ... på högkant på den reella tallinjen så, att bricka D_n står ovanpå talet n. Till höger om bricka D_n står då bricka D_{n+1}. Om vi puttar den första brickan åt höger, mot den andra brickan, så trillar förstås alla brickorna så småningom. Låt oss se närmare på, hur man kan resonera för att komma till denna insikt, mest för att genomskåda exemplet ovan.

Vi döper påståendet, att en viss bricka trillar.

$P(n)$: "Bricka D_n faller."

Vi inser omedelbart att, oberoende av vilket heltal n det är fråga om, så gäller följande implikation.

$P(n) \Rightarrow P(n+1)$: "Om bricka D_n faller, så faller även bricka D_{n+1}."

Detta är ju förstås epistemologiskt uppenbart, ty det säger ju självt, att om en bricka ramlar åt höger, så kommer brickan närmast till höger att få sig en knuff och därför också ramla. Hur trivialt det än låter, så sammanfattar den logiska implikationen $P(n) \Rightarrow P(n+1)$ oändligt många påståenden, nämligen samtliga implikationer

$$P(1) \Rightarrow P(2), \qquad P(2) \Rightarrow P(3), \qquad P(3) \Rightarrow P(4), \qquad \ldots.$$

Nu startar vi kedjereaktionen genom att ge den första brickan en puff:

$P(1)$: "Bricka D_1 faller."

Bricka D_1 trillar åt höger. Varför trillar alla brickorna? Att $P(1)$ är sann *och* implikationen $P(1) \Rightarrow P(2)$, så följer att $P(2)$ också är sann, det vill säga bricka D_2 faller. Eftersom $P(2) \Rightarrow P(3)$ gäller, så följer att $P(3)$ är sann, det vill säga bricka D_3 faller. Fortsätter vi detta resonemang, så når vi så småningom, att $P(1000)$ är sann, det vill säga bricka D_{1000} faller. Och det slutar förstås inte där. Vi deducerar, att alla utsagorna $P(n)$ är sanna. Alltså trillar alla brickorna.

Låt oss sammanfatta detta som en allmän princip, eftersom det uppenbarligen fungerar oavsett naturen av påståendet $P(n)$.

Induktionsprincipen. Låt $P(n)$ vara en familj av påståenden, där $n = 1, 2, 3, \ldots$.
Antag följande.

- $P(1)$ är sann.
- Implikationen $P(n) \Rightarrow P(n+1)$ är sann för varje n.

Då kan vi dra slutsatsen, att $P(n)$ är sann för alla $n = 1, 2, 3, \ldots$.

Påståendena (dominobrickorna) måste inte börja numreras från $n = 1$. Det går lika bra att starta från $n = 0$. Det viktiga är, att det finns något initialvärde.

Matematisk induktion är alltså en metod för att visa satser, som gäller för alla naturliga tal genom att visa dem "ett steg i taget".

§3. Identiteter

Summan av de n första heltalen ges av en behändig formel

$$s_n = 1 + 2 + 3 + 4 + \cdots + n = \frac{1}{2}n(n+1),$$

ett specialfall av den allmänna summaformeln för en aritmetisk talföljd. Nu söker vi en formel för summan av de n första kvadraterna

$$k_n = 1^2 + 2^2 + 3^2 + \cdots + n^2.$$

Denna summa är varken aritmetisk eller geometrisk, och kan därför ej beräknas med de metoder vi hittills sett. En bra början är att göra några experiment, d.v.s. räkna lite:

n	1	2	3	4	5	6
s_n	1	3	6	10	15	21
k_n	1	5	14	30	55	91
$\frac{k_n}{s_n}$	1	$\frac{5}{3}$	$\frac{7}{3}$	3	$\frac{11}{3}$	$\frac{13}{3}$

Tydligen växer k_n snabbare än s_n, och för att jämföra har vi infört en rad med kvoten mellan de två följderna. Kvoten växer stadigt med $\frac{2}{3}$ i varje steg, så med vår tabell som grund framkastar vi hypotesen $\frac{k_n}{s_n} = \frac{2n+1}{3}$. Eftersom vi har en formel för s_n, leder detta till en förmodad formel för kvadratsumman:

$$k_n = \frac{k_n}{s_n} \cdot s_n = \frac{2n+1}{3} \cdot \frac{n(n+1)}{2} = \frac{n(n+1)(2n+1)}{6}.$$

Vi kan enkelt kontrollera, att vår formel verkligen är giltig för $n = 1, \ldots, 6$.

Påståendet, att denna formel är riktig för kvadratsumman k_n, betecknar vi med $P(n)$.

$P(n)$: Summan av de första n kvadraterna är

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

Det finns ett påstående för varje positivt heltal n. Påståendet $P(3)$ lyder t.ex.

$$1^2 + 2^2 + 3^2 = \frac{3(3+1)(2\cdot 3+1)}{6}.$$

Uträkningen tidigare visar att påståendena $P(1)$ till $P(6)$ alla är sanna. Än så länge vet vi ingenting om sanningshalten hos $P(7)$ och därutöver.

Om vi nu kan visa att $P(n) \Rightarrow P(n+1)$, för varje positivt heltal n, så är vi precis i situationen med dominobrickorna ovan, och vi kan dra slutsatsen, att $P(n)$ är sann för alla n. Så låt oss försöka göra detta. Detta är *induktionssteget*.

Metoden för att visa en implikation som $P(n) \Rightarrow P(n+1)$ är, att *antaga att förledet $P(n)$ är sant*, och sedan *ge ett argument för att då också $P(n+1)$ är sann*. Antagandet, att $P(n)$ är sann, kallas i dessa sammanhang för *induktionsantagandet*.

Antag alltså, att n är ett fixt positivt heltal, för vilket $P(n)$ är sann. Utsagan $P(n)$, som vi alltså antager sann, säger förstås att

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}, \tag{1}$$

medan $P(n+1)$, som vi önskar visa, säger att

$$\begin{aligned} 1^2 + 2^2 + 3^2 + \cdots + n^2 + (n+1)^2 &= \frac{(n+1)(n+1+1)(2(n+1)+1)}{6} \\ &= \frac{(n+1)(n+2)(2n+3)}{6}. \end{aligned}$$

Denna likhet kan vi nu visa som följer.

$$\begin{aligned} 1^2 + 2^2 + 3^2 + \cdots + n^2 + (n+1)^2 &= \frac{n(n+1)(2n+1)}{6} + (n+1)^2 \\ &= \frac{n+1}{6}\Big(n(2n+1) + 6(n+1)\Big) \\ &= \frac{n+1}{6}\left(2n^2 + 7n + 6\right) \\ &= \frac{(n+1)(n+2)(2n+3)}{6}. \end{aligned}$$

I första steget nyttjade vi induktionsantagandet (1), resten är bara algebra. Alltså har vi visat induktionssteget.

Vi har då både $P(1)$ och implikationen $P(n) \Rightarrow P(n+1)$. Enligt Induktionsprincipen gäller då $P(n)$ för alla n. Alla dominobrickor har fallit!

Resonemanget var ganska mångordigt och trygferat med våra förklaringar och anmärkningar. Ett renskrivet bevis, mötande Livsmedelsverkets striktaste kvalitetskrav, lyder som följer:

Exempel 1.
Vi önskar bevisa formeln

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

för alla positiva heltal n. Formeln gäller då $n = 1$, för

$$1^2 = 1 = \frac{1(1+1)(2\cdot 1+1)}{6}.$$

Antag, att formeln

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

gäller för *något* värde på n. Då gäller formeln även för *nästa* värde på n, ty

$$\begin{aligned} 1^2 + 2^2 + 3^2 + \cdots + n^2 + (n+1)^2 &= \frac{n(n+1)(2n+1)}{6} + (n+1)^2 \\ &= \frac{n+1}{6}\Big(n(2n+1) + 6(n+1)\Big) \\ &= \frac{n+1}{6}\left(2n^2 + 7n + 6\right) \\ &= \frac{(n+1)(n+2)(2n+3)}{6}. \end{aligned}$$

Enligt Induktionsprincipen gäller då formeln för *alla* värden på n.

Induktionsbeviset är en sträng ritual att underkasta sig. Man gör klokt i, att skriva det enligt ett noggrant utstakat formulär, till exempel som vi skrivit det ovan. Observera särskilt de kursiverade orden *något* och *alla*, som förhindrar cirkelresonemang. Beviset avslutas med ett åkallande av den heliga Induktionsprincipen.

Exempel 2.
Låt oss formulera ett strikt induktionsbevis för att talföljden

$$a_n = \begin{cases} \dfrac{1}{2 - a_{n-1}} & \text{om } n \geq 1 \\ 0 & \text{om } n = 0 \end{cases}$$

har den explicita formeln $a_n = \frac{n}{n+1}$. Här tar induktionen sin början i $n = 0$.

Formeln gäller för $n = 0$, för direkt från definitionen har vi $a_0 = 0 = \frac{0}{1}$. Antag, att formeln $a_n = \frac{n}{n+1}$ gäller för *något* värde på n. Då gäller formeln även för *nästa* värde på n, ty

$$a_{n+1} = \frac{1}{2 - a_n} = \frac{1}{2 - \frac{n}{n+1}} = \frac{n+1}{2(n+1) - n} = \frac{n+1}{n+2}.$$

Enligt Induktionsprincipen gäller formeln för *alla* n.

Exempel 3.
Dyrkad inom differentialkalkylen är formeln för derivatan av en potensfunktion:

$$D(x^n) = nx^{n-1}, \qquad n \geq 1.$$

I detta fall har vi återigen ett påstående för varje positivt heltal n. För små värden på n kan påståendet vara lätt att inse, medan svårigheterna växer med n. Vi behöver därför en allmän metod.

Om vi till exempel vill visa att $D(x^6) = 6x^5$, så ligger det nära till hands att använda produktregeln:

$$D(x^6) = D(x \cdot x^5) = 1 \cdot x^5 + x \cdot D(x^5).$$

Om vi nu redan vet motsvarande regel för $n = 5$, d.v.s. att $D(x^5) = 5x^4$, så kan vi fortsätta:

$$D(x^6) = x^5 + xD(x^5) = x^5 + x \cdot 5x^4 = 6x^5.$$

Om vi i stället vill visa $D(x^{97}) = 97x^{96}$, så är det kanske inte så troligt, att vi redan känner till motsvarande formel av lägre ordning, d.v.s. för $n = 96$. Genom att fortsätta proceduren nedåt kommer vi dock så småningom ner till ett känt påstående (t.ex. $D(x^2) = 2x$). Detta är ett alternativt sätt att tänka på matematisk induktion. Motsvarigheten i exemplet med dominobrickorna är att bricka 97 trillar, om bricka 96 trillar, om bricka 95... om bricka 2 trillar, om bricka 1 trillar. Emellertid brukar man, som vi sett, utgå från fallet $n = 1$ (eller 0) och arbeta uppåt, snarare än att börja med ett fixt n och arbeta sig nedåt.

Så här ser då ett induktionsbevis ut. Vi skall visa formeln $D(x^n) = nx^{n-1}$ för varje positivt heltal n. Formeln stämmer för $n = 1$, ty $Dx = 1 = 1 \cdot x^{1-1}$. Antag nu formeln stämma för ett *visst* n, det vill säga $D(x^n) = nx^{n-1}$ för detta n. Formeln är då också sann för $n + 1$, ty vi beräknar med produktregelns tillhjälp

$$D(x^{n+1}) = D(x \cdot x^n) = D(x)x^n + xD(x^n) = x^n + nxx^{n-1} = (n + 1)x^n,$$

där vi i tredje steget nyttjade induktionsantagandet. Enligt Induktionsprincipen är då formeln sann för *alla* positiva heltal.

§4. Kombinatoriska problem

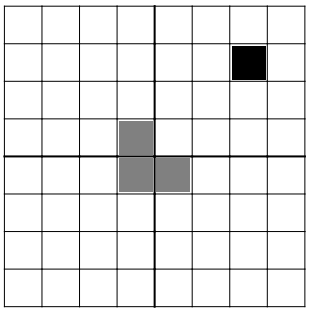
Exempel 4.
Givet är ett kvadratiskt rutnät G_n med 2^n rutor längs varje sida. Rutnätet är *defekt*, då en av dess rutor saknas. Till vårt förfogande har vi en massa **triomino-brickor**: tre kvadrater, lika stora som rutorna, sammanfogade i formen av bokstaven L. Vi påstår, att det defekta rutnätet alltid kan **tesselleras** (täckas utan överlappning) med triomino-brickor.

Det påstående som skall bevisas kan alltså formuleras som: Det defekta rutnätet G_n kan tesselleras av triomono-brickor, oavsett den felande rutans position. Bassteget är trivialt: G_1 har format 2×2 och är själv identisk med en triomino-bricka.

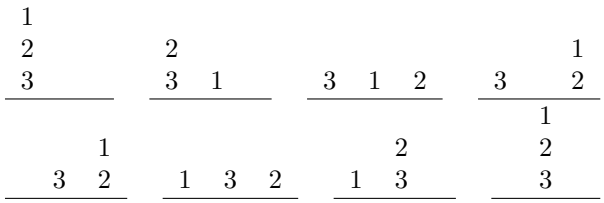
Induktionssteget: Vi antar att, för något fixt n, alla defekta rutnät G_n kan tesselleras, oberoende av den felande rutans läge. Målet är att visa, hur även G_{n+1} kan tesselleras. Splittra G_{n+1} i fyra lika stora kvadratiska rutnät; se Figur 1. Vart och ett av dessa har sidlängd 2^n rutor. I en av dessa fyra kvadrater ligger den saknade rutan från G_{n+1}, i figuren den övre högra. Denna delkvadrat är alltså av typ G_n, och kan därmed, enligt induktionsantagandet, övertäckas med triomino-brickor.

Återstår de tre övriga delkvadraterna, som vi inte kan betrakta som G_n, eftersom de inte är defekta. Men detta kan avhjälpas på följande vis. Tag en triomino och placera den i mitten, där de tre hela kvadraterna möts (se figuren). Triomino-brickan täcker då precis en ruta i varje kvadrat och, om vi bortser från dessa tre rutor, så har vi åstadkommit tre defekta rutnät av typ G_n. Dessa kan, åter enligt induktionsantagandet, övertäckas med triomino-brickor. Alltså kan vi täcka hela brädet G_{n+1}, oavsett den felande rutans position.

Enligt Induktionsprincipen kan då varje defekt $2^n \times 2^n$-bräde tesselleras.



FIGUR 1: Kvadratiskt rutnät av storlek $2^{n+1} \times 2^{n+1}$.



FIGUR 2: Tornen i Hanoi ($n = 3$).

Exempel 5.
Tornen i Hanoi är ett klassiskt pussel. Det består av tre vertikala pinnar. På den ena är trädda n hålförsedda runda skivor, olika stora och sorterade i storleksordning, med den minsta skivan överst och den största skivan underst. Målet är, att överflytta tornet av skivor till en av de tomma pinnarna. Härvidlag flyttas en skiva i taget, och en större skiva må aldrig placeras över en mindre. Figur 2 visar en lösning i sju drag med $n = 3$ skivor.

Vi bevisar med induktion, att $2^n - 1$ drag räcker för att flytta n skivor till en annan pinne. Basfallet $n = 1$ är klart. En skiva kan flyttas till en annan pinne på $1 = 2^1 - 1$ drag.

Antag, att vi redan vet, att $2^n - 1$ drag räcker för att flytta n skivor, för något visst värde på n. Betrakta ett spel med $n + 1$ skivor, där skivorna startar på en pinne döpt till A (de övriga pinnarna döper vi till B och C). Vi påstår, att detta kan lösas på $2^{n+1} - 1$ drag. Spelet löses i tre faser.

- Flytta först de n översta skivorna från pinne A till pinne B. Det kan ske på precis $2^n - 1$ drag enligt induktionsantagandet.
- Flytta sedan den understa, största skivan, i ett enda drag, från A till C.
- Flytta slutligen de n skivorna på B över till C. Det kan ske på precis $2^n - 1$ drag enligt induktionsantagandet.

Sammantaget kan alltså de n skivorna flyttas från A till C i

$$(2^n - 1) + 1 + (2^n - 1) = 2^{n+1} - 1$$

drag. Enligt Induktionsprincipen kan då spelet med n skivor lösas i precis $2^n - 1$ drag för varje positivt heltal n, och beviset är färdigt.

Enligt sägnen står de ursprungliga Tornen i Hanoi i ett uråldrigt indiskt tempel. Skivorna, sextiofyra till antalet, är av rent guld, kringgärdande trenne pelare huggna i diamant. Skivorna är i ständig rörelse och flyttas runt, en i sekunden, av det brahminska prästerskapet. Det står skrivet i de heliga skrifterna, att då hela tornet överflyttats till en annan pelare, skall världens undergång tima.

§5. Olikheter

Exempel 6.
För ett heltal $n \geq 4$, visar vi olikheten

$$3^n > n^3.$$

Olikheten stämmer för $n = 4$, ty $3^4 = 81 > 64 = 4^3$. Detta är basen för induktionen. Antag olikheten gälla för ett visst värde på n, så att $3^n > n^3$. Då vill vi visa att den också gäller för $n+1$, d.v.s. att $3^{n+1} > (n+1)^3$. Räkningen

$$\begin{aligned} 3^{n+1} = 3 \cdot 3^n &> 3 \cdot n^3 = n^3 + 2n^2 \cdot n && \text{(induktionsantagandet)} \\ &\geq n^3 + 2n^2 \cdot 4 = n^3 + 3n^2 + 3n^2 + 2n^2 && (n \geq 4) \\ &> n^3 + 3n^2 + 3n + 2n^2 && (n^2 > n \text{ för } n \geq 4) \\ &> n^3 + 3n^2 + 3n + 1 = (n+1)^3 && (2n^2 > 1 \text{ för } n \geq 4) \end{aligned}$$

visar induktionssteget. Av Induktionsprincipen följer det, att olikheten är sann för alla heltal $n \geq 4$.

Exempel 7.
Vi griper oss an ett litet mera exotiskt uppdrag, nämligen att försöka bestå den oändligt itererade kvadratroten

$$\sqrt{a + \sqrt{a + \sqrt{a + \sqrt{a + \cdots}}}}, \qquad a > 0,$$

med en mening. Vi löser problemet som följer. Rekursionsformeln

$$r_{n+1} = \begin{cases} \sqrt{a + r_n} & \text{om } n \geq 1 \\ 0 & \text{om } n = 0, \end{cases}$$

ger upphov till talföljden

$$r_0 = 0, \qquad r_1 = \sqrt{a}, \qquad r_2 = \sqrt{a + \sqrt{a}}, \qquad r_3 = \sqrt{a + \sqrt{a + \sqrt{a}}}, \qquad \ldots,$$

och det vore mer än rimligt, att definiera

$$\sqrt{a + \sqrt{a + \sqrt{a + \sqrt{a + \cdots}}}} = \lim_{n \to \infty} r_n,$$

förutsatt förstås, att detta gränsvärde existerar. Det bevisar vi strax genom att nyttja den sats från analysen, vilken anför, att *en växande, uppåt begränsad talföljd har ett gränsvärde.* Givet existensen av gränsvärdet $\alpha = \lim_{n\to\infty} r_n$, kan vi beräkna det genom att låta $n \to \infty$ i rekursionsformeln $r_{n+1} = \sqrt{a + r_n}$. Vi får $\alpha = \sqrt{a + \alpha}$, så att det icke-negativa talet α löser ekvationen $\alpha^2 = a + \alpha$; sålunda är

$$\alpha = \frac{1}{2} + \sqrt{a + \frac{1}{4}}.$$

Vi visar nu, att talföljden är uppåt begränsad av α, alltså $r_n \leq \alpha$. Det gör vi med induktion. Det är klart att olikheten gäller för $n = 0$, ty

$$r_0 = 0 < \frac{1}{2} + \sqrt{a + \frac{1}{4}} = \alpha.$$

Antag den övre begränsningen $r_n \leq \alpha$ vara sann för ett visst index n. Samma övre begränsning gäller ävenledes för nästa index $n+1$, vilket inses genom räkningen

$$r_{n+1} = \sqrt{a + r_n} \leq \sqrt{a + \alpha} = \alpha.$$

Alltså har vi visat induktionssteget. Induktionsprincipen ger då, att $r_n \leq \alpha$ gäller för alla n.

Låt oss nu undersöka, huruvida följden växer. Detta innebure att

$$r_{n+1} = \sqrt{a + r_n} \geq r_n \quad \Leftrightarrow \quad a + r_n \geq r_n^2 \quad \Leftrightarrow \quad r_n^2 - r_n - a \leq 0$$

för varje n. Definiera funktionen

$$f(x) = x^2 - x - a = \left(x - \frac{1}{2}\right)^2 - \frac{1}{4} - a,$$

med två skilda reella nollställen $\frac{1}{2} \pm \sqrt{a + \frac{1}{4}}$. Olikheten ovan är sann om och endast om $f(r_n) \leq 0$, vilket är sant precis då r_n ligger mellan nollställena, alltså

$$\frac{1}{2} - \sqrt{a + \frac{1}{4}} \leq r_n \leq \frac{1}{2} + \sqrt{a + \frac{1}{4}} = \alpha.$$

Den övre begränsningen har allaredan visats. Den undre begränsningen gäller trivialt, ty uppenbarligen är

$$\frac{1}{2} - \sqrt{a + \frac{1}{4}} < 0 \leq r_n.$$

Talföljden är alltså både växande och uppåt begränsad och har därför gränsvärdet α enligt ovan. Sålunda är

$$\sqrt{a + \sqrt{a + \sqrt{a + \sqrt{a + \cdots}}}} = \alpha = \frac{1}{2} + \sqrt{a + \frac{1}{4}}.$$

Gyllene snittet, som ständigt tycks ploppa upp i alla möjliga märkvärdiga sammanhang, är

$$\sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + \cdots}}}} = \frac{1 + \sqrt{5}}{2}.$$

Ett annat exempel är

$$\sqrt{2 + \sqrt{2 + \sqrt{2 + \sqrt{2 + \cdots}}}} = 2.$$

Ibland krävs det två steg bakåt för att kunna gå ett steg framåt:

Induktionsprincipen (variant). Låt $P(n)$ vara en familj av påståenden, där $n = 1, 2, 3, \ldots$. Antag följande.

- $P(1)$ och $P(2)$ är sanna.
- Om $P(n-1)$ och $P(n)$ är sanna, så medför det att även $P(n+1)$ är sann.

Då kan vi dra slutsatsen, att $P(n)$ är sann för alla $n = 1, 2, 3, \ldots$.

Exempel 8.
Fibonacci-talen definieras av rekursionsformeln

$$F_n = \begin{cases} 0 & \text{om } n = 0 \\ 1 & \text{om } n = 1 \\ F_{n-1} + F_{n-2} & \text{om } n \geq 2. \end{cases}$$

Fibonacci-följden inleds

$$0,\ 1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ \ldots.$$

Fibonacci-talen tillväxer exponentiellt och har en övre begränsning $F_n < \left(\frac{7}{4}\right)^n$, vilken vi nu bevisar med induktion.

Eftersom varje Fibonacci-tal beror av sina *två* föregångare, måste induktionsbasen och induktionssteget båda vara tudelade. Induktionsbasen blir att verifiera

$$F_0 = 0 < 1 = \left(\frac{7}{4}\right)^0 \qquad \text{jämte} \qquad F_1 < \frac{7}{4} = \left(\frac{7}{4}\right)^1.$$

Nu kommer induktionssteget. Antag, att *både*

$$F_{n-2} < \left(\frac{7}{4}\right)^{n-2} \qquad \text{och} \qquad F_{n-1} < \left(\frac{7}{4}\right)^{n-1}.$$

Då gäller olikheten även F_n, ty

$$\begin{aligned} F_n = F_{n-1} + F_{n-2} &< \left(\frac{7}{4}\right)^{n-2} + \left(\frac{7}{4}\right)^{n-1} \\ &= \left(\frac{7}{4}\right)^{n-2}\left(1 + \frac{7}{4}\right) = \left(\frac{7}{4}\right)^{n-2} \cdot \frac{11}{4} < \left(\frac{7}{4}\right)^{n-2} \cdot \frac{49}{16} = \left(\frac{7}{4}\right)^{n}. \end{aligned}$$

Enligt Induktionsprincipen (varianten ovan) gäller då olikheten för alla naturliga tal n.

Övningar

11.1. Talföljden a_n definieras genom rekursionsformeln

$$a_n = \begin{cases} \dfrac{1}{\sqrt{1 + \frac{1}{a_{n-1}^2}}} & \text{om } n \geq 2; \\ 1 & \text{om } n = 1. \end{cases}$$

(a) Beräkna a_1, a_2, a_3, a_4 och a_5.

(b) Uppställ och bevisa en explicit formel för a_n medelst induktion.

11.2. Låt $a_0 = 1$ och definiera rekursivt talföljden

$$a_n = \frac{a_{n-1}}{1 + a_{n-1}}, \qquad n = 1, 2, 3, \ldots$$

Skriv upp några av de första termerna i denna följd, gissa en allmängiltig explicit formel för a_n, samt visa sedan, att gissningen är riktig (eller revidera gissningen till något mindre kasst…).

11.3. Definiera en talföljd genom rekursionsformeln

$$a_1 = 3, \quad a_2 = 5, \qquad \text{och} \qquad a_{n+1} = 3a_n - 2a_{n-1}, \quad n = 2, 3, 4, \ldots$$

Skriv upp några av de första termerna i denna följd, hypotetisera en allmängiltig formel för a_n, och visa sedan denna.

11.4. Visa summaformeln

$$\sum_{k=1}^{n} (k+1)(k+5) = \frac{1}{6} n(n+7)(2n+7)$$

för alla positiva heltal n.

11.5. Visa identiteten

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{n(n+1)} = \frac{n}{n+1}.$$

för alla positiva heltal n.

11.6. Visa identiteten

$$1^3 + 2^3 + \cdots + n^3 = (1 + 2 + \cdots + n)^2$$

för alla positiva heltal n.

11.7. Visa summaformeln

$$\sum_{k=1}^{n} \frac{k}{2^k} = 2 - \frac{n+2}{2^n}$$

för alla positiva heltal n.

11.8. Visa summaformeln

$$\sum_{k=1}^{n} \frac{1}{1 + 2 + \ldots + k} = \frac{2n}{n+1}$$

för alla positiva heltal n.

11.9. Visa att

$$2^n > n^3$$

för alla heltal $n \geq 10$.

11.10. Visa summaolikheten

$$\sum_{k=1}^{n} \frac{1}{\sqrt{k}} < 2\sqrt{n}$$

för alla positiva heltal n.

11.11. En samling böcker står huller om buller i en bokhylla. I varje steg får du byta plats på två närstående böcker. Visa, att n böcker kan sorteras i alfabetisk ordning med högst $\frac{1}{2}n(n-1)$ sådana byten.

11.12. Visa produktolikheten

$$\frac{1}{2} \cdot \frac{3}{4} \cdot \frac{5}{6} \cdot \cdots \cdot \frac{2n-1}{2n} \leq \frac{1}{\sqrt{3n+1}}$$

för alla positiva heltal n.

11.13. (a) Vilket är talet

$$\sqrt{6 + \sqrt{6 + \sqrt{6 + \sqrt{6 + \cdots}}}} \, ?$$

(b) Vilka tal $b > 0$ kan skrivas på formen

$$\sqrt{a + \sqrt{a + \sqrt{a + \sqrt{a + \cdots}}}} = b \, ?$$

Kapitel 12

Kombinatorik

Kombinatorik, eller antalsräkning, är konsten att räkna saker och ting: möjligheter, urval, kombinationer och arrangemang. Typiska frågor för kombinatoriken är följande.

- Hur många olika "ord" kan bildas genom omkastning av bokstäverna ANTANANARIVO?
- På hur många sätt kan vi välja nattlektyr från bokhyllan, denna och två kvällar framöver?
- Hur många styrelser kan vi utse i en bostadsrättsförening, om den skall vara jämställd och ett visst kverulerande äkta par vägrar ingå samtidigt?

Dessa och liknande problem kommer vi att lösa genom att studera olika sätt att *räkna antal* och ge ett antal fundamentala principer. Teknikerna är användbara på många olika problem, inte minst i tillämpningar i fysik och datalogi. På sedvanligt sätt berör vi inte dessa, utan studerar enklare, mer basala problem, ofta formulerade på det speciellt artificiella (d.v.s. nerdiga) sätt, som matematiker är så förtjusta i.

§1. Matematikerns morgonrutin

Exempel 1.
Vi stiger upp om morgonen och står inför bryderiet att klä på oss. Vi har fyra T-shirts att välja på, tre par byxor fem par skor och har dessutom valet att komplettera vår outfit med en tuff keps. På hur många olika sätt kan vi kombinera dessa plagg?

Vi kan se det som resultatet av en följd av val. Först väljer vi T-shirt. Det kan ske på fyra sätt. Sedan väljs byxor. Oavsett vilken T-shirt vi nyss valde, så kan detta ske på tre sätt, och alltså kan dessa två val ske på $3+3+3+3 = 4\cdot 3$ olika sätt. Sedan väljer vi skor. För varje val av T-shirt och byxor kan detta ske på fem sätt, alltså sammanlagt $4\cdot 3\cdot 5$ sätt. Slutligen väljer vi, om vi skall bära solglasögon eller inte, d.v.s. för varje val av de tidigare plaggen har vi två möjligheter. Svaret blir alltså att det finns

$$4\cdot 3\cdot 5\cdot 2 = 120$$

möjliga klädkombinationer. Vi klädde förstås på oss i en viss ordning. Hade vi startat i någon annan ände, t.ex. med att välja keps eller inte, så hade faktorerna ovan kommit i en annan ordning, men produkten hade blivit densamma, således samma antal möjliga klädkombinationer. (Vissa praktiska svårigheter kan dock vara förknippade med att ta på kepsen före T-shirten.)

Vi formulerar tekniken i exemplet som en allmän idé.

Sats 12.1: Multiplikationsprincipen

En följd av k val, där det första beslutet kan träffas på m_1 sätt, det andra på m_2 sätt o.s.v., kan sammantaget ske på $m_1 m_2 \cdots m_k$ olika sätt, förutsatt att antalet av valen är sinsemellan oberoende.

Exempel 2.
Du skall inreda ett rum med en färg på väggarna och en färg till möblerna, med 100 färgnyanser att välja bland. Inredningen kan då ske på $100 \cdot 100 = 10\,000$ olika sätt.

Exempel 3.
Antalet "ord" med fem bokstäver, valda från det svenska alfabetet med 29 bokstäver är enligt Multiplikationsprincipen 29^5. Den första bokstaven kan väljas på 29 sätt, den andra på 29 sätt, och så vidare.

Ett "ord" är för matematikern synonymt med bokstavskombination. Det behöver alltså inte betyda någonting, varken på svenska eller något annat språk. Naturligtvis kommer de flesta av dessa 29^5 "ord" inte att betyda någonting alls (t.ex. QWERT).

Exempel 4.
Du skall färgkoda 10 000 olika sorters elektriska motstånd, ungefär som registreringsskyltar på bilar, men med färger istället för bokstäver och siffror. Detta skall ske medelst små färgade fyrkanter i fyra färger: gult, blått, grönt, rött. Hur många fyrkanter måste tryckas på motstånden?

Om du bara använder två fyrkanter är antalet möjligheter $4^2 = 16$, med tre fyrkanter $4^3 = 64$. Eftersom $4^6 = 4096$, men $4^7 = 16384$, är svaret är alltså att sju färgade fyrkanter räcker för att särskilja motstånden.

Låt oss vända på problemet. Hur många färger behövs för att fem fyrkanter skall räcka? Kalla antalet färger för n. Då skall alltså $n^5 \geq 10\,000$, eller $n \geq 10000^{\frac{1}{5}} \approx 6.30957$. Svaret är alltså att det behövs minst sju färger för att fem fyrkanter skall vara nog.

§2. Arrangemang

Exempel 5.
På hur många sätt kan vi ordna sex olika böcker i en bokhylla? Vi har 6 val för platsen längst till vänster. När vi valt en bok att ställa på den platsen, har vi 5 val till nästa plats, 4 till den tredje o.s.v. Totalt har vi

$$6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 6! = 720$$

möjligheter att ordna böckerna.

Tre av böckerna är på svenska, tre på franska. Låt oss säga, att vi vill ställa dem så, att språken alternerar i bokhyllan. På hur många sätt låter det sig göras? Det finns två sätt att välja hur språken skall stå: antingen börjar vi med en svensk bok eller med en fransk. Då är språken, så att säga, utplacerade. Inom den svenska gruppen kan de tre böckerna ordnas på 3! sätt; likaså inom den franska gruppen. Antalet möjligheter är $2 \cdot (3!)^2 = 72$ enligt Multiplikationsprincipen.

Om vi har n objekt, vilka som helst, kan vi med samma resonemang som i exemplet ordna dem i en rad på $n! = n \cdot (n-1) \cdot \cdots \cdot 2 \cdot 1$ olika sätt. Det finns n möjligheter att besätta den första platsen. När denna är tillsatt, finns det $n-1$ möjligheter för den andra platsen, och så vidare.

Sats 12.2

En samling av n olika objekt kan ordnas på rad på $n!$ olika sätt.

Vi kan alltså se $n!$ som den kombinatoriska funktion, som ger svaret till frågan "på hur många sätt kan n föremål ordnas i en rad?". Det är förresten det som motiverar, ur en matematiknerds skruvade perspektiv, konventionen $0! = 1$. Vi kan ju placera 0 studenter i en bostadskö på ett enda unikt sätt — genom att inte placera *någon alls* i kön.

Exempel 6.
Hundra personer strömmar in på Systembolaget i Sundbyberg och ställer sig i den (naturligtvis) enda kön. Detta kan ske på $100! \approx 10^{158}$ olika sätt. Vi kan se hur lördagskön på Systemet i Sundbyberg platsar i kosmiska sammanhang; Wikipedia (denna postmoderna Gröngölingsbok) skattar antalet atomer i universum till jämförelsevis försumbara 10^{80}...

Nå, för att komplicera problemet något, så befinner sig både herr och fru von Tratt bland kunderna. De vägrar absolut stå i kön *intill* varandra. Hur många möjligheter finns det nu?

Enklast är att börja i den andra änden genom att först räkna de köer, där von-Trattarna verkligen står intill varandra. Vi låtsas därför att vi klistrar ihop herrn och frun, betraktande dem som en odelbar enhet. Då har vi nittionio "personer" att ordna på led, vilket kan ske på $99!$ olika sätt. Inom sin lilla bubbla måste dock herrn och frun kivas om, vem som skall stå främst och vem som skall stå bakerst. Det finns här alltså 2 möjligheter. Antalet köer med von-Trattarna intill varandra är därför enligt Multiplikationsprincipen $2 \cdot 99!$. Detta subtraherar vi från det totala antalet köer, vilket ger

$$100! - 2 \cdot 99!$$

köer, där herr och fru von Tratt besparas obehaget av varandras sällskap.

Exempel 7.
Hur många "ord" kan bildas med de fem bokstäverna VIETA? Svaret är $5! = 120$.

Det var enkelt. För enkelt, faktiskt. Vad händer, om vissa bokstäver är lika?

Exempel 8.
Hur många "ord" kan bildas genom omkastning av de fem bokstäverna i PAPPA? Problemet nu är att bokstäverna inte är olika. Det finns alltså färre än de $5!$ möjligheterna vi hade ovan.

Man kan lösa problemet genom att först låtsas alla bokstäver vara olika. Säg att vi indexerar dem som $\mathrm{P}_1\mathrm{A}_1\mathrm{P}_2\mathrm{P}_3\mathrm{A}_2$. Då kan verkligen $5!$ olika ord bildas genom omkastning. Om vi tar bort märkningen på P_1, P_2 och P_3, kommer vissa ord som tidigare var olika att sammanfalla, t.ex. kommer $\mathrm{P}_1\mathrm{P}_2\mathrm{P}_3\mathrm{A}_1\mathrm{A}_2$ och $\mathrm{P}_2\mathrm{P}_3\mathrm{P}_1\mathrm{A}_1\mathrm{A}_2$ båda att bli $\mathrm{PPPA}_1\mathrm{A}_2$. Hur många olika ord klumpas samman, när vi avlägsnar märkningen på P:na? Det finns $3! = 6$ ord som sammanfaller, för det finns så många sätt att omordna P_1, P_2, P_3 sinsemellan. Det blir alltså nu $5!/3!$ olika ord kvar. Sedan tar vi bort märkningen på A_1 och A_2, vilket gör att kvarvarande ord kommer att klumpas ihop två och två. Resultatet blir alltså, att man bara kan bilda $\frac{5!}{3!2!} = 10$ ord.

För att kanske tydliggöra resonemanget, eller i vart fall förklara det på annat sätt, kan vi också resonera baklänges. Kalla antalet ord som kan bildas av bokstäverna PAPPA för N. Om vi nu märker upp bokstäverna PPP till P_1, P_2, P_3, kan vi ur varje ord bilda $3!$ nya, eftersom det, givet ett ord utan märkning, finns $3!$ sätt att fördela etiketterna 1, 2, 3 på de tre P:na. Det finns alltså $3!N$ ord med bokstäverna $\mathrm{P}_1\mathrm{P}_2\mathrm{P}_3\mathrm{AA}$. Av varje sådant ord kan vi sedan bilda två nya genom att sätta index på de två A:na. Det finns alltså $2!3!N$ ord med bokstäverna $\mathrm{P}_1\mathrm{A}_1\mathrm{P}_2\mathrm{P}_3\mathrm{A}_2$. Men nu är alla bokstäver olika, varav $3!2!N = 5!$ och $N = \frac{5!}{3!2!}$.

Den allmänna formeln ser ut så här:

Sats 12.3

En samling bestående av n_1 sinsemellan lika objekt, n_2 sinsemellan lika objekt, ..., n_k sinsemellan lika objekt, kan ordnas på rad på

$$\frac{(n_1 + n_2 + \cdots + n_k)!}{n_1! n_2! \cdots n_k!}$$

olika sätt.

För exemplet ovan ger alltså formeln direkt $\frac{(3+2)!}{3!2!}$ olika ord.

Exempel 9.
Hur många ord kan bildas genom omkastning av bokstäverna i ANTANANARIVO? Det finns fyra A, tre N och dessutom bokstäverna T, R, I, V, O en gång vardera. Enligt ovan är svaret

$$\frac{12!}{4!3!1!1!1!1!1!} = \frac{12!}{4!3!}.$$

Här är för övrigt ett ord, som händelsevis betyder någonting. ANTANANARIVO är huvudstaden på Madagaskar, mycket populär i kombinatoriska kretsar. WOOLLOOMOOLOO, en stadsdel i Sydney, är också föremål för intensivt kombinatoriskt studium.

§3. Ordnade urval

Exempel 10.
En tävling med sju deltagare skall resultera i en prislista med en guld-, silver- och bronsmedaljör. Det givs 7 möjligheter att välja förstapristagaren, därefter 6 att välja andrapristagaren och 5 att välja tredjepristagaren. Totalt alltså $7 \cdot 6 \cdot 5 = 210$ möjligheter enligt Multiplikationsprincipen.

Vi kan se detta som att vi har valt ut tre personer i ordning bland de sju deltagarna. Ett sådant val kallas ett **ordnat urval**.

Mer allmänt, låt oss välja k objekt, i ordning, från en samling med n objekt. Det första objektet kan väljas på n sätt, det andra på $n-1$ sätt, och så vidare. Det k:te objektet kan väljas på $n-k+1$ sätt. Antalet möjligheter är enligt Multiplikationsprincipen

$$\begin{aligned} &n(n-1)\cdots(n-k+1) \\ &\quad= \frac{\Big(n\cdot(n-1)\cdots(n-k+1)\Big)\Big((n-k)\cdot(n-k-1)\cdot\cdots\cdot 2\cdot 1\Big)}{(n-k)(n-k-1)\cdot\cdots\cdot 2\cdot 1} \\ &\quad= \frac{n\cdot(n-1)\cdot\cdots\cdot 2\cdot 1}{(n-k)\cdot(n-k-1)\cdot\cdots\cdot 2\cdot 1} = \frac{n!}{(n-k)!}. \end{aligned}$$

Sats 12.4

Antalet ordnade urval av k objekt från n givna är

$$n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}.$$

Exempel 11.
Kom ihåg de sex böckerna i bokhyllan, tre svenska och tre franska. På hur många sätt kan vi välja ut tre av dessa att läsa, en i dag, en i morgon och en i övermorgon? Detta är ett ordnat urval. Vi kan välja våra böcker på $6 \cdot 5 \cdot 4 = \frac{6!}{3!} = 120$ olika sätt.

Exempel 12.
Bara för att klura till bokproblemet litet, hur många möjligheter finns det, att läsa två böcker från det ena språket och en från det andra? Åter är det enklast att tänka utifrån de oönskade möjligheterna. Vi vill alltså inte välja alla de tre svenska böckerna, vilket kan göras på 3! sätt. Inte heller vill vi välja samtliga tre franska böcker, vilket också kan ske på 3! sätt. Svaret är differensen

$$120 - 2 \cdot 3! = 108.$$

§4. Oordnade urval

Exempel 13.
Åter tänker vi oss en tävling med sju deltagare. På hur många sätt kan vi välja ut tre, som skall erhålla stipendium? Alla får samma stipendium, så här är det alltså fråga om att välja ut en grupp om tre personer, där ordningen är oväsentlig.

Vi drar oss till minnes, att antalet prislistor med en första-, en andra- och en tredjepristagare var $7 \cdot 6 \cdot 5 = \frac{7!}{4!}$. För varje grupp av tre pristagare finns det 3! sådana listor som ger identiska uppsättningar stipendiater, när vi glömmer ordningen inför stipendieutdelningen. Svaret blir alltså $\frac{7!}{4!3!}$.

Detta är ett exempel på ett **oordnat urval**. Här handlar det om att välja ut en samling eller mängd objekt från en given uppsättning, utan hänsyn tagen till ordningen. Antalet oordnade urval av k objekt från en samling av n objekt betecknas med **binomialkoefficienten**

$$\binom{n}{k},$$

vilket utläses "n över k" eller "n välj k". Vi har alltså t.ex. $\binom{7}{4} = \frac{7!}{4!3!}$.

Den allmänna formeln för binomialkoefficienterna, som bevisas på samma sätt som i det inledande exemplet, är

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

Antalet *ordnade* urval är nämligen, enligt vad vi redan vet, $\frac{n!}{(n-k)!}$. Men $k!$ sådana ordnade urval blir ju identiska, så fort vi glömmer ordningen i vilken vi valde objekten; de kan ju sinsemellan ordnas just på $k!$ sätt. Denna måste därför divideras bort från antalet ordnade urval.

Sats 12.5

Antalet oordnade urval av k objekt från n givna är

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Exempel 14.
I ett hus i Sundbyberg bor fjorton personer: åtta kvinnor och sex män. Till bostadsrättsföreningen vill de utse en styrelse, bestående av fyra personer. Den vill alltså göra ett oordnat urval av fyra personer bland fjorton, vilket kan ske på

$$\binom{14}{4} = \frac{14!}{4!10!} = 1001$$

sätt.

Kräver vi en jämställd styrelse (två kvinnor, två män), finns det färre valmöjligheter. Kvinnorna och männen måste nu väljas var för sig. Två kvinnor av åtta kan väljas på $\binom{8}{2}$ sätt, två män av sex kan väljas på $\binom{6}{2}$ sätt, och antalet möjligheter är enligt Multiplikationsprincipen

$$\binom{8}{2}\binom{6}{2} = 420.$$

Men så har vi då makarna von Tratt. Under förtäckta hot om stambyte vägrar de absolut att medverka i föreningen samtidigt, vilket förstås ställer till en hel del bryderi. För att räkna dessa möjligheter, är det återigen lämpligt att tänka subtraktivt. Vilka styrelser är icke-önskevärda? Dem i vilka båda makarna ingår. Då återstår en kvinna att välja bland de övriga sju, och vidare en man att välja bland de övriga fem, vilket ger $\binom{7}{1}\binom{5}{1} = 7 \cdot 5 = 35$

$\binom{0}{0}$	1
$\binom{1}{0}\ \binom{1}{1}$	1 1
$\binom{2}{0}\ \binom{2}{1}\ \binom{2}{2}$	1 2 1
$\binom{3}{0}\ \binom{3}{1}\ \binom{3}{2}\ \binom{3}{3}$	1 3 3 1
$\binom{4}{0}\ \binom{4}{1}\ \binom{4}{2}\ \binom{4}{3}\ \binom{4}{4}$	1 4 6 4 1
.....	.....

FIGUR 1: Pascals triangel.

förbjudna styrelser. Svaret är därför

$$\binom{8}{2}\binom{6}{2} - \binom{7}{1}\binom{5}{1} = 420 - 35 = 385.$$

§5. Pascals triangel

Binomialkoefficienterna $\binom{n}{k}$ inordnas gärna i formen av en triangel med namnet *Pascals triangel*, trots att den var känd av de gamla kineserna och studerad även i Europa långt före Pascal (1600-talet). Se Figur 1. Talet $\binom{n}{k}$, där $0 \le k \le n$, återfinns på plats k i rad n (observera att både k och n räknas från 0).

Tre observationer låter sig omedelbart göras. Varje rad i triangeln börjar och slutar med 1. Triangeln är vertikalt symmetrisk (kring mittlinjen). Dessutom kan varje tal fås som summan av de två talen närmast ovanför, t.ex. är $\binom{4}{2} = \binom{3}{1} + \binom{3}{2}$. Vi formulerar dessa egenskaper i en sats.

Sats 12.6

Binomialkoefficienterna har följande egenskaper.

(a) $\binom{n}{0} = \binom{n}{n} = 1.$

(b) $\binom{n}{k} = \binom{n}{n-k}.$

(c) $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ (Pascals triangelschema).

Bevis

(a) Vi har formeln $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. Speciellt är alltså $\binom{n}{0} = \frac{n!}{0!(n-0)!} = 1$ och $\binom{n}{n} = \frac{n!}{n!(n-n)!} = 1$. Detta svara emot att det finns precis ett sätt att välja 0 objekt av n givna (välj inga alls) och precis ett sätt att välja n objekt av n givna (välj alla). (Det är för övrigt nödvändigheten att $\binom{n}{0} = 1$ som gör att man väljer definitionen $0! = 1$.)

(b) Ett kombinatoriskt bevis lyder som följer. Studera antalet sätt att välja k av våra n kompisar att leka med efter skolan? Svaret är förstås $\binom{n}{k}$. Men vi kan lika gärna välja, vilka $n-k$ av våra kompisar, som *inte* skall få vara med och leka. Detta går på $\binom{n}{n-k}$ sätt. Eftersom de båda talen är svaren på samma kombinatoriska problem, måste de vara

lika. Ett mer intetsägande algebraiskt bevis är

$$\binom{n}{n-k} = \frac{n!}{(n-k)!k!} = \frac{n!}{k!(n-k)!} = \binom{n}{k}.$$

(c) Här är ett kombinatoriskt bevis för denna formel. Vi vet, att $\binom{n}{k}$ räknar antalet sätt att göra ett oordnat urval av k personer från en grupp av n personer, säg en lärare och $n-1$ elever. De k personerna kan nu antingen inkludera läraren eller ej. I det första fallet skall man välja ut $k-1$ bland de $n-1$ eleverna för att sammanlagt få ihop k personer. Detta kan göras på $\binom{n-1}{k-1}$ sätt. I det andra fallet skall man välja alla k personerna från de $n-1$ eleverna, vilket kan göras på $\binom{n-1}{k}$ sätt. Adderar vi dessa två varandra uteslutande möjligheter, får vi just formeln i (c). Ett algebraiskt bevis, med grund i formeln $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ är inte heller särskilt komplicerat.

§6. Bevis för Binomialsatsen

Vi betraktar nu följande formler (de är säkert välkända vid det här laget):

$$\begin{aligned}(x+y)^2 &= x^2 + 2xy + y^2\\ (x+y)^3 &= x^3 + 3x^2y + 3xy^2 + y^3\\ (x+y)^4 &= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4\end{aligned}$$

Koefficienterna för $(x+y)^n$ bär en kuslig likhet rad nummer n i Pascals triangel. Mera precist, så verkar koefficienten för $x^k y^{n-k}$ vara lika med $\binom{n}{k}$. Låt oss se varför och hur detta fungerar.

Uttrycket

$$(x+y)^3 = (x+y)(x+y)(x+y)$$

kan utvecklas genom att man, på alla möjliga sätt, väljer ut x eller y ur varje parentes, multiplicerar ihop dessa, och summerar produkterna. Man får alltså

$$\begin{aligned}(x+y)^3 = x\cdot x\cdot x + x\cdot x\cdot y + x\cdot y\cdot x + x\cdot y\cdot y\\ + y\cdot x\cdot x + y\cdot x\cdot y + y\cdot y\cdot x + y\cdot y\cdot y.\end{aligned}$$

Alla termer är inte olika, t.ex. är $xxy = xyx = yxx$. Antalet termer som är en produkt av två x och ett y, d.v.s. de som blir x^2y, är detsamma som antalet sätt, att välja ut två av de tre parenteserna, ur vilka bokstaven x skall tagas. Detta kan göras på $\binom{3}{2} = 3$ sätt.

Resonemanget går att göra allmängiltigt. Om man multiplicerar ihop faktorerna i

$$(x+y)^n = (x+y)(x+y)\cdots(x+y),$$

så får vi termer som är produkter av sammanlagt n stycken bokstäver, vardera ett x eller ett y. Antalet termer med k stycken x och $n-k$ stycken y, d.v.s. de som blir $x^k y^{n-k}$, är antalet sätt att välja ut de k parenteser $(x+y)$, ur vilka vi plockar bokstaven x. Från övriga $n-k$ parenteser väljs bokstaven y. Detta kan göras på $\binom{n}{k}$ sätt. Koefficienten för $x^k y^{n-k}$ är alltså $\binom{n}{k}$. Följande sats är visad.

Sats 12.7: Binomialsatsen

Låt x och y vara komplexa tal, och låt n vara ett naturligt tal. Då gäller att

$$(x+y)^n = \binom{n}{0}y^n + \binom{n}{1}xy^{n-1} + \cdots + \binom{n}{n-1}x^{n-1}y + \binom{n}{n}x^n = \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}.$$

Exempel 15.
Betrakta åter Pascals triangel. Vi gjorde ovan tre uppenbara upptäckter i denna. En något mindre uppenbar observation är, att summan av talen i rad n tycks vara lika med 2^n, alltså

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

Vi bevisar denna ekvation på två sätt.

Kombinatoriskt kan vi studera problemet, att välja ut några objekt av n givna. Observera: *några*. Vi har inte specificerat antalet. Att välja ut ett specifikt antal k av objekten, går ju som bekant på $\binom{n}{k}$ sätt, så svaret är alltså summan $\sum_{k=0}^{n} \binom{n}{k}$ av alla dessa möjligheter. Vi har nu räknat med möjligheten att välja inga ($k = 0$) och alla objekt ($k = n$). Men vi kan också tänka som så, att vi för vart och ett av de n objekten specificerar, huruvida det skall väljas eller ej. För första objektet har vi då två möjligheter (väljas eller ratas), för andra objektet två möjligheter, och så vidare. Totala antalet möjligheter är enligt Multiplikationsprincipen 2^n, och vi är klara.

Det algebraiska beviset är elegant och sofistikerat, men totalt intetsägande. Nyttja Binomialsatsen och sätt $x = y = 1$:

$$2^n = (1+1)^n = \sum_{k=0}^{n} \binom{n}{k} 1^k 1^{n-k} = \sum_{k=0}^{n} \binom{n}{k}.$$

Övningar

12.1. Hur många bilar kan det maximalt finnas i Sverige, om varje bil skall ha ett unikt bilnummer? Ett bilnummer består av tre bokstäver, vilka som helst, följda av tre siffror, vilka som helst. Till skillnad från Bilregistret utesluter vi inte vissa bokstavskombinationer (som DUM, FAN, GUD, etc.). Eftersom vi just nu känner oss ovanligt lata, antar vi ett alfabet om trettio bokstäver.

12.2. Hur många tal finns det, som har högst fem siffror i tiosystemet? Precis fem siffror?

12.3. Hur många fembokstaviga "ord", kombinerade av bokstäver från svenska alfabetet, finns det?

12.4. Hur många sexsiffriga tal kan bildas med hjälp av siffrorna 0, 1 och 2, om talen skall innehålla minst en nolla?

12.5. Hur många olika "ord" på nio bokstäver (meningslösa bokstavskombinationer) kan bildas genom omkastning av bokstäverna

SASSAFRAS ?

12.6. (a) Hur många "ord" kan bildas genom omkastning av de sex bokstäverna

SYMBOL ?

(b) I hur många av dessa ord står Y inte omedelbart mellan S och M?

12.7. (a) På hur många olika sätt kan åtta personer ställa sig på led?

(b) På hur många sätt kan de placeras kring ett runt bord? Två utplaceringar som endast skiljer sig med en rotation anses vara lika.

12.8. (a) Hur många olika åttabokstaviga "ord" kan bildas av bokstäverna

GEOMETRI ?

(b) Hur många trebokstaviga ord kan bildas?

12.9. Som en förberedelse för en ny läroplan, bestämmer myndigheterna sig (i en oemotsägligt blodig diktatur) för, att inventera det numera rätt åldersstigna beståndet av naturliga tal. De ger därför Matematiska institutionen i uppdrag att räkna alla dessa, samt kassera de som är alltför slitna. Institutionen skall för detta nyanställa tre räknare, och dessa kan väljas från hundra sökanden. På hur många sätt kan urvalet ske?

12.10. Elsa väljer på caféet mellan åtta olika smörgåsar.

(a) På hur många sätt kan hon välja fyra smörgåsar av dessa?

(b) Hur många sätt finns det, om hon dessutom skall välja ordningen, i vilken hon äter smörgåsarna?

12.11. Elsa har åtta identiska chokladbitar och åtta identiska kolor att smaska i sig under fredagsmyset. På hur många sätt kan hon välja ordningen mellan dessa?

12.12. I en stryktipsrad tippar man, för vardera av tretton olika matcher, endera 1, X eller 2. Hur många olika stryktipsrader finns det med sju 1, två X och fyra 2?

12.13. En kortlek består, som bekant, av 52 kort i fyra olika färger, vardera med 13 kort i olika valörer. I en pokergiv får en spelare en hand om fem kort (ordningsföljden är oväsentlig).

(a) Hur många olika pokerhänder finns det?

(b) Hur många av dessa innehåller ett fyrtal, d.v.s. fyra kort i samma valör?

(c) Hur många händer innehåller två par, d.v.s. två kort i en valör, två kort i en annan valör, samt ett av en tredje valör?

12.14. (a) På hur många sätt kan en mängd av fyra *olika* bokstäver väljas ur svenska alfabetet (med 29 bokstäver)?

(b) Hur många olika ord kan bildas med fyra bokstäver, ej nödvändigtvis olika, ur svenska alfabetet?

(c) Hur många olika ord kan bildas med fyra bokstäver, där första bokstaven är ett X och andra bokstaven inte är ett Y?

12.15. Vi har nio (identiska) röda och sex (identiska) svarta bollar. På hur många sätt kan vi arrangera dessa bollar i en rad, givet att de första r bollarna skall vara röda? (För vilka r är problemet lösbart?)

12.16. Elsa bakar. På en plåt med elva bullar, av litet olika färg och façon, skall hon garnera sex med glasyr och fem med pärlsocker (en bulle kan få bäggedera).

(a) På hur många sätt kan hon göra detta?

(b) Hur många sätt finns det, om minst två bullar skall få både glasyr och pärlsocker?

12.17. Elsa skall intaga middag å restaurant. Denna erbjuder fjorton förrätter, arton huvudrätter och åtta efterrätter.

(a) Hur många trerättersmenyer kan hon välja bland?

(b) Elsa har ganska god aptit (och ganska gott om pengar), och bestämmer sig för att äta tre förrätter och fyra huvudrätter. Hur många sätt finns det för henne att välja dessa rätter (ordningen är oväsentlig och de rätter hon väljer skall vara olika)?

(c) Vid närmare granskning visar sig fem av huvudrätterna vara vegetariska, och sådana vill Elsa helst undvika. Hur många sätt finns det då för henne att välja tre förrätter och fyra huvudrätter, om hon får äta *högst en* vegetarisk huvudrätt (ordningen är oväsentlig och rätterna skall vara olika)?

12.18. Elsa skall laga middag och väljer länge i sin kokbok mellan sexton maträtter. Åtta av dem serveras med Béarnaise-sås, fyra med majonnäs, och till fyra rekommenderas bägge såserna.

På hur många sätt kan hon välja ut fyra rätter att äta, om ordningen är oväsentlig? Elsa är sugen på både Béarnaise-sås och majonnäs.

12.19. Bevisa att

$$\sum_{k=0}^{n} 2^k \binom{n}{k} = 3^n.$$

12.20. Beräkna summan

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k}.$$

12.21. Till en fest har värdparet bjudit fyra gifta par. En del personer skakar hand med varandra, men ingen skakar hand med sin partner. Värden frågar alla andra hur många de skakat hand med och får idel olika svar. Hur många hälsade på värdinnan? (Det här är ett tuffare problem endast för elitidrottare, där man inte kan använda sig av teorien.)

Kapitel 13

Satslogik

§1. Utsagor

Matematisk text består av *utsagor* eller *påståenden*, som vi nu skall betrakta närmare. Exempel på matematiska utsagor är följande.

(a) 13 är ett primtal.

(b) 15 är ett primtal.

(c) Varje triangel är rätvinklig.

(d) Varje primtal är udda.

(e) $0 + 1 < 3$.

(f) π är irrationellt.

(g) Det finns inga primtal av formen $2^{2^n} + 1$, där $n \geq 5$.

Matematiska påståenden, som dessa ovan, är *sanna eller falska*. Det kan tyckas banalt, men det gäller ju inte i vardagslivet. (Är Kina en demokrati eller inte?) Matematiska påståenden kan inte tillåtas, genom luddighet i formulering eller begreppsinnebörd, att hamna i någon gråzon mellan sant och falsk, som ju är fallet i vanligt språk.

Påståendena (b), (c) och (d) är falska, och kan lätt motbevisas. Utsagorna (a) och (e) inses lika enkelt vara sanna. Även (f) är sann, men beviset är komplicerat och kläcktes inte förrän på 1700-talet. Status för det sista påståendet (g) ovan är okänd. Trots massivt arbete inom talteorien under de senaste tre århundraden, är det ännu inte känt, huruvida några primtal $2^{2^n} + 1$, så kallade *Fermat-primtal*, existerar för $n \geq 5$. Men utsagan måste vara endera, antingen sann eller falsk.

Några exempel på fraser, som inte är utsagor:

(a) Derivatan av vänstra ledet.

(b) $x^2 + y^2$.

(c) x är mycket större än y.

Det är nämligen omöjligt att tillordna dem ett sanningsvärde "sant" eller "falskt". I (a) och (b) möter oss nominalfraser, ett samling lösrycka termer. Ingenting händer i dessa fraser (verb eller predikat saknas), och de uttalar sig inte om någonting alls. Ingenting påstås. I (c) anträffar vi visserligen ett påstående, men icke ett matematiskt stringent sådant, ty vad skulle det betyda, att x vore *mycket större* än y? Att skillnaden $x - y > 100$ eller $x - y > 10^{99}$? Eller kanske rentav att $\frac{x}{y} > 10^{10^{100}}$? Att vara *större* är ett absolut begrepp och giltig matematisk terminologi, men att vara *mycket större* relativiserar bort den matematiska stringensen.

Vi skall också syssla med så kallade *öppna utsagor*, som innehåller en eller flera variabler och vars sanningsvärde beror på dessa:

(a) x är ett primtal.

P	$\neg P$
S	F
F	S

TABELL 1: Sanningstabell för negation.

(b) $x < y$.

(c) $x^2 - y^2 = (x + y)(x - y)$.

(d) $x^2 + 1 < 0$.

Här underförstår man, att variablerna må tilldelas värden från någon universalmängd. I (a) förmodas nog x vara ett heltal, eller kanske naturligt tal, medan x och y i (b)–(d) nog antages vara reella tal. Vilken denna mängd är måste framgå av sammanhanget. Först efter att ha tilldelat alla i påståendet ingående variabler värden, kan man undersöka frågan, om påståendet är sant eller falskt. Svaret kan naturligtvis bero på vilka variablernas värden. Utsagorna (a) och (b) ovan beror i allra högsta grad på de aktuella värdena av x och y, medan (c) alltid är sann och (d) alltid falsk.

§2. "Och", ock! ock!, och Icke, sa Nicke!

Givet några utsagor, som vi kommer att beteckna med bokstäver som P, Q, R, ..., så kan vi kombinera dem till mer komplicerade utsagor med hjälp av *logiska konnektiver*. Vi inför symboler för dessa, men det förtjänar att påpekas, att de inte nyttjas alltför mycket utanför den matematiska logiken, som å andra sidan verkligen frossar i dem.

Definition 13.1

Negationen $\neg P$ av en utsaga utläses "icke P". Den definieras vara sann, precis då P är falsk.

Sanningstabellen i Tabell 1 ger en systematisk överblick över sanningshalten för $\neg P$ utifrån sanningshalten för P. (Bokstaven S betecknar "sann", bokstaven F "falsk".) Denna tabell har bara två rader, men det blir strax mer avancerat.

Exempel 1.
Negationen $\neg(x = 1)$ brukar vi vanligen skriva $x \neq 1$. Negationen $\neg(x < 1)$ skriver vi bekvämast $x \geq 1$.

Definition 13.2

Konjunktionen $P \wedge Q$ av två utsagor utläses "P och Q". Den definieras vara sann, precis då både P och Q är det.

Att $P \wedge Q$ är sann precis då både P och Q är sanna, är också är den vanliga användningen i språket. Inga överraskningar där. Sanningstabellen i Tabell 2 ger en systematisk överblick över sanningshalten för $P \wedge Q$ utifrån de fyra olika möjligheterna för sanningsvärdena för de ingående utsagorna P och Q.

Exempel 2.
Vi kan skriva

$$x^2 = 4 \ \wedge \ x \geq 0 \quad \Leftrightarrow \quad x = 2.$$

P	Q	$P \wedge Q$
S	S	S
S	F	F
F	S	F
F	F	F

Tabell 2: Sanningstabell för konjunktion.

P	Q	$P \vee Q$
S	S	S
S	F	S
F	S	S
F	F	F

Tabell 3: Sanningstabell för disjunktion.

§3. Antingen "antingen eller" eller "eller"

Ordet *eller* används av icke-matematiker på ett tvetydigt sätt. I frasen "vän eller fiende" skall det rimligen tolkas som "antingen eller" — det ena eller det andra, men inte gärna bäggedera. Detta är det *exklusiva eller*. Det finns också ett *inklusivt eller*. Du kan t.ex. säga: "Om jag får högsta betyg på tentamen eller vinner på Lotto, så blir det firande och svirande i kväll." Kanske du prickar in alla rätt på både tentamen *och* Lotto, och det verkar rimligt, att firandet och svirandet skulle äga rum även i detta fall. Här öppnas alltså för att båda möjligheterna kan vara sanna samtidigt.

Matematiken, som föraktar de oklarheter och mångtydigheter, varav det vardagliga språket är bemängt, måste besluta sig för (antingen!) den ena eller den andra varianten av *eller*. Den tolkning man har valt som standard är det *inklusiva eller*.

Definition 13.3

Disjunktionen $P \vee Q$ av två utsagor utläses "P eller Q". Den definieras vara sann, precis då endera av P och Q är det (eller kanske bäggedera).

Disjunktionen $P \vee Q$ är sålunda sann, precis när åtminstone något av påståendena P och Q är sant, inklusive fallet, då bägge är sanna. Den är därmed falsk endast när både P och Q är falska. Åter är det behändigt att åskådliggöra de olika fallen i en sanningsvärdestabell, Tabell 3, som visar precis när disjunktionen är sann.

Exempel 3.
Lösningarna till $x^2 = 4$ är ± 2. I vardagligt tal säger vi nog, att lösningarna är $x = 2$ *och* $x = -2$. En stunds eftertanke visar dock, att detta egentligen är nonsens. Utsagan

$$x = 2 \quad \wedge \quad x = -2$$

är alltid falsk, ty x kan förstås inte vara lika med 2 och -2 samtidigt. Korrekt glosa är *eller*, så att ekvationslösningen bör skrivas symboliskt

$$x^2 = 4 \quad \Leftrightarrow \quad x = 2 \ \vee \ x = -2.$$

§4. Implikation och ekvivalens

I samband med ekvationslösning mötte vi implikationen $P \Rightarrow Q$ och ekvivalensen $P \Leftrightarrow Q$. De har samma karaktär som $\vee$ och $\wedge$, i den meningen att man med deras hjälp bygger upp mer komplicerade påståenden från enklare, och faller därför också i kategorien logiska konnektiver.

P	Q	$P \Leftrightarrow Q$
S	S	S
S	F	F
F	S	F
F	F	S

TABELL 4: Sanningstabell för ekvivalens.

Definition 13.4

Ekvivalensen $P \Leftrightarrow Q$ av två utsagor utläses "P är ekvivalent med Q" eller "P om och endast om Q". Den definieras vara sann, precis då P och Q har samma sanningsvärden, det vill säga de är båda sanna eller båda falska.

Frasen "P om och endast om Q" är så frekvent i matematiska sammanhang, att den ofta förkortas till "P *omm* Q". På engelska skriver man "P *iff* Q". Sanningstabell för ekvivalensen finns i Tabell 4.

Så har vi då implikationen $P \Rightarrow Q$, som uttalas "P medför Q", "om P så Q" eller "P implicerar Q". Man säger också att "P är tillräckligt för Q" (det räcker att P gäller för att Q skall gälla) eller att "Q är nödvändigt för P" (för att P skall gälla är det nödvändigt att Q gäller). Om Q inte gäller, så gör inte P det heller.

Hur skall dess sanningsvärde definieras för olika P och Q? Om P är sann, skall Q också vara det, för att implikationen $P \Rightarrow Q$ skall gälla. Men om P är falsk? Vad skall då gälla för Q? Vi tittar på ett exempel.

Exempel 4.
I samband med ekvationslösning tillät vi oss att skriva

$$x \geq 2 \quad \Rightarrow \quad x^2 \geq 4.$$

Det tolkar vi som så, att om $x \geq 2$, så måste nödvändigtvis $x^2 \geq 4$. Men vad händer om $x \geq 2$ är falsk, det vill säga $x < 2$? Då finns både möjligheten att $x^2 < 4$ (t.ex. för $x = 1$) och att $x^2 \geq 4$ (t.ex. för $x = -3$). När vi alltså skriver $x \geq 2 \Rightarrow x^2 \geq 4$, så ignorerar vi fallet $x < 2$. Vad som sker med x^2 i detta fall anses ointressant. Implikationen hade varit falsk om det funnes ett värde på x, för vilket $x \geq 2$ vore sann och $x^2 \geq 4$ falsk.

Mer allmänt, så kan vi säga, i en implikation $P \Rightarrow Q$, att vi väljer att inte bry oss så mycket om situationen, då P är falsk. Om P är sann och dessutom Q är sann, så bör implikationen $P \Rightarrow Q$ vara sann, medan om P är sann och Q ändå är falsk, så bör implikationen vara falsk. Detta leder till följande definition.

Definition 13.5

Implikationen $P \Rightarrow Q$ av två utsagor utläses "P medför Q" eller "om P, så Q". Den definieras vara sann, precis då P och Q båda är sanna, *eller* P är falsk.

Sanningstabell för implikationen finns i Tabell 5. De i praktiken viktigaste raderna är alltså den första och den andra. De två sista raderna handlar om när P är falsk. Då är implikationen sann, oberoende av om Q är sann eller falsk. Det avspeglar, som sades ovan, att vi mest är intresserade av att använda implikationer, när vi redan vet att P är sann.

Exempel 5.
Från ekvationslösningen minns vi, hur kvadrering gav upphov till implikationer av typen

$$\sqrt{x+1} = 3 \quad \Rightarrow \quad x + 1 = 9,$$

som faktiskt är en ekvivalens, eftersom $x = 8$ kan verifieras vara en äkta rot i den ursprungliga ekvationen, och

$$\sqrt{x+1} = -3 \quad \Rightarrow \quad x + 1 = 9,$$

P	Q	$P \Rightarrow Q$
S	S	S
S	F	F
F	S	S
F	F	S

Tabell 5: Sanningstabell för implikation.

P	Q	$P \Rightarrow Q$	$Q \Rightarrow P$	$(P \Rightarrow Q) \wedge (Q \Rightarrow P)$
S	S	S	S	S
S	F	F	S	F
F	S	S	F	F
F	F	S	S	S

Tabell 6: Sanningstabell för $(P \Rightarrow Q) \wedge (Q \Rightarrow P)$.

som inte är en ekvivalens. Den högra utsagan är sann för $x = 8$, men inte den vänstra, så de har olika sanningsvärden.

Exempel 6.
Vi har de sanna implikationerna

$$x^2 = y^2 \quad \Leftrightarrow \quad (x = -y \ \vee \ x = y),$$

om x och y tilldelas värden i $\mathbb{R}$ (likgiltigt vilka), och

$$q \text{ är irrationellt} \quad \Rightarrow \quad q \text{ är inte ett heltal.}$$

Exempel 7.
Tabell 6 ger en sanningstabell för påståendet $(P \Rightarrow Q) \wedge (Q \Rightarrow P)$. Vi ser att denna har samma sanningsvärden som ekvivalensen $P \Leftrightarrow Q$. Dessa två utsagor är därför i sin tur ekvivalenta, det vill säga utsagan

$$\Big((P \Rightarrow Q) \wedge (Q \Rightarrow P)\Big) \Leftrightarrow (P \Leftrightarrow Q)$$

är alltid sann, oavsett sanningsvärdena för de ingående påståendena P och Q.

Beskrivningen $(P \Rightarrow Q) \wedge (Q \Rightarrow P)$ talar om en möjlig väg att bevisa en ekvivalens $P \Leftrightarrow Q$, som ofta används i matematiska bevis. Först visas att $P \Rightarrow Q$, och sedan visas separat att $Q \Rightarrow P$.

Exempel 8.
I Tabell 7 har sanningstabellen för ekvivalensen $P \Leftrightarrow Q$ utökats med en kolumn för dess negation $\neg(P \Leftrightarrow Q)$. Tydligen betyder denna utsaga "antingen P eller Q (men inte båda)". Detta är just det *exklusiva eller*, som nämndes tidigare i samband med disjunktioner. Av denna anledning betecknas denna *exklusiva disjunktion* ofta med symbolen $P \not\Leftrightarrow Q$.

P	Q	$P \Leftrightarrow Q$	$\neg(P \Leftrightarrow Q)$
S	S	S	F
S	F	F	S
F	S	F	S
F	F	S	F

Tabell 7: Sanningstabell för exklusiv disjunktion.

P	Q	$P \Rightarrow Q$	$\neg Q$	$\neg P$	$\neg Q \Rightarrow \neg P$
S	S	S	F	F	S
S	F	F	S	F	F
F	S	S	F	S	S
F	F	S	S	S	S

TABELL 8: Det indirekta bevisets logik.

P	Q	$P \Rightarrow Q$	$\neg Q$	$P \wedge (\neg Q)$	$P \wedge (\neg Q) \Rightarrow \mathrm{F}$
S	S	S	F	F	S
S	F	F	S	S	F
F	S	S	F	F	S
F	F	S	S	F	S

TABELL 9: Motsägelsebevisets logik.

Exempel 9.
Allra först i denna bok diskuterade vi bevis för satser av implikationstypen $P \Rightarrow Q$. I det *direkta beviset* antar man att P är sann, och härleder därefter Q genom något resonemang. Då är implikationen visad. Vi bryr oss ju inte om vad som händer om P är falsk; då är implikationen automatiskt sann.

Det *indirekta beviset* baseras på, att implikationen $P \Rightarrow Q$ är logiskt ekvivalent med **kontrapositionen** $\neg Q \Rightarrow \neg P$. Se Tabell 8. Notera, hur $P \Rightarrow Q$ och $\neg Q \Rightarrow \neg P$ har exakt samma sanningsvärden. Att utgå från P och visa Q, är alltså logiskt ekvivalent med att utgå från $\neg Q$ och visa $\neg P$.

Exempel 10.
Motsägelsebeviset baseras på att implikationen $P \Rightarrow Q$ är logiskt ekvivalent med utsagan

$$P \wedge (\neg Q) \Rightarrow \mathrm{F}.$$

Bokstaven F betecknar här en **kontradiktion** eller motsägelse, alltså ett falskt påstående eller omöjlighet (vilken som helst). Se Tabell 9. Notera, hur $P \Rightarrow Q$ och $P \wedge (\neg Q) \Rightarrow \mathrm{F}$ har exakt samma sanningsvärden. Att utgå från P och visa Q, är logiskt ekvivalent med att utgå från P och $\neg Q$, och härleda en motsägelse.

§5. KVANTIFIKATORER

Vi skall slutligen införa två operationer på utsagor av en ny typ, kallade **kvantifikatorer**. Betrakta en *öppen utsaga* $P(x)$, en utsaga, som (eventuellt) innehåller en variabel x, varvid x tar värden i en mängd A, till exempel någon av talmängderna $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ eller $\mathbb{C}$.

Definition 13.6

Den **universella kvantifikationen**

$$\forall x \in A : P(x)$$

utläses "för varje x i mängden A gäller $P(x)$".

Exempel 11.
Formeln $|x| = \sqrt{x^2}$ för absolutbeloppet av ett reellt tal är en öppen utsaga. Eftersom den gäller för alla *reella* x (men inte irreella!) kan den formuleras symboliskt som

$$\forall x \in \mathbb{R} : |x| = \sqrt{x^2}.$$

Konjugatregeln

$$(x+y)(x-y) = x^2 - y^2$$

är en öppen utsaga. Den gäller för alla komplexa tal x och y, och kan skrivas symboliskt

$$\forall x \in \mathbb{C} : \forall y \in \mathbb{C} : (x+y)(x-y) = x^2 - y^2,$$

eller kortare

$$\forall x, y \in \mathbb{C} : (x+y)(x-y) = x^2 - y^2.$$

Definition 13.7

Den **existentiella kvantifikationen**

$$\exists x \in A : P(x)$$

utläses "det existerar ett x i mängden A, sådant att $P(x)$ gäller".

En alternativ formulering är "för något x i A gäller $P(x)$".

Exempel 12.
Påståendet

$$\exists x \in \mathbb{R} : x \geq 0 \wedge x^2 = 2$$

uttrycker existensen av kvadratroten ur 2.

Exempel 13.
Uppenbarligen är både

$$\forall x \in \mathbb{R} : x^2 \geq 0 \qquad \text{och} \qquad \exists x \in \mathbb{R} : x^2 \geq 0$$

sanna påståenden. Påståendet

$$\exists x \in \mathbb{R} : x^2 \leq 0$$

är också sant, eftersom $0^2 = 0$.

Exempel 14.
Förefinnes flera kvantifikatorer är det inte egalt i vilken ordning de kommer. Påståendet

$$\forall x \in \mathbb{R} : \exists y \in \mathbb{R} : x = y$$

är sant, ty det kan utläsas "för varje reellt tal x finns ett reellt tal y, sådant att $x = y$". Det skenbart snarlika påståendet

$$\exists y \in \mathbb{R} : \forall x \in \mathbb{R} : x = y$$

utläses "det finns ett reellt tal y, som för alla reella x har egenskapen $x = y$", och är tydligen falskt.

Om det av sammanhanget är klart ur vilken mängd A värdena för variabeln x skall hämtas, kan vi drista oss att bara skriva $\forall x$ eller $\exists x$.

Exempel 15.
Om $\exists x : f(x) = 0$, säger vi att ekvationen $f(x) = 0$ är *lösbar*. Om $\forall x : f(x) \neq 0$, säger vi att ekvationen $f(x) = 0$ är *olösbar*.

Den öppna utsagan $x \leq y$ är varken sann eller falsk, innan vi tilldelat variablerna x och y (reella) värden. Likaså har påståendet $\exists x \in \mathbb{R} : x \leq y$ inget sanningsvärde förrän vi angivit vad y betecknar. Men

$$\forall y \in \mathbb{R} : \exists x \in \mathbb{R} : x \leq y$$

är definitivt ett sant påstående; det innehåller inga variabler som skall ges värden. Man säger att x och y i det sista påståendet är *bundna* variabler. Jämför med $\int_0^1 x^3\ dx$, som betecknar ett reellt tal, och där den bundna variabeln x inte skall ges något värde.

Exempel 16.
Följande logiska ekvivalenser gäller alltid, oberoende av vilka påståenden $P(x)$ och $Q(x)$ betecknar.

$$\begin{aligned}
\neg\forall x : P(x) &\quad\Leftrightarrow\quad \exists x : \neg P(x)\\
\neg\exists x : P(x) &\quad\Leftrightarrow\quad \forall x : \neg P(x)\\
\forall x : (P(x) \wedge Q(x)) &\quad\Leftrightarrow\quad \forall x : P(x) \wedge \forall x : Q(x)\\
\exists x : (P(x) \vee Q(x)) &\quad\Leftrightarrow\quad \exists x : P(x) \vee \exists x : Q(x)
\end{aligned}$$

Dessa ekvivalenser svarar egentligen bara mot sunt förnuft. Antag t.ex. att vi vill visa "det är falskt, att alla hela tal är delbara med 2". Det kan vi formulera som $\neg\forall x\ P(x)$, där $P(x)$ står för påståendet "x är delbart med 2". Det är nu bara att visa upp ett tal, t.ex. 1, som inte är delbart med 2. Detta sista innebär att vi har demonstrerat $\exists x : \neg P(x)$. Den första ekvivalensen ovan uttrycker bara att detta är en logiskt korrekt metod.

Övningar

13.1. För vilka reella tal x är följande påstående sant:

$$x^2 - 2x - 3 < 0 \quad \vee \quad 0 \leq x \leq 10\ ?$$

13.2. Skriv upp (de åttaradiga) sanningsvärdestabellerna för

$$P \vee (Q \wedge R) \qquad \text{och} \qquad (P \vee Q) \wedge R.$$

Resultatet visar på vikten av att, med hjälp av parenteser, utmärka hur påståendena skall fogas ihop.

13.3. (a) Visa, med hjälp av sanningstabell, att $P \wedge (Q \vee R)$ är logiskt ekvivalent med $(P \wedge Q) \vee (P \wedge R)$.

(b) Samma uppgift för $P \vee (Q \wedge R)$ och $(P \vee Q) \wedge (P \vee R)$.

Dessa är de *distributiva lagarna*.

13.4. (a) Visa, att $\neg P \wedge \neg Q$ är logiskt ekvivalent med $\neg(P \vee Q)$.

(b) Samma uppgift för $\neg P \vee \neg Q$ och $\neg(P \wedge Q)$.

Dessa är *de Morgans lagar*.

13.5. (a) Visa, att utsagan $P \vee (\neg P)$ alltid är sann, en *tautologi*.

(b) Visa, att utsagan $P \wedge (\neg P)$ alltid är falsk, en *kontradiktion*.

13.6. Illustrera de övriga ekvivalenserna i Exempel 13.16 med konkreta exempel. Ge argument för att de gäller.

13.7. Ange för alla reella värden på a och b när följande påståenden är sanna.

(a) $\exists x \in \mathbb{R} : ax = b$.

(b) $\forall x \in \mathbb{R} : ax \neq b$.

13.8. Negationen till påståendet "alla läroböcker i matematik är oläsbara" kan givetvis formuleras "det är inte sant, att alla läroböcker i matematik är oläsbara", men det är naturligare att säga "någon lärobok i matematik är läsbar".

Formulera på detta sätt negationen till följande påståenden.

(a) “Alla läroböcker i matematik och logik är oläsbara.”

(b) “Alla klassens elever fick godkänt i alla språkämnen.”

(c) “Minst en av klassens elever fick godkänt i alla språkämnen.”

(d) “Minst sju av sjömännen var sjuka.”

(e) “Högst sju av sjömännen var sjuka.”

Kapitel 14

Mängdlära

§1. Mängder och delmängder

En **mängd** är en samling objekt. Välbekanta är ju nedanstående talmängder, som är så ofta förekommande att de givits standardbeteckningar:

$$\begin{aligned}\mathbb{Z}^+ &= \text{mängden av positiva heltal}\\ \mathbb{N} &= \text{mängden av naturliga tal (inklusive 0)}\\ \mathbb{Z} &= \text{mängden av heltal}\\ \mathbb{Q} &= \text{mängden av rationella tal}\\ \mathbb{R} &= \text{mängden av reella tal}\\ \mathbb{C} &= \text{mängden av komplexa tal}\end{aligned}$$

Men en mängd behöver inte bestå av tal. Alla slags matematiska objekt kan sammanföras till mängder, exempelvis mängden av alla polynom, mängden av alla deriverbara funktioner eller mängden av alla trianglar i planet. Är vi på generöst humör, kan vi även inkludera konkreta, icke-matematiska objekt och få åskådliga exempel som mängden av alla nu levande människor.

Om x tillhör mängden A, säger man att x är ett **element** i A, och skriver detta $x \in A$ (uttalat "x tillhör A"). Om x inte tillhör A, skriver man $x \notin A$. Att x är ett reellt tal kan alltså kortfattat uttryckas $x \in \mathbb{R}$.

Den mängd som består av elementen $x_1, \ldots, x_n$ tecknas $\{x_1, \ldots, x_n\}$ (utläses "mängden av elementen $x_1, \ldots, x_n$"). Måsvingarna kallas i matematiken *mängdklammer*.

Två saker förtjänar här att anmärkas. För det första, så definieras ju en mängd bara som en samling objekt. Därför spelar elementens ordning ingen som helst roll. Mängden av de tre första positiva heltalen kan skrivas på flera sätt:

$$\{1,2,3\} = \{2,3,1\} = \{3,2,1\} = \cdots$$

För det andra, så finns det bara två möjligheter för ett objekt: antingen tillhör det en viss mängd eller också inte. Mängden ovan kan därför också skrivas

$$\{1,2,3\} = \{1,1,2,3\} = \{1,1,1,2,2,3\} = \cdots$$

Även om samma element skrivs flera gånger, så räknas det ändå bara en gång.

I regel kan vi inte räkna upp elementen i en mängd. Det vanligaste sättet att specificera en mängd är därför, att ange ett definierande villkor. Givet en mängd A, definierar vi

$$\{\, x \in A \mid P(x) \,\}$$

som mängden av alla $x \in A$, som uppfyller villkoret $P(x)$. Det vertikala strecket $\mid$ kallas i detta sammanhang för *mängdbyggaren*. Även andra tecken, som ett kolon eller semikolon, kan nyttjas för detta ändamål.

Exempel 1.
Låt mängden A bestå av alla reella tal x, sådana att $x^2 - 2x + 1 = 0$. Det är lätt att se, att A bara består av det enda elementet 1. Det skriver vi

$$A = \{\, x \in \mathbb{R} \mid x^2 - 2x + 1 = 0 \,\} = \{1\}.$$

Exempel 2.
Inom den endimensionella analysen arbetar man vanligen med intervall av reella tallinjen. De är så vanligt förekommande, att man infört speciella beteckningar för dem. Det *slutna* intervallet definieras av

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

och det *öppna* intervallet av

$$(a, b) = \{\, x \in \mathbb{R} \mid a < x < b \,\}.$$

Exempel 3.
Låt mängden B bestå av alla funktioner $f\colon \mathbb{R} \to \mathbb{R}$, sådana att $f'(x) = x^2$ och $f(0) = 1$. Från analysen vet vi, att B har det enda elementet $\frac{1}{3}x^3 + 1$, och alltså är

$$B = \{\, f\colon \mathbb{R} \to \mathbb{R} \mid f'(x) = x^2 \wedge f(0) = 1 \,\} = \left\{ \frac{1}{3}x^3 + 1 \right\}.$$

Låt mängden C innehålla alla funktioner $f\colon \mathbb{R} \to \mathbb{R}$, sådana att $f'(x) = x^2$. Den består av alla funktioner f sådana att $f(x) = \frac{1}{3}x^3 + C$, där C är ett reellt tal, och detta kan skrivas

$$C = \{\, f\colon \mathbb{R} \to \mathbb{R} \mid f'(x) = x^2 \,\} = \left\{ f(x) = \frac{1}{3}x^3 + C \;\middle|\; C \in \mathbb{R} \right\}.$$

Exempel 4.
Betrakta mängden F av alla positiva heltal n med egenskapen, att det finns positiva heltal x, y, z, sådana att $x^n + y^n = z^n$, symboliskt

$$F = \{\, n \in \mathbb{Z}^+ \mid \exists x, y, z \in \mathbb{Z}^+ : x^n + y^n = z^n \,\}.$$

Det tog 350 år och ca 10 000 sidor ytterst avancerad matematisk text för att visa att $F = \{1, 2\}$; detta är Fermats stora sats.

Exempel 5.
Låt slutligen G vara mängden av alla jämna positiva heltal större än 2, som kan skrivas som summan av två primtal. Det är fortfarande okänt, huruvida denna består av *alla* jämna tal större än 2; detta är Goldbachs förmodan från 1700-talet.

Mängderna i exemplen ovan var av vitt skilda slag. Men alla är förvisso mängder, och tankemässigt vinner man något i påtaglighet genom att sammanfatta deras element i ett enda begrepp och ge dem ett namn.

Definition 14.1

Om varje element i mängden A också är ett element i mängden B, säges A vara en **delmängd** av B, och man skriver $A \subseteq B$ eller $B \supseteq A$. Om A är en delmängd av B, och dessutom $A \neq B$, så säges A vara en **äkta delmängd** av B, och man skriver $A \subset B$ eller $B \supset A$.

Exempel 6.
Talmängderna ovan är alla delmängder av varandra och låter sig ordnas i en kedja:

$$\mathbb{Z}^+ \subseteq \mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}.$$

Alla är äkta delmängder av varandra, så för att betona detta, skulle vi också kunna använda symbolen $\subset$.

§2. Mängdoperationerna

Av givna mängder kan vi konstruera nya.

Definition 14.2

Unionen eller **föreningsmängden** $A \cup B$ av mängderna A och B består av de objekt, som tillhör minst en av mängderna A och B:

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

Exempel 7.
Vi har $\mathbb{N} = \mathbb{Z}^+ \cup \{0\}$.

Definition 14.3

Snittet eller **skärningsmängden** $A \cap B$ av mängderna A och B består av de objekt, som tillhör båda mängderna A och B:

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

Exempel 8.
Snittet av de två intervallen $[0, 2]$ och $[1, 3]$ är $[0, 2] \cap [1, 3] = [1, 2]$.

Vad är $\{1, 2, 3\} \cap \{4, 5, 6\}$? Uppenbarligen finns inget element som tillhör bägge mängderna. Vi står nu inför två alternativ. En väg är att förklara, att skärningen mellan två mängder helt enkelt inte är definierad, då det saknas gemensamma element. Den utvägen känns frånstötande eftersom vi gärna vill kunna skriva upp $A \cap B$, utan att varje gång kontrollera om det är tillåtet. En annan väg är att införa en tom mängd $\emptyset$ och sedan låta $\{1, 2, 3\} \cap \{4, 5, 6\} = \emptyset$. Detta är det naturliga tillvägagångssättet och det är just det man har valt.

Definition 14.4

Den **tomma mängden** är den unika mängd, som saknar element:

$$\emptyset = \{\}.$$

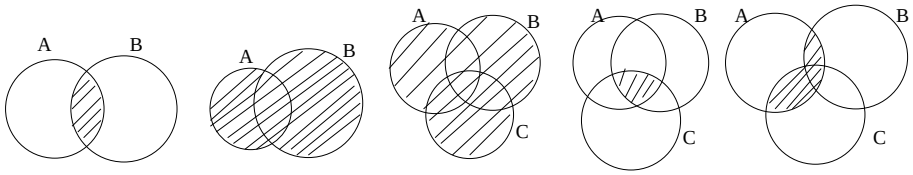
Denna är alltså ingen onödig abstraktion, eftersom den erfordras om man vill kunna använda skärning ohämmat. Den tomma mängden är unik, emedan två mängder med samma element (i detta fall inga element) är lika.

Exempel 9.
Den tomma mängden är en delmängd av varje mängd A, så att $\emptyset \subseteq A$ alltid gäller. Varje element i $\emptyset$ (och det finns ju inga) är också med i A.

Vidare har vi tydligen alltid

$$A \cap \emptyset = \emptyset \qquad \text{och} \qquad A \cup \emptyset = A.$$

(Jämför med $a \cdot 0 = 0$ och $a + 0 = a$.)



FIGUR 1: Venn-diagram.

Om $A \cap B = \emptyset$, säges mängderna A och B vara **disjunkta**. Mängderna $\{1, 2, 3\}$ och $\{4, 5, 6\}$ är alltså disjunkta.

För att åskådliggöra operationer på mängder användes ofta så kallade *Venn-diagram*. Se Figur 1. Mängden A, vad den än månde bestå av (tal, funktioner, polynom, människor, ...), tänkes schematiskt som cirkelskivan märkt med A (randen jämte innanmätet). I de två första Venn-diagrammen finns situationen för två mängder A och B. Det första diagrammet visar snittet $A \cap B$, det andra unionen $A \cup B$.

Sedan visas en schematisk överblick över tre mängder. Först kommer unionen $A \cup B \cup C$, därefter snittet $A \cap B \cap C$. Det femte diagrammet, slutligen, demonstrerar identiteten

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$

distributiva lagen för snitt över union. (Jämför med den distributiva lagen $a(b + c) = ab + ac$.) Det finns sjokvis av sådana räknelagar för mängdoperationer, och alla låter sig bevisas genom Venn-diagram. Normalt kan man inte dra några slutsatser ur en figur, men Venn-diagrammet är inte någon *geometrisk* figur, som vi gör avläsningar i, utan en *logisk* och *schematisk* överblick över situationen, giltig för alla möjliga mängder A, B (och C). Eventuellt kan förstås något av områdena i Venn-diagrammet vara tomt.

Exempel 10.
För den läsare, som inte känner sig riktigt övertygad av det bildliga beviset ovan, visar vi också den distributiva lagen genom ett logiskt resonemang. För att visa att två mängder M och N är lika, kan det vara klokt att visa $M \subseteq N$ och $N \subseteq M$ i två separata argument. Vi vill alltså visa, att vi har både

$$A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C) \quad \text{och} \quad A \cap (B \cup C) \supseteq (A \cap B) \cup (A \cap C).$$

Om $x \in A \cap (B \cup C)$, så är $x \in A$, och dessutom tillhör x en av mängderna B och C. Om $x \in B$ har vi $x \in A \cap B$; om $x \in C$ har vi $x \in A \cap C$. I varje fall har vi $x \in (A \cap B) \cup (A \cap C)$. Därmed är den första inklusionen riktig.

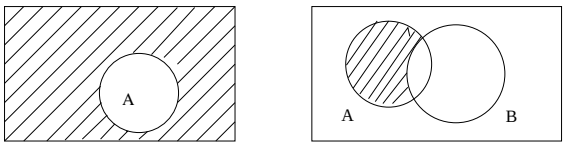
För att visa den andra, antar vi att $x \in (A \cap B) \cup (A \cap C)$ och får då att (minst) ett av alternativen $x \in A \cap B$ eller $x \in A \cap C$ gäller. Om $x \in A \cap B$, måste tydligen x tillhöra både A och B, därför tillhör x både A och $B \cup C$, alltså $x \in A \cap (B \cup C)$. Om $x \in A \cap C$, ses på samma sätt att $x \in A \cap (B \cup C)$. Likheten är därmed visad.

Komplementet till en mängd A består av alla objekt, som *inte* tillhör A. Men det är ofta ganska mycket som inte tillhör A (troligen läsaren själv till exempel), och detta skulle bringa oss i filosofiska bryderier. Dessutom är vi vanligen intresserade endast av objekt av ett visst slag. Därför menar man i definitionen inte bokstavligen *alla* objekt utanför A, utan inskränker sig till någon på förhand angiven **universalmängd** eller **universum** U. Universalmängden varierar mellan olika områden av matematiken. Inom den reella analysen tänker man sig antagligen $U = \mathbb{R}$. När vi arbetade med polynom var $U = \mathbb{C}$; inom talteorien är $U = \mathbb{Z}$.

Definition 14.5

Låt A vara en mängd, inkluderad i någon universalmängd U. Mängden av alla objekt i U, som *inte* tillhör A, kallas **komplementet** till A, och tecknas A^c:

$$A^c = \{\, x \in U \mid x \notin A \,\} = \{\, x \in U \mid \neg x \in A \,\}.$$



FIGUR 2: Venn-diagram för A^c och $A \setminus B$.

Vi ser kvickt att $\emptyset^c = U$ och $U^c = \emptyset$.

Definition 14.6

Det **relativa komplementet** $A \setminus B$ av A med avseende på B är mängden alla element i A, som inte ligger i B:

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

Vi ser enkelt att $A \setminus B = A \cap B^c$ (givet en universalmängd, så att B^c är definierad).

Exempel 11.
De negativa heltalen skriver vi $\mathbb{Z} \setminus \mathbb{N}$. De irrationella talen tecknas $\mathbb{R} \setminus \mathbb{Q}$ och de irreella talen $\mathbb{C} \setminus \mathbb{R}$.

I ett Venn-diagram ritas universalmängden oftast som en stor rektangel, som innesluter de övriga mängderna. I Figur 2 finns komplementet A^c och det relativa komplementet $A \setminus B$ schematiskt avbildade.

§3. Ändliga mängder

Med en **ändlig** mängd menas förstås en mängd som har ändligt många element. (Motsatsen är en **oändlig** mängd.) Om A är en ändlig mängd betecknar vi med $|A|$ antalet element i A.

Exempel 12.
Tydligen är $|\emptyset| = 0$ och $|\{x\}| = 1$. Ytterligare är $|\{x, y\}| = 2$ (förutsatt förstås att $x \neq y$).

Exempel 13.
Herr Varians, Statistiska centralbyråns i särklass klantigaste marknadsundersökare, producerade en rapport med följande lydelse. Antalet undersökta personer var 100. Av dessa drack 77 personer kaffe, 69 personer te, 54 personer bådadera och 9 personer ingetdera. Rapporten underkändes dock och hamnade i papperskorgen. Herr Varians fick sparken. Varför?

Situationen kan beskrivas med hjälp av mängder. Låt K vara mängden av kaffedrickare och T mängden av tedrickare. Vi har givet att $|K| = 77$ och $|K \cap T| = 54$. Antalet personer, som bara drack kaffe, är därför $|K \setminus T| = 77 - 54 = 23$. Tedrickarna är $|T| = 69$, och antalet personer, som endast drack te, är $|T \setminus K| = 69 - 54 = 15$. Antalet personer, som drack någon av dryckerna kaffe eller te, är

$$|K \cup T| = |K \setminus T| + |T \setminus K| + |K \cap T| = 23 + 15 + 54 = 92.$$

De, som drack varken kaffe eller te var enligt rapporten 9. Sammanlagt ger detta $92 + 9 = 101$ personer, stridande mot förutsättningarna.

Exemplet förmedlar det felaktiga intrycket, att detta och liknande problem måste analyseras med list. Följande sats avlägsnar behovet av slughet och ger en systematisk formel för antalet element i snitt och unioner, som är nyttig vid konkreta problem. Vi nöjer oss, i denna första kurs, med att ge den för två och tre mängder, men den kan generaliseras till fler.

Sats 14.7: Principen om inklusion–exklusion

För två mängder A och B gäller formeln

$$|A \cup B| = |A| + |B| - |A \cap B|,$$

och för tre mängder A, B och C formeln

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |B \cap C| - |A \cap C| + |A \cap B \cap C|.$$

Bevis

Vi betraktar fallet med två mängder A och B. Talet $|A \cup B|$ räknar antalet element, som tillhör någon av A eller B. För att få dessa kan vi naivt addera elementen i A och elementen i B. De element, som tillhör både A och B, kommer då att ha räknats två gånger. För att kompensera för detta, subtraherar vi $|A \cap B|$, och formeln är bevisad. Fallet med tre mängder kan bevisas med ett motsvarande resonemang.

Exempel 14.
I exemplet ovan med kaffe- och tedrickare följer att antalet personer, som drack kaffe eller te, är

$$|K \cup T| = |K| + |T| - |K \cap T| = 77 + 69 - 54 = 92.$$

Tillsammans med de 9 personer, som drack ingetdera, blir det $92 = 9 = 101$ personer, i strid med förutsättningarna.

Exempel 15.
Principen om inklusion–exklusion ger en exakt formel, men ibland har man ofullständig information eller nöjer sig med en uppskattning. Betrakta n stycken ändliga mängder $A_1, \ldots, A_n$. Tydligen har vi

$$|A_1 \cup \cdots \cup A_n| \leq |A_1| + \cdots + |A_n|, \tag{1}$$

där likhet inträffar om och endast om $A_1, \ldots, A_n$ är parvis disjunkta.

Betrakta nu n mängder $B_1, \ldots, B_n$ inuti ett ändligt universum U. Av (1), parad med det faktum att $|B_k^c| = |U| - |B_k|$ för varje k, följer

$$\begin{aligned} |B_1 \cap \cdots \cap B_n| &= |U| - |(B_1 \cap \cdots \cap B_n)^c| = |U| - |B_1^c \cup \cdots \cup B_n^c| \\ &\geq |U| - |B_1^c| - \cdots - |B_n^c| = |B_1| + \cdots + |B_n| - (n-1)|U|. \end{aligned} \tag{2}$$

När gäller likhet?

§4. Matematikens grundvalar

Mängder — ja, vad sysslar vi egentligen med för något? Existensen av matematiska objekt är långtifrån oproblematisk, om man börjar rota litet. Var finns alla dessa oändliga decimaltal, räta linjer eller ens s.k. “naturliga” tal?

Vad är egentligen talet 2? Ingenstans finns det en tvåa gjuten i en dyr legering av platina–iridium, som skulle kunna vara det riktiga officiella *Två*, i stil med den meterförebild, arkivmetern, som förvaras i Internationella byrån för mått och vikt i Paris. I stället får man kanske nöja sig med att säga, att talet 2 är en abstrakt konvention, som används på ett visst precist sätt, och det känns kanske inte alltför tryggt. Å andra sidan fungerar begreppet synnerligen väl i den matematiska vardagen.

För att tänka och kommunicera effektivt med våra matematiska abstraktioner behöver vi tänka på dem som begrepp med konkret och påtaglig existens, som vi kan peka på, hantera och

visa upp, inte bara som långa kedjor av språkliga definitioner och konstruktioner. Det är detta problem, som införandet av mängder löser, eller rättare förpassar någon annanstans.

En definition av matematiken är just som *läran om strukturer på mängder*. En av dessa mängder är $\mathbb{R}$, som kommer utrustad med vissa egenskaper. Mängdlärans verktyg och konstruktioner genomsyrar hela matematiken. Räcker de gamla inte till, kan vi med hjälp av snitt och unioner skapa nya mängder. Vi slipper därmed, att gång på gång göra oprecisa definitioner i stil med att "en punkt är läge utan utsträckning", och kan istället säga att en punkt är det ordnade paret (x, y), där $x, y \in \mathbb{R}$. Med ett elegant, men litet ohederligt trick har vi förpassat spörsmål om den ofrånkomliga vagheten och osäkerheten om matematikens grundvalar till specialisterna på logik, där de rätteligen hör hemma. Själva kan vi övergå till något roligare, t.ex. att använda matematik på väsentliga problem.

Till mängdläran skyfflar vi alla våra grubbel och svårigheter för att sedan låta filosofer och logiker, i sin existentiella 2000-årskris, bry sina huvuden över dem. Mängdläran har blivit de allmänt accepterade grundvalarna för vår vetenskap. Matematiker säger numera högtidligt, att matematiska objekt *är* mängder.

Vi skojar inte. Den tomma mängden spelar den hyllade huvudrollen. Talet noll definieras då helt enkelt som den tomma mängden: $0 = \emptyset$. Talet ett definieras som $1 = \{\emptyset\}$, talet två som $2 = \{\emptyset, \{\emptyset\}\}$, och så vidare. Ingen matematiker skulle förstås få för sig att *tänka* på talet 2 på detta makabra och artificiella vis, men det bygger en logisk grund för matematiken med mängdläran som grund. Naturliga tal definieras som mängder, närmare bestämt kombinationer av den tomma mängden, därefter skapas de hela talen från de naturliga, och så har maskineriet stånkat igång.

Då vet vi vad vi sysslar med: att existera är att vara ett element av en mängd, och när man dör flyttar man över till dess komplement. Trots kritik visar sig detta med mängder vara ett mycket praktiskt sätt att tänka, inte minst för att vi kan konstruera en mängd så behändigt, bara genom att ange dess särskiljande eller definierande egenskap, och på det sättet laborera med den, som om vi verkligen hölle den i vår hand och visste (nästan) allt om den.

§5. Uppsagda ur paradiset

Några avslutande ord om vad som hände med dem, som tog mängdläran på för stort allvar. Av vad vi just sett, verkar det som om vi kan konstruera mängder rätt vilt och hämningslöst. För varje öppen utsaga $P(x)$ skapade vi ju mängden $\{\, x \mid P(x) \,\}$ av alla objekt x, som uppfyller villkoret $P(x)$. Men detta naiva förhållningssätt ledde till katastrof. År 1901 upptäckte filosofen Bertrand Russell en paradox, som krävde evakuering av mängdlärans härliga paradis.

Den infekterade frågan gällde mängder, som var medlemmar av sig själva. Mängden T av alla trianglar är förstås inte själv en triangel, och alltså inte ett element i sig själv: $T \notin T$. Detta är den vanligaste situationen för en mängd, men, med vårt naiva förhållningssätt, skulle vi kunna tänka oss mängder, som faktiskt vore medlemmar av sig själva. Ett exempel skulle kunna vara "mängden" S, vars element är de mängder, som kan beskrivas med ändligt många (svenska) ord. (Vi skriver "mängd", för den leder till trubbel.) Mängden $N = \{1, 2, 3\}$ kan beskrivas med ändligt många ord som "mängden av heltal större än 0 och mindre än 4". Alltså är den en av de mängder som skulle ingå som element i S, d.v.s. $N \in S$. Men uttrycket "mängden av de mängder, som kan beskrivas med ändligt många ord" är ju en beskrivning av S i ett ändligt antal ord, och S är alltså kvalificerad för medlemskap i sig själv: $S \in S$.

Nu till Russells paradox. Sorgebarnet är

$$M = \{\, A \mid A \notin A \,\},$$

mängden av alla mängder, som inte är medlemmar av sig själva. Hur är det då med själva mängden M? Är det så att $M \in M$? Om svaret är nej, alltså $M \notin M$, så är ju M kvalificerad för medlemskap i sig själv, och vi borde alltså ha $M \in M$. Men i samma stund som vi låter $M \in M$, så diskvalificerar sig M för medlemskap i sig själv, och vi borde alltså ha $M \notin M$. Denna ohjälpliga motsägelse klarar vi oss inte ur. Inget tycks fungera här, utan vi har hittat ett påstående som varken är sant eller falskt. Detta ledde förstås till massjälvmord bland matematiker, eller?

Nej, ty denna paradox berör inte det matematiska verkstadsgolvet i någon större utsträckning. Man har inte något större intresse och användning för den typ av "mängd", som konstrueras i Russells paradox. Den kom visserligen den första euforien över matematikens logiska och mängdteoretiska grundvalar att falna, men det gick utmärkt att undvika Russells paradox, utan att behöva ge upp mängdlärans praktiska flexibilitet och fröjder, genom att på olika fiffiga sätt bygga in förbud mot konstruktioner i stil med den som nyss gav en paradox. I en matematisk variant av "Vi flytt' int'", sade Hilbert: "Ingen skall driva ut oss ifrån det paradis, som Cantor skapade åt oss." Cantor var alltså den ursprunglige arkitekten bakom mängdläran.

Problemen kan bl.a. lösas genom att man kräver, att nya mängder hela tiden måste skapas inuti någon given universalmängd. Man får alltså inte ha för stora ambitioner, utan snällt hålla sig inom ett någorlunda litet universum. Detta mängdteoretiska universum är dock av så ofattbara och kataklysmiska proportioner, att alla rimliga mängder och all matematik man är intresserad av får plats. Detta exkluderar mängden i Russells paradox, och vi kan fortsätta jaga älg i paradiset.

Avslutningsvis en annan variant på Russells paradox i ett väldigt litet universum — som tur är går den inte att formulera strikt matematiskt. Frissan i byn hemma säger: "Jag klipper precis de människor, som inte klipper sig själva." Är det så att hon klipper sitt eget hår?

Övningar

14.1. Skriv upp alla delmängder till $\{1, 2, 3\}$.

14.2. Betrakta mängderna

$$A = \{1, 3, 5, 7, 9\} \qquad \text{och} \qquad B = \{6, 7, 8, 9, 10\}.$$

Ange följande:

$$A \cup B, \qquad A \cap B, \qquad B \setminus A, \qquad |A|.$$

14.3. Betrakta de tre mängderna

$$A = \{1, 2, 3\}, \qquad B = \{3, 4, 5\} \qquad \text{och} \qquad C = \{1, 3, 5\}.$$

(a) Rita ett Venn-diagram för mängderna, där talen $1, \ldots, 5, 6$ är korrekt utplacerade.

(b) Ange elementen i mängderna

$$A \cup (B \cap C) \qquad \text{och} \qquad (B \cup C) \setminus A.$$

14.4. Låt

$$A = \{\, x \in \mathbb{R} \mid |x - a| \leq 1 \,\} \qquad \text{och} \qquad B = \{\, x \in \mathbb{R} \mid |x - b| \geq 2 \,\}.$$

Ange villkor på a och b för att $A \cap B = \emptyset$ resp. $A \cap B = \{0\}$.

14.5. Skriv upp alla mängder som kan bildas med union och snitt från mängderna (intervallen) $[0, 2]$, $[1, 3]$ och $[2, 4]$. (Det blir nio stycken om de tre givna räknas med.)

14.6. Den *distributiva lagen för union över snitt* säger

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

Visa den med dels Venn-diagram, dels logiskt resonemang.

14.7. Skriv upp alla mängder som kan bildas med union, skärning och komplement relativt universalmängden $\mathbb{R}$ från $[0, 2]$ och $[1, 2]$. (Det blir åtta stycken.)

14.8. *De Morgans lagar* säger

$$(A \cup B)^c = A^c \cap B^c \qquad \text{och} \qquad (A \cap B)^c = A^c \cup B^c.$$

Visa dem med dels Venn-diagram, dels logiskt resonemang.

14.9. De två mängderna A och B uppfyller

$$|A \cap B| = 999, \qquad |A| = 1000 \qquad \text{och} \qquad |A \cup B| = 1001.$$

Beräkna $|B|$.

14.10. I ett häftigt slag förlorade minst 70% av de stridande ett öga, minst 75% ett öra, minst 80% en arm och minst 85% ett ben. Använd Exempel 14.15 för att finna det minsta antal, som drabbades av alla fyra skadorna.

14.11. En marknadsundersökning meddelar, att av 1000 tillfrågade tycker 816 om konfekt, 725 om glass och 645 om tårta medan 562 tycker om både konfekt och glass, 463 om både konfekt och tårta, 470 om både glass och tårta och slutligen 310 om alla tre sorterna. Är resultatet troligt?

Kapitel 15

Facit

1.1a $a^3 - 3a^2b + 3ab^2 - b^3$.

1.1b $a^2 + b^2 + c^2 + 2ab + 2ac + 2bc$.

1.2a $x(x-3)(x+3)$.

1.2b $(a-9)(a+b)$.

1.2c $(x+a+b)(x+a-b)$.

1.2d

$$x^6 - x^4 + x^2 - 1 = x^4(x^2-1) + (x^2-1) = (x^4+1)(x^2-1) = (x^4+1)(x-1)(x+1).$$

Med ett trick kan faktorn x^4+1 faktoriseras ytterligare som

$$x^4 + 1 = x^4 + 2x^2 + 1 - 2x^2 = (x^2+1)^2 - (\sqrt{2}x)^2 = (x^2 + 1 + \sqrt{2}x)(x^2 + 1 - \sqrt{2}x).$$

1.3 $\frac{7}{6(x+3)}$.

1.4a $(x^2+y^2)(x-y)$.

1.4b $\left(\frac{4x^2}{9} + y^2\right)\left(\frac{2x}{3} - y\right) = \frac{1}{27}(4x^2+9y^2)(2x-3y)$.

1.4c x.

1.5a $\frac{1}{a+b}$.

1.5b $\frac{1}{a}$.

1.5c $\frac{a+b}{a-b}$.

1.6 $-(9+4\sqrt{5})$.

1.7 $a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$.

1.8 Enligt Binomialsatsen är $(2x^2 + \frac{1}{4x})^{99}$ en summa av termer på formen

$$\binom{99}{k} \left(2x^2\right)^k \left(\frac{1}{4x}\right)^{99-k}.$$

I den konstanta termen skall exponenten för x vara 0, vilket inträffar då $2k = 99 - k$, alltså $k = 33$. Konstanta termen är då

$$\binom{99}{33} 2^{33} \left(\frac{1}{4}\right)^{66} = 2^{-99}\binom{99}{33}.$$

1.9 Enligt Binomialsatsen är $(\frac{x}{3} - \frac{1}{x^3})^{10}$ en summa av termer på formen

$$\binom{10}{k}\left(\frac{x}{3}\right)^k\left(-\frac{1}{x^3}\right)^{10-k}.$$

Denna har grad 2 om $k - 3(10-k) = 2$, alltså $k = 8$. Koefficienten är $3^{-8}\binom{10}{8}$.

1.10 Använd Aritmetisk–geometriska olikheten.

1.11a $A(1,2) = \frac{3}{2}$, $G(1,2) = \sqrt{2}$, $H(a,b) = \frac{4}{3}$.

1.11b Alla tal är positiva, så $a \le \frac{2}{\frac{1}{a}+\frac{1}{b}}$ är ekvivalent med $1 + \frac{a}{b} \le 2$, vilken är ekvivalent med $\frac{a}{b} \le 1$, vilket är sant då $a \le b$. Den andra olikheten visas på liknande sätt.

1.11c Observera att $\frac{1}{H(a,b)} = A(\frac{1}{a}, \frac{1}{b})$. Använd Aritmetisk–geometriska olikheten.

1.12b $a^3 + b^3 = (a+b)(a^2 - ab + b^2)$.

2.1a $\sum_{k=1}^{n} k$ och $\prod_{k=1}^{n} k$.

2.1b $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$.

2.2a $\sum_{k=1}^{n} 2k = 2 + 4 + 6 + \cdots + 2n = \frac{1}{2}n(2+2n) = n(n+1)$.

2.2b $\sum_{k=1}^{n}(2k-1) = 1 + 3 + 5 + \cdots + (2n-1) = n^2$.

2.3 $\frac{1}{2}n(3n-1)$.

2.4 Först får du 512 böcker. Sedan lämnar du in dessa och får 256 nya, o.s.v. Alltså är svaret $512 + 256 + \cdots + 1 = 1023$.

2.5a 1023.

2.5b $\frac{121}{27}$.

2.5c 683.

2.5d $\frac{x^{10}-1}{x-1}$.

2.5e $\frac{1-x^{10}}{1+x}$.

2.6a $\frac{1}{1} + \frac{1}{2} + \frac{1}{3} = \frac{11}{6}$;

2.6b Summan är

$$3\frac{1-2^{11}}{1-2} = 3(2^{11}-1) = 6141.$$

2.6c Följden i summan startar med $3\cdot 2^m$, slutar med $3\cdot 2^n$, så att det supersista talet är $3\cdot 2^{n+1}$. Vidare är kvoten 2. Alltså är summan

$$\sum_{i=m}^{n} 3\cdot 2^i = \frac{3\cdot 2^m - 3\cdot 2^{n+1}}{1-2} = 3(2^{n+1} - 2^m).$$

2.6d $\displaystyle\sum_{i=0}^{2n} 3\cdot x^{-i} = \frac{3(x^{-(2n+1)}-1)}{x^{-1}-1} = \frac{-3(x-x^{-2n})}{1-x}.$

2.6e $\displaystyle\sum_{i=0}^{2n} 3\cdot x^{-n} = 3(2n+1)x^{-n}.$

2.6f 4^n.

2.6g $\frac{1}{n+1}$.

2.7 Nej. Av $ak - a = ak^2 - ak$ följer antingen $a = 0$ eller $k = 1$.

3.1a $\frac{1}{25}$.

3.1b 3.

3.1c $3\sqrt{3}$.

3.1d 9.

3.1e $2^8 = 256$.

3.1f $2^6 = 64$.

3.2a 1.

3.2b $a^{\frac{7}{12}}$.

3.2c $y^{\frac{1}{4}}$.

3.2d $x^{\frac{4}{3}}$.

3.3a $26 \cdot 5^x$.

3.3b $(e+1)e^x$.

3.3c 2.

3.3d $\frac{6}{5^{x+1}}$.

3.4a $\frac{x-x^{n+1}}{1-x}$.

3.4b $x^{\frac{1}{2}n(n+1)}$.

4.1 $p(x) = 2\left(\left(x - \left(\frac{1}{4}\right)\right)^2 - \frac{41}{16}\right)$, nollställena är $x = \frac{1 \pm \sqrt{41}}{4}$, och det minsta värdet är $-\frac{41}{8}$ (för $x = \frac{1}{4}$).

4.2a $x = 3$.

4.2b $x = -1$.

4.2c $x = 2$.

4.3 Dubbel kvadrering ger

$$\begin{aligned}
& \sqrt{x} + \sqrt{x+1} = \frac{3}{2} \\
\Rightarrow \quad & \frac{9}{4} = \left(\sqrt{x} + \sqrt{x+1}\right)^2 = 2x + 1 + 2\sqrt{x^2+x} \\
\Leftrightarrow \quad & \sqrt{x^2+x} = \frac{5}{8} - x \\
\Rightarrow \quad & x^2 + x = \left(\frac{5}{8} - x\right)^2 = \frac{25}{64} - \frac{5}{4}x + x^2 \\
\Leftrightarrow \quad & \frac{9}{4}x = \frac{25}{64}
\end{aligned}$$

$$\Leftrightarrow \quad x = \frac{25}{144}.$$

Vid kontroll av roten befinnes denna vara äkta:

$$\sqrt{\frac{25}{144}} + \sqrt{\frac{25}{144} + 1} = \sqrt{\frac{25}{144}} + \sqrt{\frac{169}{144}} = \frac{5}{12} + \frac{13}{12} = \frac{3}{2}.$$

Ekvationen har alltså lösningen $x = \frac{25}{144}$.

4.4a $x = \pm\sqrt{2}$.

4.4b $x = 0$.

4.4c $x = \frac{\pi}{2} + n \cdot \pi$.

4.5a $x < -\frac{1}{2}$.

4.5b Alla x.

4.5c $x \geq -\frac{2}{3}$.

4.6a $-2 < x < 2$.

4.6b $x > 2$ eller $x < -2$.

4.6c $x < -3$.

4.7 Kvadratkomplettera till $x^2 + xy + y^2 = (x + \frac{y}{2})^2 + \frac{3}{4}y^2 \geq 0$, med likhet precis då $x = y = 0$.

4.8 $-2 < x < 3$.

4.9a $x < 0$.

4.9b $x < -2$.

4.9c Alla x.

4.10 $-3 < x < \frac{1}{3}(-4 - \sqrt{19})$ eller $-2 < x < \frac{1}{3}(-4 + \sqrt{19})$.

4.11 $0 < x < 1$ eller $2 < x < 3$ eller $4 < x < 5$ eller $x > 6$.

4.12a 5.

4.12b 5.

4.12c 5.

4.12d 5.

4.12e $|a|$.

4.12f $|a|$.

4.13a $x = \pm 3$.

4.13b $x = 0$.

4.13c Saknar lösning.

4.13d $x = -2$ och $x = 4$.

4.13e $x = -1$ och $x = 2$.

4.13f $x = -2$ och $x = 4$

4.14a Intervallet $[-3, 3]$.

4.14b Intervallen $(-\infty, -2]$ och $[2, \infty)$.

4.14c Intervallet $[-4, 6]$.

4.14d Intervallet $(-5, 1)$.

4.15a $x = -1$.

4.15b $x = -1$ och $x = 1$.

4.15c $x = -1$.

4.16 $x = \pm 2$.

4.17 Kvadrering ger

$$\begin{aligned} & 1 = (\sqrt{x+a} - \sqrt{x-a})^2 = (x+a) - 2\sqrt{x^2-a^2} + (x-a) \\ \Leftrightarrow\ & 2x - 1 = 2\sqrt{x^2-a^2} \\ \Rightarrow\ & 4x^2 - 4x + 1 = (2x-1)^2 = 4(x^2-a^2) = 4x^2 - 4a^2 \\ \Leftrightarrow\ & x = a^2 + \frac{1}{4}, \end{aligned}$$

vilken alltså är ekvationens enda möjliga rot. Kontroll av denna:

$$\begin{aligned} \sqrt{a^2+\frac{1}{4}+a} - \sqrt{a^2+\frac{1}{4}-a} &= \sqrt{\left(a+\frac{1}{2}\right)^2} - \sqrt{\left(a-\frac{1}{2}\right)^2} \\ &= \left|a+\frac{1}{2}\right| - \left|a-\frac{1}{2}\right| = \begin{cases} -\left(a+\frac{1}{2}\right) + \left(a-\frac{1}{2}\right) = -1 & \text{om } a < -\frac{1}{2}; \\ \left(a+\frac{1}{2}\right) + \left(a-\frac{1}{2}\right) = 2a < 1 & \text{om } -\frac{1}{2} \le a < \frac{1}{2}; \\ \left(a+\frac{1}{2}\right) - \left(a-\frac{1}{2}\right) = 1 & \text{om } \frac{1}{2} \le a. \end{cases} \end{aligned}$$

Roten $x = a^2 + \frac{1}{4}$ är alltså äkta precis när $a \ge \frac{1}{2}$. Då $a < \frac{1}{2}$ saknas reella lösningar.

4.18 Substituera $t = \sqrt{x-1}$.

4.19 Kvadrera olikheterna för att avlägsna absolutbeloppen. Addera dem sedan.

5.1 $1110 = 2 \cdot 3 \cdot 5 \cdot 37$.

5.3 $2^2 \cdot 113 \cdot 1117^2$.

5.4 73.

5.5a $\frac{7}{17}$.

5.5b $\frac{29}{37}$.

5.6a Saknar lösning, ty $\text{SGD}(14, -21) = 7 \nmid 333$.

5.6b Vi har $\text{SGD}(114, 303) = 3$, så dela bägge sidor av ekvationen med 3 för att få den ekvivalenta ekvationen $38x + 101y = 56$. Med Euklides algoritm baklänges fås partikulärlösningen $(8, -3)$ till ekvationen $38x + 101y = 1$, och partikulärlösningen $(x_0, y_0) = (8 \cdot 56, -3 \cdot 56)$ till $38x + 101y = 56$. Den allmänna lösningen är $(x, y) = (8 \cdot 56 - 101n, -3 \cdot 56 + 38n)$ för godtyckligt heltal n.

5.6c $(x, y) = (16 - 19n, -36 + 43n)$ för godtyckligt heltal n.

5.7 Om antalet semlor kallas x och antalet mandelkakor y, så gäller alltså $17x + 6y = 250$. Denna ekvation har den allmänna lösningen $(x, y) = (250 \cdot (-1) + 6n, 250 \cdot 3 - 17n)$. Att både x och y skall vara icke-negativa innebär $-250 + 6n \geq 0$ och $750 - 17n \geq 0$, vilket reduceras till $\frac{750}{17} \geq n \geq \frac{125}{3}$. Detta ger möjligheterna $n = 42, 43, 44$. Av dessa är det endast $n = 43$ som ger ett udda y. Svaret är $x = 8$ och $y = 19$.

5.8 Skriv $k^2 + k = k(k + 1)$. Precis ett av talen k och $k + 1$ är jämnt och därmed är produkten $k(k + 1)$ jämn.

Alternativt: Om k är jämnt kan det skrivas $k = 2a$, och $k^2 + k = 4a^2 + 2a$ är uppenbarligen är jämnt. Om k udda kan det skrivas $k = 2b + 1$ och $k^2 + k = 4b^2 + 6b + 2$ är jämnt.

5.9 Skriv $n = 2k + 1$ och använd föregående uppgift.

5.11a Att n är sammansatt betyder att $n = ab$, där $1 < a, b < n$. Nu kan inte både $a > \sqrt{n}$ och $b > \sqrt{n}$.

5.11b Följer omedelbart ur (a).

5.11d Det räcker att pröva delbarhet upp till $\sqrt{8521} < 93$. Om du kan multiplikationstabellen, och därmed vet vilka tal under 100 som är primtal, så räcker det att pröva delbarhet med de 24 primtal, som understiger 93.

5.12 Summan av delarna till $16p$ (utom $16p$ själv) är

$$1 + 2 + 4 + 8 + 16 + p + 2p + 4p + 8p = 31 + 15p.$$

Om detta skall vara $16p$, så är $p = 31$, vilket mycket riktigt är ett primtal.

5.13 Nej.

6.1 $3^{100} - 1 = 81^{25} - 1 \equiv 1^{25} - 1 = 0 \pmod{16}$.

6.2 2.

6.3a Räkna ut resten modulo 10. Svaret är 1.

6.3b 1.

6.4a Ja.

6.4b Nej, kontrollsiffran borde vara X.

6.4c 4.

6.5 Det finns exakt tre möjligheter, som behöver undersökas: $n \equiv 0, 1, 2 \pmod{3}$.

6.6b Varje tal är modulo 7 kongruent med något av talen $0, 1, 2, 3, 4, 5, 6$.

6.7 $g = 3$ och $g = 5$.

6.8 Alla där $a + b + c$ är delbart med 3.

6.9 Ja.

7.1a $3 + 5i$.

7.1b $3 + i$.

7.1c $1 + 5i$.

7.1d $-4 + 7i$.

7.1e $-5 + 12i$.

7.1f $110 + 74i$.

7.1g -4.

7.2a $\frac{1}{2} - \frac{i}{2}$.

7.2b $\frac{2}{29} + \frac{5i}{29}$.

7.2c i.

7.2d $\frac{-17}{26} + \frac{19i}{26}$,

7.2e $-i$.

7.2f $\frac{-i}{2}$.

7.3a $\operatorname{Re} z = 2$ och $\operatorname{Im} z = 3$.

7.3b $\operatorname{Re} z = 2$ och $\operatorname{Im} z = -3$.

7.3c $\operatorname{Re} z = 2$ och $\operatorname{Im} z = 0$.

7.3d $\operatorname{Re} z = 0$ och $\operatorname{Im} z = 5$.

7.3e $\operatorname{Re} z = 0$ och $\operatorname{Im} z = -1$.

7.4a $2 - 3i$.

7.4b $1 + 3i$.

7.4c $-3i$.

7.4d 13.

7.4e 5.

7.4f 5.

7.4g 4.

7.4h 7.

7.4i 1.

7.5a $2\sqrt{205}$.

7.5b $\sqrt{\frac{5}{2}}$.

7.6a Ansätt $z = a + bi$. Svaret är $z = -1 + 3i$.

7.6b $z = 1 + i$.

7.6c Ekvationen saknar lösningar, eftersom vänsterledet alltid är reellt.

7.7a En vertikal linje.

7.7b En horisontell linje.

7.7c Övre halvplanet.

7.7d Det blir linjen av de $z = x + iy$, som uppfyller $y = 2 - x$.

7.7e Imaginära axeln $\operatorname{Re} z = 0$.

7.7f Reella axeln $\operatorname{Im} z = 0$.

7.8a $z = -2 - 3i$ och $z = -1 - i$.

7.8b $z = -3i$ och $z = -1 - i$.

7.8c $z = -1 + 3i$ och $z = -1 - i$.

7.9 Rektangulär form $z = a + bi$ ger

$$z + \frac{1}{z} = (a+bi) + \frac{1}{a+bi} = (a+bi) + \frac{a-bi}{(a+bi)(a-bi)}$$
$$= \frac{(a+bi)(a^2+b^2)}{a^2+b^2} + \frac{a-bi}{a^2+b^2} = \frac{(a^3+ab^2+a)+(a^2b+b^3-b)i}{a^2+b^2}$$

Uttrycket är reellt då

$$0 = a^2b + b^3 - b = b(a^2+b^2-1) \qquad \Leftrightarrow \qquad b = 0 \quad \text{eller} \quad a^2+b^2 = 1,$$

alltså då z är reellt eller $|z| = 1$.

Uttrycket är rent imaginärt då

$$0 = a^3 + ab^2 + a = a(a^2+b^2+1) \qquad \Leftrightarrow \qquad a = 0,$$

alltså då z är rent imaginärt.

Alternativ lösning. Polär form $z = re^{\theta i}$ ger

$$z + \frac{1}{z} = re^{\theta i} + \frac{1}{re^{\theta i}} = re^{\theta i} + \frac{1}{r}e^{-\theta i}$$
$$= r(\cos\theta + i\sin\theta) + \frac{1}{r}(\cos\theta - i\sin\theta) = \left(r + \frac{1}{r}\right)\cos\theta + i\left(r - \frac{1}{r}\right)\sin\theta.$$

Uttrycket är reellt då

$$r - \frac{1}{r} = 0 \quad \text{eller} \quad \sin\theta = 0 \qquad \Leftrightarrow \qquad r = 1 \quad \text{eller} \quad \theta = 0, \pi,$$

alltså då $|z| = 1$ eller z är reellt.

Uttrycket är rent imaginärt då

$$r + \frac{1}{r} = 0 \quad \text{eller} \quad \cos\theta = 0 \qquad \Leftrightarrow \qquad \theta = \pm\frac{\pi}{2},$$

alltså då z är rent imaginärt.

7.10 Ansätt $z = ci$ och $p = a + bi$, där $a, b, c \in \mathbb{R}$. Då fås

$$\left|\frac{z-p}{z+\overline{p}}\right| = \left|\frac{ci-(a+bi)}{ci+(a-bi)}\right| = \left|\frac{-a-(b-c)i}{a-(b-c)i}\right| = \frac{\sqrt{a^2+(b-c)^2}}{\sqrt{a^2+(b-c)^2}} = 1.$$

Alternativ lösning. Emedan z är rent imaginärt, gäller det att $\overline{z} = -z$, varav följer att

$$|z+\overline{p}| = |\overline{z+\overline{p}}| = |\overline{z}+p| = |-z+p| = |z-p|,$$

vilket är ekvivalent med det givna påståendet, förutsatt att $z \neq p$.

Alternativ lösning. Ekvationen är, förutsatt att $z \neq p$, ekvivalent med

$$|z-p| = |z+\overline{p}|.$$

Eftersom z ligger på imaginära axeln, har detta tal samma avstånd till de bägge talen p och $-\overline{p}$, vilka ju är varandras spegelbilder i imaginära axeln. Detta är ekvationens geometriska tolkning.

8.1a $1+i$.

8.1b -1.

8.1c $1+i$.

8.1d i.

8.1e 1.

8.1f $\frac{1}{2}-\frac{i}{2}$.

8.1g 1.

8.2a $17e^{0i}$.

8.2b $11e^{\pi i}$.

8.2c $1e^{\frac{\pi i}{2}}$.

8.2d $\sqrt{2}e^{\frac{3\pi i}{4}}$.

8.2e $2e^{\frac{2\pi i}{3}}$.

8.2f $2\sqrt{3}e^{\frac{\pi i}{3}}$.

8.3a 1 resp. $\frac{\pi}{8}$

8.3b 1 och θ.

8.4 Övergå till polär form. Svaret är $-\frac{1}{2}-i\frac{\sqrt{3}}{2}$.

8.6 $\sin\alpha\cos\beta=\frac{1}{2}\big(\sin(\alpha+\beta)+\sin(\alpha-\beta)\big)$.

8.7 Avbildningen svarar mot multiplikation med i. Talen 1 och $-3+2i$ övergår till i respektive $-2-3i$, och allmännare övergår $a+bi$ till $-b+ai$.

8.8 Rotationen svarar mot multiplikation med $e^{\frac{5\pi i}{6}}=-\frac{\sqrt{3}}{2}+\frac{i}{2}$. Alltså svarar avbildningen mot multiplikation med $-\sqrt{3}+i$. Talen 1 och $-1+i$ har bilderna $-\sqrt{3}+i$ respektive $\sqrt{3}-1-(\sqrt{3}+1)i$.

8.9a 1.

8.9b i.

8.9c $1+i$.

8.9d -1.

8.9e $e^3(\cos 1-i\sin 1)$.

8.10 Av $e^z=e^xe^{iy}$ ser vi direkt att $|e^z|=|e^x|\cdot|e^{iy}|=e^x\cdot 1=e^x$ och $\arg e^z=y$.

8.11a $z=-3-2i$ och $z=3+2i$.

8.11b $z=-2-i$ och $z=4+3i$.

8.12a $z=\pm\frac{-1+i}{\sqrt{2}}$.

8.12b $z=\pm\sqrt[4]{2}e^{\frac{\pi i}{8}}$.

8.13a $e^{\frac{\pi i}{6}},\quad e^{\frac{5\pi i}{6}},\quad e^{\frac{9\pi i}{6}}$.

8.13b $\sqrt[6]{2}e^{\frac{\pi i}{12}}$, $\sqrt[6]{2}e^{\frac{9\pi i}{12}} = \sqrt[6]{2}e^{\frac{3\pi i}{4}}$, $\sqrt[6]{2}e^{\frac{17\pi i}{12}}$.

8.13c $\sqrt[5]{4}e^{\frac{\pi i}{10}}$, $\sqrt[5]{4}e^{\frac{5\pi i}{10}} = \sqrt[5]{4}i$, $\sqrt[5]{4}e^{\frac{9\pi i}{10}}$, $\sqrt[5]{4}e^{\frac{13\pi i}{10}}$, $\sqrt[5]{4}e^{\frac{17\pi i}{10}}$.

9.1a Kvot $x^2 + 3x - 3$, rest $-4x^2 + 2x + 2$.

9.1b Kvot $x^5 + x^4 + x^3 + x^2 + x + 1$, rest 0.

9.1c Kvot $x^2 - 2x + 3$, rest $-2x + 10$,

9.1d Kvot 0, rest $x^2 + 4x + 5$.

9.2 Skriv $p(x) = k(x)(x-1)(x-2) + ax + b$ enligt Divisionsalgoritmen. Vi har $1 = p(1) = a + b$ och $2 = p(2) = 2a + b$ enligt Restsatsen.

9.3a $\frac{1}{x-2} - \frac{1}{x-1}$.

9.3b $\frac{5}{x-2} - \frac{4}{x-1}$.

9.3c $\frac{3+\sqrt{2}}{2\sqrt{2}(x-\sqrt{2})} + \frac{-3+\sqrt{2}}{2\sqrt{2}(x+\sqrt{2})}$.

9.3d $\frac{1}{2(x-2)} - \frac{1}{x-1} + \frac{1}{2x}$.

9.4a $\frac{1}{x} - \frac{x}{1+x^2}$.

9.4b $\frac{1}{x^2+4} - \frac{1}{x^2+5}$.

9.5 Alla polynom på formen $p(x) = x^n$.

10.1 Enligt Faktorsatsen räcker det att $p(1) = -20 + a = 0$, alltså $a = 20$. Faktoriseringen blir $(x-1)(x+4)(x-5)$.

10.2 Enligt Faktorsatsen delar $x - 1$ polynomet. Genomför divisionen och fortsätt sedan att kontrollera, om kvoten är delbar med $x-1$, återigen med hjälp av Faktorsatsen. Upprepa samma procedur. Resultatet blir $p(x) = (x-1)^3(x^2+3x+6)$. (Sista faktorn har inga reella nollställen.) Multipliciteten är alltså 3.

10.3 Sätt $t = z^3$. Då blir ekvationen $t^2 - 2t + 2 = 0$ med lösningarna $t = 1 \pm i = \sqrt{2}e^{\pm\frac{\pi i}{4}}$. Sedan löser man $z^3 = t$ och får de sex rötterna $\sqrt[6]{2}e^{\frac{\pi i}{12}}$, $\sqrt[6]{2}e^{\frac{7\pi i}{12}}$, $\sqrt[6]{2}e^{\frac{3\pi i}{4}}$, $\sqrt[6]{2}e^{\frac{5\pi i}{4}}$, $\sqrt[6]{2}e^{\frac{17\pi i}{12}}$, $\sqrt[6]{2}e^{\frac{23\pi i}{12}}$.

10.4 Sätt roten till ir. Då är också $-ir$ en rot och polynomet $p(z)$ är delbart med $(z-ir)(z+ir) = z^2 + r^2$. Utföres division fås resten $(-9 + r^2)z - 18 - r^2(7 - r^2)$. För att denna skall vara noll, måste $r = \pm 3$, och räkningarna ger att $p(z) = (z^2 + r^2)(z^2 - z - 2)$. De resterande rötterna är alltså rötter till $z^2 - z - 2 = 0$ och är $-1, 2$.

Ett alternativ är att konstatera att om polynomet betecknas med $p(z)$ så är

$$p(ri) = (r^4 - 7r^2 - 18) + (r^3 - 9r)i.$$

Eftersom r är reellt måste därmed $r^4 - 7r^2 - 18 = r^3 - 9r = 0$, vilket är ekvivalent med att $r = \pm 3$. Detta ger oss att $z^2 + 9$ delar $p(z)$. Med hjälp av polynomdivision får vi sedan de 2 övriga rötterna, som ovan.

10.5 Ansätt en lösning $z = 1 + yi$ och räkna på. Svaret är $1 + \pm i, \pm\sqrt{7}i$.

10.6 Utför divisionen av $p(x)$ med $x^2 + 2x + 2$. För att resten skall vara 0, ser man att $a = 2$, och nollställena till de två andragradsfaktorer man fått fram är $\pm i, -1 \pm i$.

10.7 Inget arbete med rätt sats: $z = 2 \pm i, -1 \pm 2i$.

10.8 T.ex. $(z - 2 + i)(z - 2 - i)(z - i)^2(z + i)^2 = z^6 - 4z^5 + 7z^4 - 8z^3 + 11z^2 - 4z + 5$.

10.9a $x^2 - 4 = (x-2)(x+2)$.

10.9b $x^2 + 2x + 1 = (x+1)^2$.

10.9c $x^3 - x = x(x-1)(x+1)$.

10.9d $x^3 - 3x + 2 = (x-1)^2(x+2)$.

10.9e $2 - x - x^2 = -(x-1)(x+2)$.

10.9f $x^4 + 27x = x(x+3)(x^2 - 3x + 9)$.

10.9g $x^6 - 64 = (x-2)(x+2)(x^2 - 2x + 4)(x^2 + 2x + 4)$.

10.10 Det rationella nollstället är 1. Faktoriseringen är $p(x) = (x-1)(x^2-2x+2)(x^2+2x+2)$.

10.11 Lös ekvationen $p(x) = 0$ genom övergång till polär form. Para sedan ihop konjugerade faktorer (rötter) till reella, till faktoriseringen $p(x) = (x-\sqrt{2})(x+\sqrt{2})(x^2-\sqrt{2}x+2)(x^2+\sqrt{2}x+2)$.

10.12a Rötterna är $z = 2, 5$ med summa 7 och produkt 10.

10.12b Rötterna är $z = 2, 5$ med summa 7 och produkt 10.

10.12c Röttera är $z = 0, 2, 5$ med summa 7 och produkt 0.

10.13 Summan är $-3 + i$, produkten $-e$.

10.14 21. (Vilken grad har polynomekvationen? Vad säger Algebrans fundamentalsats?)

10.15 Rötterna är 3, -2 och $\pm\left(\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\right)$.

10.17 Enligt Rationella rotsatsen må en rationell rot $\frac{a}{b}$ (slutförkortad) uppfylla $a \mid q$ och $b \mid 1$. Enda möjligheterna är $\frac{a}{b} = \pm 1, \pm q$. Prövning av dessa ger

$$\begin{aligned}
(x = -q) \qquad & (-q)^3 + p(-q) = q \quad \Leftrightarrow \quad q^2 + p = -1 \\
(x = q) \qquad & q^3 + pq = q \quad \Leftrightarrow \quad q^2 + p = 1 \\
(x = -1) \qquad & (-1)^3 + p(-1) = q \quad \Leftrightarrow \quad 1 + p = -q \\
(x = 1) \qquad & 1^3 + p \cdot 1 = q \quad \Leftrightarrow \quad 1 + p = q.
\end{aligned}$$

De tre första ekvationerna är omöjliga för positiva heltal p och q. Den fjärde ekvationen löses däremot av $p = 2$ och $q = 3$, men inga andra primtal.

Tredjegradsekvationen saknar alltså rationella rötter, utom i fallet $x^3+2x = 3$, då det existerar en rationell rot $x = 1$.

11.1a

$$a_1 = 1 = \frac{1}{\sqrt{1}}, \quad a_2 = \frac{1}{\sqrt{2}}, \quad a_3 = \frac{1}{\sqrt{3}}, \quad a_4 = \frac{1}{2} = \frac{1}{\sqrt{4}}, \quad a_5 = \frac{1}{\sqrt{5}}$$

11.1b Formeln är $a_n = \frac{1}{\sqrt{n}}$, vilken stämmer för $1 \leq n \leq 5$ enligt ovan.

Antag, att $a_n = \frac{1}{\sqrt{n}}$ gäller för något n. Då gäller ävenledes

$$a_{n+1} = \frac{1}{\sqrt{1 + \frac{1}{a_n^2}}} = \frac{1}{\sqrt{1 + \frac{1}{\left(\frac{1}{\sqrt{n}}\right)^2}}} = \frac{1}{\sqrt{1+n}}.$$

Enligt Induktionsprincipen gäller formeln för alla n.

11.4 För $n = 1$ är $VL = 2 \cdot 6 = 12 = \frac{1}{6} \cdot 9 \cdot 8 = HL$. Antag att påståendet gäller för n, så att

$$2 \cdot 6 + 3 \cdot 7 + 4 \cdot 8 + \cdots + (n+1)(n+5) = \frac{n}{6}(n+7)(2n+7).$$

Då gäller påståendet även för $n+1$, ty

$$\begin{aligned} 2 \cdot 6 + 3 \cdot 7 + 4 \cdot 8 + \cdots + (n+1)(n+5) + (n+2)(n+6) \\ = \frac{n}{6}(n+7)(2n+7) + (n^2+8n+12) = \frac{1}{6}(2n^3+27n^2+97n+72) \\ = \frac{n+1}{6}(n+8)(2n+9). \end{aligned}$$

Enligt Induktionsprincipen gäller påståendet för alla n.

11.5 Bassteget: För $n = 1$ är $VL = \frac{1}{1 \cdot 2} = \frac{1}{2} = HL$.
Induktionssteget: Antag att påståendet är sant för $n = k$, d.v.s. att

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{k(k+1)} = \frac{k}{k+1}.$$

Vi måste visa att påståendet är sant också för $n = k+1$, d.v.s. att

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{k(k+1)} + \frac{1}{(k+1)(k+2)} = \frac{k+1}{k+2}.$$

Detta följer ur induktionsantagandet enligt

$$\begin{aligned} \left(\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{k(k+1)}\right) + \frac{1}{(k+1)(k+2)} \\ = \frac{k}{k+1} + \frac{1}{(k+1)(k+2)} = \frac{k(k+2)+1}{(k+1)(k+2)} = \frac{(k+1)^2}{(k+1)(k+2)} = \frac{k+1}{k+2}. \end{aligned}$$

Enligt Induktionsprincipen gäller påståendet för alla positiva heltal n.

11.6 Använd formeln $1 + 2 + \cdots + n = \frac{1}{2}n(n+1)$ för en aritmetisk summa, därefter induktion.

11.7 Standardförfarande.

11.8 Formeln $1 + 2 + \cdots + n = \frac{1}{2}n(n+1)$ för en aritmetisk summa kan komma väl till pass även här.

11.11 Boken, som kommer först i alfabetet, kan flyttas längst till vänster i högst $n-1$ byten. Använd nu induktion på de återstående $n-1$ böckerna.

11.13a 3.

11.13b Av $a > 0$ följer $b = \frac{1}{2} + \sqrt{a + \frac{1}{4}} > 1$. I sådant fall kommer

$$a = \left(b - \frac{1}{2}\right)^2 - \frac{1}{4} = b^2 - b$$

att ge det önskade resultatet.

12.1 $30^3 \cdot 10^3 = 27 \cdot 10^6$.

12.2 Precis fem siffror har $9 \cdot 10^4$ olika tal. Högst fem siffror har talen 0 till 99999, d.v.s. 100 000 tal.

12.3 29^5.

12.4 Första siffran måste vara 1 eller 2, de andra kan vara 0, 1 eller 2. Sådana tal finns det $2 \cdot 3^5$ av. De, som inte innehåller en nolla, är 2^6, så svaret blir $2 \cdot 3^5 - 2^6$.

12.5 Antalet ord är

$$\frac{9!}{4!3!} = 9 \cdot 8 \cdot 7 \cdot 5 = 2520.$$

12.6a Det finns $6! = 720$ arrangemang av de sex symbolerna.

12.6b Vi är inte intresserade av de ord, som innehåller sekvenserna SYM eller MYS. Hur många ord innehåller SYM? Klistra ihop dessa tre bokstäver till SYM. Vi har då fyra symboler: SYM, B, O, L, som skall arrangeras. Dessa bildar 4! förbjudna ord. Lika många ord finns det med den förbjudna sekvensen MYS. Svaret är därför $6! - 2 \cdot 4! = 672$.

12.7a $8! = 40\ 320$.

12.7b Den första personen kan vi sätta var som helst. Varje permutation av de övriga sju ger en entydig utplacering runt bordet räknat från den första (sittande) personen, räknat motsols runt bordet. Svaret är därmed $7! = 5040$.

12.8a $\frac{8!}{2!} = 20\ 160$.

12.8b Den enda dubbletten är E. Bestäm först antalet trebokstaviga ord med båda E. Platserna för de två E kan väljas på $\binom{3}{2} = 3$ sätt och den tredje platsen kan fyllas med en av de andra sex bokstäverna. Totalt alltså $3 \cdot 6 = 18$ ord. Orden med högst ett E kan behandlas som om vi skulle välja tre av de sju olika bokstäverna (G, E, O, M, T, R, I) i ordning, vilket går på $7 \cdot 6 \cdot 5 = 210$ sätt. Svaret är alltså $18 + 210 = 228$ ord.

12.9 $\binom{100}{3} = \frac{100 \cdot 99 \cdot 98}{3 \cdot 2 \cdot 1} = 100 \cdot 33 \cdot 49$.

12.10a $\binom{8}{4} = \frac{8 \cdot 7 \cdot 6 \cdot 5}{1 \cdot 2 \cdot 3 \cdot 4} = 70$.

12.10b $8 \cdot 7 \cdot 6 \cdot 5 = 1680$.

12.11 Elsa har sexton godbitar framför sig. Hon skall välja, vilka åtta av dessa, som skall vara chokladbitar (övriga åtta blir automatiskt kolor). Svaret är $\binom{16}{8}$.

12.12 Matcherna, som skall tippas 1, kan väljas på $\binom{13}{7}$ sätt. Sedan finns det sex matcher kvar, och att välja ut fyra av dessa för att där tippa X, kan ske på $\binom{6}{4}$ sätt. De återstående 2 platserna tippas 1. Svaret är $\binom{13}{7}\binom{6}{4}$.

12.13a $\binom{52}{5}$.

12.13b Fyrtalet kan väljas på 13 sätt (13 valörer), och sedan 48 sätt att välja det återstående kortet. Svaret är $13 \cdot 48 = 624$.

12.13c Det finns $\binom{13}{2}$ val av valörer för de 2 paren. Sedan finns det $\binom{4}{2}$ möjliga färger för det ena paret och lika många för det andra. Det femte kortet kan väljas bland de $52 - 8 = 44$ kort som inte har samma valör som paren. Svaret är alltså $44\binom{13}{2}\binom{4}{2}^2$.

12.14a Detta är ett oordnat urval, och det finns

$$\binom{29}{4} = \frac{29 \cdot 28 \cdot 27 \cdot 26}{1 \cdot 2 \cdot 3 \cdot 4} = 23\ 751$$

möjligheter.

12.14b Enligt Multiplikationsprincipen finns det

$$29^4 = 707\ 281$$

olika ord.

12.14c Andra bokstaven kan väljas på 28 sätt och den tredje och fjärde på 29 sätt vardera, varför svaret enligt Multiplikationsprincipen är

$$28 \cdot 29^2 = 23\,548.$$

12.15 Uppenbarligen kan r variera mellan 0 och 9. Om de första r bollarna är röda, finns det $9 - r$ röda bollar kvar att placera på de återstående $15 - r$ platserna. Dessa platser kan väljas på $\binom{15-r}{9-r}$ sätt. Placeringen av de röda bollarna bestämmer sedan följden, eftersom de svarta bollarna måste placeras ut på de platser, där det inte lagts röda. Svaret är alltså $\binom{15-r}{9-r}$.

12.16a Det kan ske på $\binom{11}{6}\binom{11}{5}$ sätt.

12.16b Vi räknar de oönskade utfallen, då högst en bulle får både glasyr och pärlsocker. Om det är noll bullar, kan först glasyren fördelas på $\binom{11}{6}$ sätt och sedan pärlsockret på $\binom{5}{5} = 1$ sätt. Om exakt en bulle får både glasyr och pärlsocker, kan Elsa välja denna bulle först, och sedan garnera de övriga tio bullarna med antingen glasyr eller pärlsocker, alltså $\binom{11}{1}\binom{10}{5}\binom{5}{4}$ möjligheter. Svaret blir alltså $\binom{11}{6}\binom{11}{5} - \binom{11}{6} - \binom{11}{1}\binom{10}{5}\binom{5}{4}$ sätt.

12.17a Totala antalet trerättersmenyer är $14 \cdot 18 \cdot 8 = 2016$.

12.17b Elsa kan välja sina måltider på

$$\binom{14}{3}\binom{18}{4} = \frac{14 \cdot 13 \cdot 12}{3!}\frac{18 \cdot 17 \cdot 16 \cdot 15}{4!} = 1\,113\,840$$

sätt.

12.17c Det finns $\binom{13}{4}$ sätt att välja huvudrätter, om hon helt utesluter vegetariska alternativ, och $5\binom{13}{3}$ sätt att välja, om hon tar exakt en vegetarisk rätt. Inkluderande valet för förrätterna, så kan hon då totalt välja bland

$$\begin{aligned}\binom{14}{3}\left(\binom{13}{4} + 5\binom{13}{3}\right) &= \frac{14 \cdot 13 \cdot 12}{3!}\left(\frac{13 \cdot 12 \cdot 11 \cdot 10}{4!} + 5\frac{13 \cdot 12 \cdot 11}{3!}\right) \\ &= 364(715 + 5 \cdot 286) = 780\,780\end{aligned}$$

menyer.

12.18 Utan restriktioner kan hon välja sina fyra rätter på $\binom{16}{4}$ sätt. En meny, som bara ger henne en enda av såserna, måste bestå antingen uteslutande av Béarnaise-rätter, vilket ger $\binom{8}{4}$ möjligheter, eller helt enkelt av alla fyra majonnäsrätterna. Svaret blir alltså $\binom{16}{4} - \binom{8}{4} - 1$.

12.19 Nyttja Binomialsatsen på $3^n = (2 + 1)^n$.

12.20 Nyttja Binomialsatsen på $(1 - 1)^n$.

12.21 Det här är ett gott exempel på ett problem, där inga satser är till någon större hjälp, utan man behöver använda påhittighet och fantasi. (Utbildningens syfte är ju annars att lära ut, hur man på ren rutin löser skitsvåra problem.)

In alles är det tio personer. De nio personerna exklusive värden har skakat hand med högst åtta personer (eftersom ingen skakar hand med sig själv eller sin partner). Eftersom alla har skakat hand med olika antal människor och det bara finns nio tal mellan noll och åtta, så är alla dessa antal förekommande. Döp personerna, förutom värden, till $P_0, \ldots, P_8$, där P_k är den som hälsat på k personer.

Person P_8 har hälsat på åtta personer, det vill säga alla utom sig själv och sin gemål. Person P_0 har hälsat på ingen, och enda möjligheten är därför, att P_0 *är* gemål till P_8.

Titta nu på person P_7. Denne har hälsat på sju personer, det vill säga alla utom sig själv, sin gemål *och* P_0. Person P_1 har hälsat på en enda person, nämligen P_8, och är alltså en av de tre, som ej hälsat på P_7. Det betyder att P_1 måste vara gift med P_7.

Fortsätter vi detta resonemang, finner vi att P_6 är gift med P_2, och P_5 med P_3. Återstår då P_4, som är värdinnan.

P	Q	R	$Q \wedge R$	$P \vee (Q \wedge R)$
S	S	S	S	S
S	S	F	F	S
S	F	S	F	S
S	F	F	F	S
F	S	S	S	S
F	S	F	F	F
F	F	S	F	F
F	F	F	F	F

TABELL 1: Sanningstabell för $P \vee (Q \wedge R)$.

13.1 $-1 < x \leq 10$.

13.2 Se Tabell 1. Den andra delen görs likadant.

13.7a Sant för alla par (a, b) utom för paren $(0, b)$ där $b \neq 0$.

13.7b Sant enbart om $a = 0$ och $b \neq 0$. Detta påstående är en negation till det föregående.

13.8a "Någon lärobok i matematik eller i logik är läsbar."

13.8b "Någon elev i klassen blev underkänd i minst ett språkämne."

13.8c "Varje elev i klassen blev underkänd i något av språkämnena."

13.8d "Högst sex av sjömännen var sjuka."

13.8e "Minst åtta av sjömännen var sjuka."

14.1 $\emptyset$, $\{1\}$, $\{2\}$, $\{3\}$, $\{1, 2\}$, $\{1, 3\}$, $\{2, 3\}$, $\{1, 2, 3\}$.

14.2

$$A \cup B = \{1, 3, 5, 6, 7, 8, 9, 10\}, \qquad A \cap B = \{7, 9\},$$
$$B \setminus A = \{6, 8, 10\}, \qquad |A| = 5.$$

14.3b Vi har $B \cap C = \{3, 5\}$ och $B \cup C = \{1, 3, 4, 5\}$, och därför

$$A \cup (B \cap C) = \{1, 2, 3, 5\} \qquad \text{och} \qquad (B \cup C) \setminus A = \{4, 5\}.$$

14.4 Utsagan $A \cap B = \emptyset$ gäller precis då $|a - b| < 1$ (vilket även kan skrivas som $b - 1 < a < b + 1$). Utsagan $A \cap B = \{0\}$ gäller precis då $(a, b) = \pm(1, 2)$.

14.5 Låt $A = [0, 2]$, $B = [1, 3]$ och $C = [2, 4]$. Utöver mängderna A, B och C får vi då

$$A \cup B = (B \cap C) \cup A = [0, 3]$$
$$A \cup C = A \cup B \cup C = [0, 4]$$
$$B \cup C = (A \cap B) \cup C = [1, 4]$$
$$A \cap B = (B \cup C) \cap A = [1, 2]$$
$$A \cap C = A \cap B \cap C = \{2\}$$
$$B \cap C = (A \cup B) \cap C = [2, 3].$$

14.6 Först visar vi att

$$A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C).$$

Låt $x \in A \cup (B \cap C)$. Då gäller att $x \in A$ eller att $x \in B \cap C$. Om $x \in A$, så tillhör x både $(A \cup B)$ och $(A \cup C)$, och därför snittet $(A \cup B) \cap (A \cup C)$. Om $x \in B \cap C$, så tillhör x både B och C, alltså både $(A \cup B)$ och $(A \cup C)$, och därför snittet $(A \cup B) \cap (A \cup C)$.

Nu visar vi att

$$A \cup (B \cap C) \supseteq (A \cup B) \cap (A \cup C).$$

Låt $x \in (A \cup B) \cap (A \cup C)$. Då tillhör x både $A \cup B$ och $A \cup C$. En möjlighet är att $x \in A$, och då tillhör förstås x unionen $A \cup (B \cap C)$. Om inte $x \in A$, så måste x tillhöra både B och C, i vilket fall x tillhör snittet $B \cap C$, och ligger även i detta fall i unionen $A \cup (B \cap C)$.

14.7 Låt $A = [0, 2]$ och $B = [1, 2]$. Utöver mängderna A och B får vi då

$$\begin{aligned} A \cup B &= [0, 2] \\ A \cap B &= [1, 2] \\ A^c = (A \cup B)^c &= A^c \cap B^c = (-\infty, 0) \cup (2, \infty) \\ B^c = (A \cap B)^c &= A^c \cup B^c = (-\infty, 1) \cup (2, \infty) \\ A^c \cup B &= (-\infty, 0) \cup [1, \infty) \\ A^c \cap B &= \emptyset \\ A \cup B^c &= \mathbb{R} \\ A \cap B^c &= [0, 1). \end{aligned}$$

14.8 Vi bevisar den första. Först visar vi att

$$A^c \cap B^c \subseteq (A \cup B)^c.$$

Låt $x \in A^c \cap B^c$. Då tillhör x både A^c och B^c, alltså varken A eller B. Följaktligen ligger x inte i unionen $A \cup B$, och ligger därför i komplementet $(A \cup B)^c$.

Nu visar vi att

$$A^c \cap B^c \supseteq (A \cup B)^c.$$

Låt $x \in (A \cup B)^c$. Då tillhör x inte unionen $A \cup B$, och ligger därför inte i vare sig A eller B. Följaktligen ligger x i komplementen A^c och B^c, och därför också i snittet $A^c \cap B^c$.

14.9 Det gäller att

$$1001 = |A \cup B| = |A| + |B| - |A \cap B| = 1000 + |B| - 999,$$

varav $|B| = 1000$.

14.10 Vi utgår ifrån en universalmängd U av 100%, och att de skadade utgjorde mängderna A, B, C och D, beroende på skador. Vi har då $|A| \geq 0.7|U|$, $|B| \geq 0.75|U|$, $|C| \geq 0.8|U|$ och $|D| \geq 0.85|U|$. Enligt Exempel 14.15 har vi

$$\begin{aligned} |A \cap B \cap C \cap D| &\geq |A| + |B| + |C| + |D| - 3|U| \\ &\geq 0.7|U| + 0.75|U| + 0.8|U| + 0.85|U| - 3|U| = 0.1|U|. \end{aligned}$$

Minst 10% av de stridande drabbades alltså av alla fyra skadorna.

14.11 Vi har $|U| = 1000$, $|A| = 816$, $|B| = 725$, $|C| = 645$, $|A \cap B| = 562$, $|A \cap C| = 463$, $|B \cap C| = 470$ samt $|A \cap B \cap C| = 310$. Enligt Principen om inklusion–exklusion är antalet personer, som tycker om något av konfekt, glass eller tårta

$$\begin{aligned} |A \cup B \cup C| &= |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C| \\ &= 816 + 725 + 645 - 562 - 463 - 470 + 310 = 1001, \end{aligned}$$

vilket är orimligt. Totala antalet personer var ju 1000.